

**CONTESTED CONCEPTS AND COMPETING  
CONCEPTIONS**

by

**Mark Edward Criley**

B.S., Symbolic Systems, Stanford University (1991)

M.A., Philosophy, University of Pittsburgh (1995)

Submitted to the Graduate Faculty of  
the Department of Philosophy in partial fulfillment  
of the requirements for the degree of

**Doctor of Philosophy**

University of Pittsburgh

2007

UNIVERSITY OF PITTSBURGH  
DEPARTMENT OF PHILOSOPHY

This dissertation was presented

by

Mark Edward Criley

It was defended on

30 November 2007

and approved by

Anil Gupta, Ph.D., Distinguished Professor of Philosophy

Mark Wilson, Ph.D., Professor of Philosophy

Kenneth Manders, Ph.D., Associate Professor of Philosophy

George Sparling, Ph.D., Professor of Mathematics

Dissertation Director: Anil Gupta, Ph.D., Distinguished Professor of Philosophy

Copyright © by Mark Edward Criley  
2007

## CONTESTED CONCEPTS AND COMPETING CONCEPTIONS

Mark Edward Criley, PhD

University of Pittsburgh, 2007

I explore and defend the distinction between an abstract concept and *conceptions* of that concept—different ways of explicating the content of that concept. In particular, I investigate *contested* concepts: concepts for which there appear to be genuine, principled disputes about which of several *competing* conceptions is the correct one.

Although philosophers (e.g., John Rawls and Ronald Dworkin) and others often employ the concept/conception distinction, it has seldom been the subject of sustained philosophical inquiry. In particular, little attention has been paid to its consequences for philosophy of language and philosophy of mind. This is unfortunate, for if they are adequately to explain certain common and important features of language and thought, theories of content must find a place for the concept/conception framework.

I begin with a presentation of some examples of contested concepts, and then offer four conditions that jointly specify contested concepts and articulate four desiderata for an account of contested concepts. Next, I assess work in this area by W.B. Gallie, Christopher Peacocke, and James Higginbotham, and briefly consider the ramifications of the concept/conception framework for any general account of concepts. After presenting a hypothetical example of an extended conceptual contest, surveying the features of such disputes, I argue that the contested concept phenomenon is theoretically novel with respect to theories of content, in the sense that it resists assimilation to similar, familiar phenomena (e.g., ambiguity, vagueness, or confusion). Finally, I develop a positive ac-

count of contested concepts, arguing that when we attend to the considerations that exert rational force in disputes over contested concepts, we see that the contents of such concepts are intimately connected with the notion of reflective equilibrium. I close with a brief survey of some areas for future research and applications.

## TABLE OF CONTENTS

<b>PREFACE</b> . . . . .	x
<b>1.0 CONTESTED CONCEPTS: AN INTRODUCTION</b> . . . . .	1
1.1 A “too crude” theory of meaning . . . . .	1
1.2 Three target cases . . . . .	4
1.3 Contested concepts: Specification and discussion . . . . .	7
1.4 Desiderata and discussion . . . . .	13
<b>2.0 EARLIER, RELATED PROJECTS</b> . . . . .	18
2.1 W.B. Gallie . . . . .	19
2.1.1 Gallie’s artificial example . . . . .	21
2.1.2 Gallie’s conditions . . . . .	22
2.1.3 Critical discussion . . . . .	26
2.1.3.1 Exemplars and equivocation . . . . .	26
2.1.3.2 The descriptive versus the evaluative . . . . .	33
2.1.3.3 Rational and genuine disputes: The vindictory burden . . . . .	37
2.2 Peacocke, Higginbotham, and “implicit conceptions” . . . . .	42
<b>3.0 WHAT CONCEPTS (AND CONCEPTIONS) JUST MIGHT BE</b> . . . . .	48
3.1 Concepts as clusters of norms . . . . .	49
3.2 Norm-like features of concepts . . . . .	52
3.2.1 Norms may be of various types and origins . . . . .	53

3.2.2	Not all norms admit of crisp and comprehensive verbal formula- tions . . . . .	54
3.2.3	Norms can be indeterminate . . . . .	56
3.2.4	Concept individuation and identification . . . . .	57
3.3	Conceptions . . . . .	60
3.3.1	Conception-relative assessment . . . . .	61
<b>4.0</b>	<b>SOME EXAMPLES: CONCEPTUAL CONTESTS IN ACTION . . . . .</b>	<b>63</b>
4.1	The concept BORBITY . . . . .	64
4.2	The concept SNORKER . . . . .	73
4.3	Discussion . . . . .	78
4.3.1	Conceptual disputes without conceptions . . . . .	80
4.3.2	The benefits of articulating conceptions . . . . .	81
4.3.3	Non-evaluative contested concepts . . . . .	82
<b>5.0</b>	<b>THEORETICAL NOVELTY . . . . .</b>	<b>84</b>
5.1	The equivocation/disambiguation strategy . . . . .	85
5.1.1	An explanation of the strategy . . . . .	85
5.2	Disjunctive concepts . . . . .	90
5.2.1	Explanation of the strategy . . . . .	90
5.2.2	How the strategy might work . . . . .	90
5.2.3	Why the strategy fails . . . . .	92
5.3	Context sensitivity . . . . .	95
5.3.1	Explanation of the strategy . . . . .	95
5.3.2	How the strategy might work . . . . .	96
5.3.3	Why the strategy fails . . . . .	97
5.4	Sorites vagueness . . . . .	99
5.4.1	Explanation of the strategy . . . . .	99
5.4.2	Why the strategy doesn't work for the target cases . . . . .	100
5.5	Merely epistemic indeterminacy . . . . .	104

5.5.1	Explanation of the strategy	104
5.5.2	A case where the strategy works	105
5.5.3	Why the strategy fails for the target cases	109
5.6	Confusion	111
5.6.1	An explanation of the strategy	111
5.6.2	How the strategy might work	112
5.6.3	Why the strategy fails for the target cases	114
5.7	Temporal externalism	115
5.7.1	An explanation of the strategy	115
5.7.2	How the strategy might work	116
5.7.3	Why the strategy fails for the target cases	117
<b>6.0</b>	<b>A POSITIVE ACCOUNT</b>	119
6.1	Rational and genuine disputes	120
6.1.1	The rational force of prior agreement	122
6.1.2	The role of hypothetical cases	126
6.1.3	Another role for general principles	129
6.1.4	Interpersonal agreement	132
6.2	Conceptual contests as exercises In reflective equilibrium	134
6.2.1	Structural similarities	136
6.2.1.1	Levels of generality	136
6.2.1.2	Multiple dimensions of interdependence	137
6.2.1.3	Conflicting judgments and commitments	138
6.2.1.4	Degrees of commitment or “weight”	138
6.2.2	The methods of argument: weighing, retraction, revision	139
6.2.3	Hypothetical cases revisited	143
6.3	Conceptual concepts and reflective equilibrium	144
6.4	Objections and replies	146
6.4.1	Internal objections	147



6.4.1.1	Mistakes in the method of reflective equilibrium? . . . .	147
6.4.1.2	No unique maximum equilibrium point? . . . . .	148
6.4.2	An “external” worry . . . . .	150
<b>7.0</b>	<b>EPILOGUE: CONCLUSION AND FUTURE RESEARCH</b> . . . . .	<b>154</b>
<b>8.0</b>	<b>BIBLIOGRAPHY</b> . . . . .	<b>159</b>

## PREFACE

It is customary in dissertation acknowledgments to credit a number of people without whom the thesis could not have been written. But I suppose that very few of those who offer those thanks have actually submitted those claims to controlled empirical testing. I speak from experience when I say of those I thank below that I could not have written the thesis without them. For long stretches of time, I tried; each time, I failed miserably. I met with success only during those times when I was willing to accept the help and support of the friends, mentors, and loved ones I acknowledge below.

I have had the good fortune of having two outstanding dissertation directors, one after the other, each of whom I greatly admire for his philosophical and pedagogical brilliance.

The first was Joe Camp. Like the projects of the dozens of students before me who worked with Joe over his nearly forty years at Pitt, this project took root and began to grow during many, many late night sessions conducted at Joe's home. As my wife can attest, I would invariably return home from meeting with Joe on what might best be described as a sort of philosophical high: exhilarated; challenged; exhausted from sustained and probing questioning; light-headed from giddy laughter at Joe's apt but outrageous examples; and eager to wake up the next day to prepare for next week's meeting. I am grateful for Joe's hospitality and ceaseless enthusiasm, and for helping through his own example to strengthen my love for philosophy to the point where no other sort of life for me seemed conceivable. This dissertation would not have been possible without Joe, and my greatest regret about this project is that I was unable to complete it under his direction. Those dozens of earlier students who worked under Joe, many of whom traveled from all over to

Pittsburgh to attend the April 2006 “Camp Out!” conference in his honor, know what I mean when I say that I will simply never have another friend or mentor like him. He is irreplaceable.

Nevertheless, upon Joe’s retirement I found myself having to replace the irreplaceable. Only Anil Gupta could have stepped into the breach created by Joe’s departure. Everything that you need to know about Anil’s character is encapsulated in his willingness to commit himself to a year of long distance work directing a Ph.D. student whom at that point he hardly knew. This was the most selfless and generous act anyone outside my family has ever undertaken for my sake. I thank him for all of the lengthy e-mail and phone conversations, and for all that he has done to encourage me, to shape and improve the project, to gently but firmly demand more and better material where it was needed, and to curb my tendency to want to pursue every tangent in every direction all at once. Without him, I would never have been able to finish this work; much of what I hope to do in the next stage of work on this project will attempt to address the questions and concerns that he has raised during our discussions. For all of this I owe Anil a debt that I can never repay.

Charlotte Brown, Ted Morris, and Jamie Tappenden have at various points been so closely involved in this project and my career that they served as de facto dissertation directors. I thank them for their philosophical acumen, guidance, humor, and deep friendship, and for their heartening conviction that I might have something worthwhile to say, even when I was dubious on that count myself.

Thanks also to Ken Manders, Mark Wilson, and George Sparling for their time, advice, and counsel, and for their service on my dissertation committee. I also wish to thank Joe’s late partner, Tamara Horowitz, who was an original member of my committee and who is sorely missed.

Lenny Clapp, Carl Gillett, Doug Lavin, and Ram Neta provided me not only with invaluable friendship but with close and careful feedback at a number of stages of the project. Comments from friends and comrades Janice Dowell, Brian Hill, Lionel Shapiro,

Tommie Shelby, and Matt Weiner on earlier versions of this document were also incredibly helpful.

I would also like to thank the members of my 2006 IWU May Term course on this topic and the audience at my January 2007 talk at IWU, including Victoria Johnson, Kenton Machina, Katherine (KP) Paton, Adam Simon, Larry Stout, Hans-Jörg Tiede, Nick Timme, and Adam Wentland.

For friendship and philosophical conversation I wish to thank Alp Aker, Alison Bailey, John Beer, David Beisecker, Tom Berry, Adam Betz, Donald Bruckner, Arthur Cunningham, Derrick Darby, Kevin Davey, Carl Davulis, David Finkelstein (the Younger), Anton Ford, Aryeh Frankfurter, Karen Frost-Arnold, Mitch Green, Jeremy Heis, Lisa Hoelle, Chris Horvath, Henry Jackman, Emily Kelahan, Ben Laurence, Hans Lottenbach, Ted McNair, John MacFarlane, Eric Marcus, Dale Miller, Carol Moeller, Jennifer Nagel, Doug Patterson, Bill Porter, Steve Puryear, Andrea Rieber, Evan Riley, John Roberts, Laura Ruetsche, Susanna Schellenberg, Lisa Shapiro, Sergio Tenenbaum, Lisa van Alstyne, Anders Weinstein, Joan Wellman, and Herb Wilson.

For friendship and support outside of the philosophical profession, I wish to thank Eli Cates, Jeremy Cantlon, Peter Chang, James and Rachel Cramton, (Granville) Scott Crawford, (William) Scott Crawford, Steve Cummer, Matthew Dresden, Kevin and Joumana Driscoll, Andy Elder, Molly Brown Forstall, Scott Forstall, Thomas J. Hunter, Keith Karkaker, Regina Linsalata, Carole Myscofski, Alison Sainsbury, and John K. Wilson.

Thank you to Kathy Rivet for administrative help, to Alp Aker for being my  $\text{\LaTeX}$  guru, and to Frederico Garcia for authoring the `pittetd`  $\text{\LaTeX}2_{\epsilon}$  class.

The sympathetic ear and support of Jennifer Whiting in my earliest days at Pitt made it possible for me to continue my graduate career. Years later, Dr. Joellen Popma and the members of the University of Pittsburgh Dissertation Support Group played much the same role. Karen Grandstrand Gervais, family friend and the first philosopher I ever knew, was an inspiration to me throughout graduate school, as she has been since I was a small child.

My family members in central and northern Illinois, Michigan, Ohio, and Wisconsin made this work possible. Thank you to my sister, Jennie Criley, for her support, commiseration, and good humor, and to my beloved late grandmother, Florence Johnson.

My parents, Bruce and Norma Criley, were my first and most prized teachers. They have been tremendously supportive, loving, and patient throughout my life, and were there from the beginning to tell me things I did not know.

Above all, I thank my wife, Amy Smith, a philosopher in intellect and temperament, even if not by trade. Amy has been my constant (and incredibly patient) partner for the last thirteen years. I cannot begin to find the words to thank her for her love and support, including her willingness to listen to various half-baked ideas on our walks with Chance and her careful editorial advice on the final version of this document. I am glad to have the rest of our lives together to try to show her the love I cannot express in words.

Finally, this work is dedicated to my son, Saed Jamal Smith, who has been there throughout its creation and who, on the day of its defense, is precisely the same age that I was on the day of his birth: 7035 days old. Among many other reasons, I love him for being precisely the sort of fellow nerd who finds that sort of pointless coincidence worth noting.

## 1.0 CONTESTED CONCEPTS: AN INTRODUCTION

### 1.1 A “TOO CRUDE” THEORY OF MEANING

In a passage in his article “Constitutional Cases,” Ronald Dworkin objects that a certain theory of legal interpretation rests upon a “too crude” theory of meaning—a theory of meaning that “ignores a distinction that philosophers have made but lawyers have not yet appreciated.” (Dworkin, 1977, p. 134). What the suspect theory ignores, according to Dworkin, is the difference between an abstract *concept* and the various, more specific *conceptions* of which that concept admits. And without that distinction in place, a theory of meaning will be too crude for failing to accommodate another, related crucial philosophical advance: the recognition that some such abstract concepts—including, say, the concepts of fairness, equality, and cruelty—are *contested* concepts, which admit of *competing* conceptions.

Upon reading Dworkin’s remark, the philosophical reader might well wonder what she has missed: *Have* philosophers in fact incorporated a distinction between contested concepts and competing conceptions into their “theories of meaning”? Of course Dworkin *himself* discusses the topic, but his remarks add up to a rather cursory but suggestive account that occupies only a few pages of text over the course of his distinguished career.<sup>1</sup> And in fact, very little further philosophical attention has been paid to a purported distinction between concepts and conceptions or to “contested concepts.” There are a few

---

<sup>1</sup>See Dworkin (1977, pp. 103, 134–136, 226–227) and (Dworkin, 1986, pp. 70–71).

noteworthy and distinguished exceptions, which I consider in due course.<sup>2</sup> However, most of these exceptions have, like Dworkin, worked in philosophical subfields other than the philosophy of language or philosophy of mind, and have not had much to say about how such a distinction plays out within a theory of meaning. It is hard to claim on the basis of these few instances that philosophers have made or appreciated the distinction with which Dworkin credits them.

Nevertheless, in another respect Dworkin's remark should strike the philosophical reader as apt. Even though philosophers have seldom had much to say *about* a putative distinction between abstract contested concepts and competing conceptions of those concepts, we actually *employ* the distinction quite regularly, as when we talk about, say, deontological conceptions of rights, reliabilist conceptions of justification, counterfactual conceptions of causation, and so on. The investigation and assessment of competing conceptions of abstract, contested concepts is part of our stock in trade. Although the phenomenon is perhaps featured most prominently in philosophical work and tracked most carefully by philosophical, vocabulary, it is not limited to philosophical discussions. Dworkin surely doesn't mean to suggest that philosophers have "made" the distinction in the sense that they have invented it out of whole cloth for their own parochial, professional purposes. Rather, he clearly means to suggest that philosophers have fastened onto an important distinction that was already exhibited in everyday, ordinary discourse.

One of my central claims in this dissertation is that the core of Dworkin's point is correct: On pain of being "too crude"—that is, on pain of failing adequately to explain certain common and important features of language and thought—theories of linguistic and mental content must find a place for the concept/conception framework and for the related notion of a "contested concept." However, I also am convinced that Dworkin is overly optimistic in his contention that philosophers have appreciated this distinction. Despite its significance for theories of meaning or content, I argue that the framework is theoretically novel with respect to present work in that field. That is, I shall argue

---

<sup>2</sup>The philosophers Dworkin likely has in mind are W.B. Gallie's work on "essentially contested concepts" (Gallie, 1962), and John Rawls's employment of the distinction for the concept of justice in (Rawls, 1971).

that none of the devices that presently come “standard” or even “special order” in the philosopher’s toolkit can adequately account for the phenomena that contested concepts are used to explain; a new tool must be developed and added to the kit. My objective in this thesis is to start that development and point the way toward its completion.

This project should be of interest for several further reasons. First, attention to this topic will provide new material for those who are interested in cases of (at least apparent) indeterminacy and their resolution. An account of contested concepts would provide us with a framework for understanding concept use and development, as well as for understanding the way that these are related to the inquiries, arguments, and objections that often shape and direct them.

Second, a successful explication and defense of the concept/conception distinction would help repay a loan taken out against theories of meaning by those philosophers working in value theory—including Dworkin and John Rawls—who have explicitly relied upon the distinction in their own work. Their fruitful exploitation of that distinction—controversial though the details of their projects might be—constitutes at least a partial vindication of it. Nevertheless, the device ought to be submitted to broader critical scrutiny in its own right. I will argue that much of our philosophical practice and everyday discourse outside of value theory also appeals to the distinction. If I am right, then a successful account would serve not only to explain but to underwrite those uses as well.

Third, and finally, I believe that coming to terms with the concept/conception distinction will recast and clarify a panoply of heated debates over philosophical questions in related areas. Among those I will briefly address in the epilogue to the thesis in Chapter 7 include the topics of conceptual analysis, analytic truth, and the appeal to intuitions and conceivability in arguments for modal claims.

Here is how I will proceed in the remainder of this chapter in order to get the discussion underway. In §1.2, I present some simple but central examples of contested concepts, introducing them within one sort of setting in which their contested character is on display: a dispute over the question whether a concept applies or fails to apply in a particular



concrete case. Then, in §1.3, I draw on features of these examples to develop a preliminary specification of my target in the thesis: I present four criteria that I think jointly characterize contested concepts and pick out the target of my inquiry. I close the chapter in §1.4 with a presentation of four desiderata that I believe any account of that target should strive to satisfy.

## 1.2 THREE TARGET CASES

To fix ideas, it will be helpful to have some very simple examples of contested concepts in place. Here are three concepts—FAIRNESS, CRUELTY, and APOLOGY<sup>3</sup>—that I think are central cases of contested concepts, in the following sense: If any concepts are contested concepts, these are. But rather than merely considering a list of such concepts, we will confront them in a setting in which they prompt the sort of “contests” from which their label derives.

The first example, concerning fairness, is adapted from [Dworkin \(1977, p. 277\)](#); the other concepts have not, to my knowledge, been discussed elsewhere.

---

<sup>3</sup>Throughout the thesis, I will alternate between talking about individual concepts by using small-capitals strings to function as names for them, as I do here, and referring to them as “the concept of . . .”. I extend both of these practices to schematic references to arbitrary concepts e.g., by discussing “the concept of C” or “the concept of F,” C or F stand in for some arbitrary concept. Also, for ease of discussion, I will restrict my consideration in the thesis to concepts that correspond to unary natural language predicates, such as “. . . is fair” or “. . . is cruel.” I believe that I can do this without any loss of generality.

### **Case 1 (FAIRNESS)**

Goofus and Gallant are each afflicted with the same horrible disease. But while Goofus's condition is merely painful, Gallant's borders on fatal. They have only one dose of medicine in their joint possession, with no prospect of getting another. Goofus and Gallant disagree over whether it would be fair for Goofus to take one half of the dose for himself. Now, Goofus admits that it would be noble and fine for him to yield the whole dose to Gallant. But he contends that since neither of them has a prior claim to the medicine, fairness requires nothing more of him than that he split the dose evenly with Gallant. When pressed to defend this contention, Goofus claims that fairness is *the equal allocation of commonly held goods*. Gallant disagrees: given the severity of his illness, fairness requires that Goofus concede to him the entire dose. When pressed to

defend his contention, Gallant says that fairness involves *allocation of commonly held goods in accordance with the needs of the parties*.

### **Case 2 (CRUELTY)**

Goofus and Gallant are discussing a mutual acquaintance named Guy. As they converse, Goofus doodles a terribly unflattering caricature of Guy performing unmentionable acts, along with a caption that calls into question Guy's character, parentage, and sense of fashion. Gallant exclaims that Goofus's act of caricature is cruel. After a moment's reflection, Goofus concedes that he should not have made the drawing, but says that his doing so was not cruel. Why not? Because, Goofus says, cruelty is *the unwarranted infliction of physical or emotional harm*, and Guy, since he was not present, could not have suffered any such harm from the caricature. Gallant demurs. Cruelty, he claims, is broader than that: to be cruel is *to compromise without warrant the dignity of the aggrieved party*, which might, but need not, involve physical or emotional harm.

### **Case 3 (APOLOGY)**

Gallant gives a piano recital; Goofus is in attendance. At the reception afterward, Goofus provokes giggles around the punchbowl by loudly speculating that Gallant might have learned his fingering techniques by observing chimpanzees at the zoo. Gallant demands that Goofus apologize. Goofus replies that while he does not think his remarks were offensive or inappropriate, he nevertheless apologizes for having made them and for having angered Gallant. Gallant hotly denies that what Goofus has offered is an apology. Goofus protests that *to apologize is to utter the phrase "I apologize" (or some equivalent) as an expression of genuine regret for the harm one's action has caused*. Gallant disagrees, on the grounds that *an apology must involve the sincere recognition of the wrongness of the act for which one is apologizing*.

### 1.3 CONTESTED CONCEPTS: SPECIFICATION AND DISCUSSION

With these specimens on the table, I would like to propose four conditions that I think characterize and specify the sort of concepts that I have in mind. I think that these conditions are individually necessary and jointly sufficient for a concept to fit into the category I have in mind, and hope over the course of the thesis to convince the reader that this is so—that is, that the concepts that I have in mind fit naturally into this class. For the moment, however, I put forward these conditions as hypotheses about what is distinctive about contested concepts.

It will be a running theme throughout the thesis that in order to begin to understand contested concepts, we will have to understand the sort of contests or disputes to which they are subject. I propose that we start by asking: What sort of disagreements are involved in the disputes in the target case? Each dispute concerns the question whether the particular case in question is an instance of fairness, cruelty, or apology—whether each case is one to which those terms apply. Loosely put, what is at issue is the question of fit between what we might (very loosely) call the “facts of the case” and what we might (again, very loosely) call the “criteria” for the correct application of the term in question.

But unlike many such disputes about fit between facts and criteria, in the target cases the disagreements over fit do not turn on the question of what the “facts of the case” are. To see the difference, we may compare the target cases with variants in which the particular facts are precisely the subject of the dispute. For instance, compare Case (1) with a variant in which Goofus and Gallant disagree about the precariousness of Gallant’s condition, and compare Case (3) with a variant in which the dispute turns on the question of whether Goofus genuinely felt bad about hurting Gallant’s feelings. In these variant cases, we may suppose that the criteria for the application of the term are fixed and held in common between Goofus and Gallant, and the issue to be settled in the dispute is the question of what the *facts* are, and whether they are such as to fit the *criteria*.

In the target cases, however, we may suppose—indeed, we can just go ahead and

stipulate—that the disputants views about all of the relevant facts are fixed and held in common. The disagreement in the target cases, however, still turns on the question of fit between facts and criteria, but in the target cases the issue that is under dispute is what the correct criteria of application are.

Each of the disputes presented in the target cases, then, turns upon conceptual questions. That is, the dispute is grounded in the parties' divergent views about the criteria imposed by the concept in question: that is, over what properly makes an arrangement count as fair, what makes an act count as cruel, or what makes a speech-act count as an apology.

The parties to these disputes understand themselves to have a single shared concept, which they employ in common and which has a history of employment within their community. The particular concepts under discussion in the target cases are ones with which it is comparatively easy to achieve competence: the very young and very dim can come to grasp these concepts. Concept-users can acquire the capacity say, to recognize certain canonical cases of unfairness, and to make rather basic inferences to and from the judgment that an act is unfair. In the present dispute, each participant treats the other as “possessing” the concept: having a grip upon it and as being capable of making judgments or inferences involving it. That is, neither lodges a charge against the other that he is incompetent with the concept, or that in denying her opponent's claim she has changed the topic or is equivocating.

These observations lead to the first condition that distinguishes a contested concept: that it must be univocal.

**Condition 1 (*The Univocality Condition*)**

A contested concept is a single concept, used univocally across a community over time, rather than a collection of several distinct or shifting concepts.

Note that this condition insists upon both synchronic and diachronic univocality. That is, the concept in question must be one and the same concept employed in common by *different members* of the community and that it must be employed in common members

of the community at *different temporal stages* of the community's concept use.

Of course, deciding whether a conceptual community has one concept associated with a general term rather than several distinct concepts is a thorny question. What makes for sameness of concept or continuity of concept? How are we to individuate concepts? Of course, the answers to these questions will depend upon how we answer the question of what concepts themselves are—a question to which I present a tentative answer only later, in Chapter 3. However, as just a rough indication of what is to follow, I want to suggest here that any answer to the questions of concept individuation and identity will turn out to place a great deal of weight on the views of the concept-users, but it is none the worse for that. Whether we are dealing with a single concept depends, I suppose, in large part upon the community's conceptual behavior with respect to that concept, as well as their own explicit or implicit commitments to univocality.

Nevertheless, questions of concept-identity cannot be resolved simply by appeal to the explicit pronouncements of community members that they are employing one and the same concept in common: The members of a community might turn out to be wrong on this count. Despite a history of overlapping use and despite a commitment to reasoning with a single concept, a conceptual community will sometimes find their claims of univocality unsustainable.

The concepts that are under discussion in the target cases—FAIRNESS, CRUELTY, and APOLOGY—are, as Dworkin suggests, abstract. By this I mean that the concepts impose only limited explicit constraints on the correct employment of their corresponding terms. Members of our own conceptual community readily employ these concepts and rest their arguments and appeals and analyses upon them, but usually with only a limited or hazy view of what those concepts involve. Without some further spelling-out—without some further specification of what fairness or cruelty or apology consists in—these concepts are thin and provide a great deal of room for disagreement and discord. This provides us with the second condition in our specification of contested concepts.

***Condition 2 (The Indeterminacy Condition)***

<sup>1</sup> The concept C is abstract or indeterminate, in the following sense: There are some

cases for which persons who are competent with respect to that concept, and who suffer from no epistemic or linguistic deficiency, may reasonably disagree over whether the concept applies or fails to apply.

This condition recognizes the fact that disputes of the sort canvassed above are different in kind from, say, disputes in which Goofus and Gallant disagree about whether it is correct to call flame-retardant objects “flammable” or where they disagree about whether the term “fruit” applies to tomatoes. Such cases are ones where the dispute can easily be resolved by appealing to settled semantic or biological facts in a way that recourse to any expert or encyclopedia could easily remedy. There don’t appear to be any such facts to appeal to in the controversies above that could provide such a resolution.

However, it is clear that the disputes in the target cases above emerge precisely because Goofus and Gallant each subscribes to a rather more detailed view about what fairness, cruelty, and apology each involves. That is, they subscribe to different precisifications or spellings-out of the concepts in question—what we shall call different conceptions of fairness, cruelty, and apology. This observation underpins the next condition in the specification:

**Condition 3 (*The Conception Condition*)**

There are a number of distinct, plausible, but incompatible ways of spelling out the concept in question—different ways of making its content thicker, more precise, or more determinate—that is, different ways of reducing the range of cases described in the Indeterminacy Condition above. Call these ways of spelling out the concept *conceptions* of C.

It is important in what follows that the reader keeps in mind that we must distinguish the use of the word “conception” here from some other, closely related senses in which the word is sometimes used. On occasion, we use the term “conception of F” to refer to someone’s stereotypes, connotations, or set of beliefs concerning F. This is the sense in which the term is used when Woody Allen reports that his “conception of Canada” is of

a cold and blustery country where criminals are simply too cold to practice their craft, or that when Polly reports that her “conception of a snake” is of a slimy, poisonous creature. A conception, as I intend it here is an answer—perhaps only a partial answer—to the constitutive conceptual question, What is it to be an F? That is, a particular conception is an attempt to specify what an object must be like in order properly to qualify as an F, what must follow from the claim that an object is an F, and similar questions with that sort of conceptual character.

Conceptions can be articulated at different levels of specificity. The target cases in §1.2 above attribute to Goofus and Gallant uncommon facility in providing linguistic formulations of their respective conceptions. For ease of discussion and to take us through the details of the dispute more quickly, the cases we will consider will principally involve disputants who are similarly adept. But disputes of this sort can take place even without that level of semantic self-awareness, and without explicitly invoking the concept/conception terminology: Nothing crucial depends upon it. Disputes over contested concepts can take place—if more awkwardly and slowly—even when the participants are unable to articulate particular conceptions. For instance, if in case (2) Goofus had merely defended himself against Gallant’s charge of cruelty by claiming that “Guy wouldn’t ever know” about the caricature, or by insisting that the drawing “hadn’t hurt anybody,” we might well understand that as an appeal to the conception he so deftly articulates in the example as it is described. Disputants often work their way to conceptions only in this gradual, piece-by-piece manner. (I return to this point below, in §4.3.1.)

A particular conception may itself in turn involve another contested concept. For instance, in example (2), Gallant’s conception of cruelty invokes the concept DIGNITY, which might well turn out to be a contested concept in its own right.

Thus, for the kind of concepts I want to discuss, there will be several plausible or what we might call “admissible” conceptions of a concept. Roughly, these are ones that a concept-user might adopt without being liable to a charge of semantic incompetence. But what is distinctive about contested concepts is that the disputes over them are oriented



around the question of which among these admissible conceptions provides the correct way of spelling out the concept in question. Hence our final condition:

**Condition 4 (*The Competition Condition*)**

The conceptions of the concept C are properly understood to be in competition with one another for status as the *proper*, *correct*, or *best* conception of C. That is, each is regarded as an attempt to explicate what thought and talk involving the concept has “been about” all along.

I contend that a concept may be contested even before the different conceptions are appreciated or made explicit. This requires some more defense and elaboration, but the gist of the idea is that the availability of hypothetical cases over which reasonable persons without epistemic or linguistic deficiency might disagree is sufficient to show that the concept is contested in the sense under discussion here. Note that the appeal to the continuity of use in this condition reinforces the insistence on diachronic univocality. A competing conception is not offered as an attempt to stipulate a new usage or to invoke a change in meaning. Rather, it is to be understood as continuous with prior employment of the concept, and as an attempt to explicate that concept.

A dispute over competing conceptions is not of the sort that is liable to satisfactory resolution by, say, a linguist’s survey of native speakers’ reports of whether they could call caricaturing someone in absentia “cruel.” It doesn’t seem as if a survey of the linguistic intuitions of native speakers should resolve the matter to anyone’s satisfaction. For I want to contend that, curiously, it seems that the apt thing to say about this sort of case is that even though the bulk of the conceptual community might articulate or embrace a conception of, say, cruelty, that does not flout any semantic norms or run counter to the judgments of any cruelty experts (for there aren’t any), that the bulk of the linguistic community might nevertheless be wrong in the fine details of its view about what “cruelty” means.

## 1.4 DESIDERATA AND DISCUSSION

The four criteria presented in the previous section specify the target of inquiry for the thesis, and they begin that inquiry by present a preliminary account of the distinctive features of contested concepts. In the remainder of the thesis, I will be elaborating and developing my account. But before embarking on that task, I would like to proffer four desiderata that it seems any account of contested concepts should try to meet.

(A) *An account of contested concepts ought to explain why there are contested concepts, and explain the conditions under which they emerge.*

I take it that contested concepts are not a necessary or ineliminable feature of a conceptual scheme: There is nothing incoherent in the idea of a community of concept users none of whose concepts are contested. We can even (very roughly) describe the adjustments that a community that presently employs contested concepts would have to make in their practices in order to eliminate them from their use. What would be required would be a lot of subscripting, or some equivalent disambiguation technique, in order to distinguish between distinct concepts when disputes arise. An account of contested concepts, then, ought to survey the conditions under which contested concepts arise, and explain how those conditions give rise to them. It would also redound to the credit of an account if it not only explained but justified communities' employment of contested concepts, perhaps by distinguishing some of the advantages that accrue to a community in virtue of the employment of contested concepts, for instance by highlighting things that a conceptual community would be able to do only if they have concepts that fit those four criteria.

(B) *An account of contested concepts and competing conceptions ought to explain the connection between contested concepts and the rest of the theoretical machinery of philosophy of language and mind.*

This desideratum is extremely broad, and points us toward more questions than I can hope to answer within the space of the present work. However, there are some pressing

questions raised under this heading that I will address. Does the phenomenon of contested concepts have anything tell us about the nature of concepts in general? What are the ramifications of the phenomenon for other theories of linguistic and mental content? What is the relationship between the concept/conception distinction and other curious semantic phenomena? In particular, why shouldn't we think that the tools that have been developed in the philosophy of language and philosophy of mind to accommodate other, similar phenomena will be able to accommodate these sorts of cases without any further theoretical innovation?

(C) *An account of contested concepts and competing conceptions ought to explain the distinctive features of disputes about contested concepts, and the connection between these features and the content of those concepts.*

In order to understand contested concepts, we will have to pay attention to the kinds of considerations that may be presented as having rational force or appeal during the course of the disputes over them. For part of what is troubling about conceptual disputes of this sort is that, despite the vehemence one tends to bring to such disputes and despite one's conviction that one's conception is correct, it is sometimes unclear (especially to an "outsider" to the dispute: someone who does not have a stake in the debate or any strong convictions concerning the conceptual issues at stake) why anyone should think that there is a fact of the matter about which party (if either!) is correct about what fairness, cruelty, or apology consists in.

Nevertheless, participants in the dispute regularly understand themselves as having a genuine disagreement, and one that is subject to the force of the better reasons. In order to take contested concepts seriously, I think that an account ought to do its best to credit the participants' convictions on this count. An account ought to try to explain the sense in which these disputes are instances of genuine disagreement, liable to resolution through principled argumentation rather than simply competing stipulations.

Finally, an account of contested concepts ought to recognize and try to explain an apparent paradox that emerges from the four conditions provided in the specification in the previous section.

On the one hand, contested concepts seem, in virtue of their abstract character, to tolerate a great deal of pluralism about their content. Indeed, the Indeterminacy Condition tells us that competent concept users who lack no relevant empirical information concerning a specific case may reasonably disagree about whether the case falls or fails to fall under the concept. This pluralism is further exhibited in the Conception Condition, which tells us that there are a variety of different plausible ways of spelling out concept—a number of what we called “admissible” conceptions. The Competition Condition appears to preclude that pluralism, since it contends that the competing conceptions are properly understood as candidates for the status of the best, proper, or correct conception of the corresponding concept—as an attempt to correctly specify the content of that concept as it has been employed within the community.

Thus, an apparent paradox emerges from the four conditions I have articulated above: The Indeterminacy and Conception Conditions appear to preclude the satisfaction of the Competition Condition, and vice versa. A successful account of contested concepts will have to address and resolve this apparent paradox. Hence, our final desideratum:

(D) *An account of contested concepts ought to resolve the apparent paradox generated by the seemingly incompatible demands of the Indeterminacy and Conception Conditions on the one hand and the Competition Condition on the other.*

Developing an account of contested concepts that satisfies these desiderata will be no easy task, and I do not claim to have fully discharged it in the present work. I intend the thesis in part as an advertisement for the importance of the phenomenon and the merits of having such an account, and in part as a first stage in the development of such an account. Here, briefly, is the road-map I will follow in the remainder of the thesis in order to begin to bring this aim to fruition.

In Chapter 2, I present a discussion of some previous forays into this area. §2.1 presents an extended discussion of the first and, to date, most extensive examination of contested concepts: W.B. Gallie's "Essentially contested concepts" (Gallie, 1962). §2.2 turns to the relatively recent work of Christopher Peacocke and James Higginbotham on the distinction between concepts and implicit conceptions.

In Chapter 3 I consider the ramifications of my preliminary specification for theories of linguistic and mental content in general and for an account of concepts in particular, with the aim of contributing to the satisfaction of desideratum (B).

Chapter 4 presents some extended hypothetical examples of disputes over contested concepts, and displays in a condensed form the sort of considerations that may be brought to bear in those disputes. The aim here is to contribute to the satisfaction of desideratum (C).

Chapter 5 consists of a sustained argument by elimination designed to show that the contested concept and competing conception framework is theoretically novel to theories of content in the philosophy of language and philosophy of mind, in the sense that the phenomena it is invoked to explain cannot be adequately accounted for by other frameworks that are currently on offer. I consider seven such alternative strategies and for each explain why it is inadequate to the task. This chapter is not strictly negative, however. Considering the ways in which these other approaches fall short of explaining contested concepts will shed further light on the distinctive characteristics of those concepts, helping us to satisfy desideratum (B).

In Chapter 6, I draw together the lessons and develop a more detailed, though still tentative, account of contested concepts. In that chapter, I draw some lessons out of our earlier discussion of the characteristics of disputes over contested concepts, and I suggest that the distinctive semantic characteristics of these concepts are due to the characteristics of those disputes. In particular, I argue that the account snaps into place, and the vexing apparent paradox from desideratum (D) may be resolved, once we understand the content of contested concepts as subject to the demands of the method of reflective equilibrium.

Chapter 7 closes the thesis with a brief summary and a brief discussion of the ways in which I foresee the concept/conception distinction can be employed to make progress on a number of other philosophical investigations.

## 2.0 EARLIER, RELATED PROJECTS

In the previous chapter, I mentioned that while many philosophers make use of the notion of contested concepts and the distinction between a more general, abstract concept on the one hand and more specific conceptions of that concept on the other, few philosophers have actually turned their critical scrutiny on these notions in order to provide a systematic treatment of them. In this chapter, I turn to consideration of two philosophical research programs that have . In §2.1, I consider the work, briefly mentioned above, by W.B. Gallie on the notion of “essentially contested concepts.” In §2.2, I consider the work of Christopher Peacocke and James Higginbotham in developing the notion of “implicit conceptions” of a concept. My aim in each of these discussions is to acknowledge these philosophers’ contributions and consider the ways in which understanding their projects can deepen our understanding of one developed here. However, in each case, we will see that much work remains to be done in order to carry out the project of the thesis. As such, this Chapter is largely oriented around desideratum (B)—explaining the connection between contested concepts and other issues in philosophy of mind and language. However, the discussion also draws lessons concerning desideratum (C) out of Gallie’s attempt to show that disputes of the kind we are interested in satisfy certain standards of genuineness and rationality.

## 2.1 W.B. GALLIE

The first philosopher to use the phrase “contested concept” was W.B. Gallie, in his 1956 article “Essentially contested concepts.”<sup>1</sup> Gallie’s article is explicitly cited by Dworkin (Dworkin, 1977, p. 103) in his own discussion of the contested concept and competing conception phenomenon. Gallie’s work is often cited as an influence on H.L.A. Hart’s treatment of justice in Chapter VIII of Hart (1998). No doubt the most widely known employment of the distinction is that of John Rawls in Rawls (1971) and his subsequent work.

The influence of Gallie’s article is perhaps even greater in the social sciences, where a number of theorists will occasionally claim that one concept or another is essentially contested. Quite recently, for instance, linguist George Lakoff has reintroduced, without revision or critical scrutiny, Gallie’s account into his own work on political discourse (Lakoff, 2006a, pp.83–84),(Lakoff, 2006b, pp. 23–24).

Despite his inclusion of the modifier “essentially,” I believe Gallie’s explanatory target overlaps significantly with my own: I suspect that we are interested in many of the same concepts and conceptual disputes, although the details of the discussion below will reveal that the match is not perfect. Although I will not independently discuss the “real life” examples of essentially contested concepts that Gallie briefly discusses—ART, DEMOCRACY, SOCIAL JUSTICE, and CHRISTIAN—I suppose that at least some of those satisfy the conditions that I presented in my earlier specification of contested concepts above in §1.3. Moreover, I suspect that the examples that I have presented in the target cases are ones that, upon reflection, Gallie would want to count as essentially contested, even though as we shall see, his conditions as they stand would seem to exclude them. Perhaps Gallie would characterize his project differently. Perhaps, although the text seems not to bear this interpretation, he means his conditions to be strictly stipulative, picking out a care-

---

<sup>1</sup>Gallie (1962). All page references are to this reprinted version rather than the original. The topic of contested concept is also discussed in Gallie (1956), although nothing in the discussion there deepens or revises the account offered in Gallie (1962).



fully circumscribed set of concepts for some specific purpose. Nevertheless, for the sake of providing a comparison with my own account, in what follows I will presume that Gallie and I are interested in roughly the same sort of conceptual phenomenon, driven by the same sort of disputes. I will consider the question which of our accounts provides the superior specification and explanation of that target.

Moreover, Gallie and I have similar justificatory or vindicatory ambitions for our projects. In particular, we each suspect, and are each concerned to demonstrate, that disputes of the kind in question may properly be understood as genuine, principled disputes, rather than as instances of equivocation, confusion, or baldly partisan squabbles. We each hope that our account of contested concepts will be able to show that these disputes are, as Gallie puts it, “of such a character that the notions of evidence, cogency and rational persuasion can properly be applied to them” (p. 138), and that those who carry out these disputes are not left with only “at best unconscious rationalizations and at worst sophistical special pleadings” at their disposal (p. 122). Gallie’s stated aim is to show that contests over contested concepts

are perfectly genuine [disputes]: which, *although not resolvable by argument of any kind*, are nevertheless sustained by perfectly respectable arguments and evidence. (p. 123, emphasis mine)

The italicized clause in the quote above suggests the reason that Gallie labels these concepts “*essentially* contested,” and highlights one supposition that Gallie makes which I do not. Gallie supposes that despite the fact that these disputes answer to standards of evidence, cogency, and argument, application of these standards will never, as a matter of fact, bring the dispute to resolution. In fact, Gallie seems to make a stronger claim than this: For essentially contested concepts the standards of evidence, cogency, and argument *are always and in principle insufficient* to fix a correct resolution to the dispute. Indeed, we will ultimately see that Gallie appears to be committed to the view that in such disputes, despite the convictions of the participants, no party to the dispute could be correct, simply because there cannot be a fact of the matter about the detailed content of a contested

concept.

### 2.1.1 Gallie's artificial example

Gallie's argumentative strategy begins with the stipulation of an artificial example involving a specialized and non-standard use of the concept CHAMPION. Drawing on the features of that artificial case, he then constructs a set of individually necessary and jointly sufficient conditions for the essential contestedness of a concept *c*, and then argues that those conditions capture the contested character of the "real life" aesthetic, political, and religious concepts that triggered his original interest in the phenomenon. But before we consider the set of conditions Gallie constructs, it is worth pausing to consider the details of his artificial example.

Gallie asks us to imagine a league dedicated to a competitive team bowling sport. Although the rules of the sport—its permissible moves, its order of turns, etc.—are clear and uncontroversial, the standards by which teams are to be evaluated are not. All of the teams and their fans agree that the number of pins a team knocks over during a period of play is important in the assessment of a team's performance; however, those tallies do not constitute a *score* for the game by which one team may decisively be said to triumph over another in a match. The consensus among all of the teams and their fans is that other elements of play—say, the speed with which the ball is pitched, the spin placed on it—are ones in virtue of which one team might be able to outperform or defeat another. However, we are to suppose that there is no consensus on the question of how these various elements are to weigh against one another in assessing a team's performance, and there are no judges or referees who are charged with the task of declaring one team the winner or loser of a match. Nevertheless, it is a competitive sport: Its teams and fans are concerned with determining "the champions" of the league—the team whose play is the best. This concern will lead to pitched and sustained disputes. Gallie asks us to further suppose that each squad excels in some particular element of the game, and that fans of that team will tend to praise performances that demonstrate those elements, and will contend that their

avored element displays what the sport is “really about.” Each of the various squads and their fans will proclaim their own team the “champions,” and will attempt win converts to their side by presenting arguments for the relative importance of those elements of the sport in which their own team excels compared to those elements in which other teams admittedly outperform their own.<sup>2</sup>

In this setting, Gallie tells us, the concept CHAMPION serves as a perfect (though admittedly artificial and somewhat strained) specimen of what he means by an “essentially contested concept.”

### 2.1.2 Gallie’s conditions

With this hypothetical example in view, Gallie elicits seven individually necessary and jointly sufficient conditions that he takes to define the essentially contested character of a concept *c*. In the interest of clarity and concision I will list them in a rather quick and compressed form, and then follow up with a few quick expository remarks concerning his conditions (2)–(5). These are the conditions which are the most elusive and in need of exposition; they are also the conditions which I think most closely approximate my own preliminary account earlier in this chapter. Once we have his account in view, I turn in section §2.1.3 to a critical assessment of Gallie’s project.

Here, then, are Gallie’s seven conditions on the essential contestedness of a concept *c*:

- (1) *c* must be appraisive or evaluative, “in the sense that it signifies or accredits some kind of valued achievement” (p. 122).
- (2) *c* must be “internally complex” and “multi-dimensional.”
- (3) *c* must be “variously describable.”
- (4) *c* must be “open,” i.e., “persistently vague” (p. 122).

---

<sup>2</sup>If the reader finds Gallie’s example too far-fetched and at too far remove from actual variants of bowling games, perhaps an easier example would be a sport like figure-skating, only absent any judges to render an official score and proclaim a recognized champion.

- (5) Community members must use c “aggressively and defensively.”
- (6) There must be an exemplar (or a set of exemplars) that all of the community members accept as authoritative instances to which the concept c is correctly applied.
- (7) The contests concerning c must “probabl[y] and plausibl[y]” allow for the achievement mentioned in condition (1) to be “sustained and/or developed” (p. 133).

Conditions (2), (3), and (4) form a tightly interconnected set, centered on the ways in which the “internal complexity” of a concept is involved in its contestedness. Let’s begin with condition (2), which demands that in order to be contested a concept c must be “internally complex” and “multi-dimensional.” By these phrases, Gallie means that the employment of the concept must be responsive to a number of different “component parts or features” of the targets to which it is applied. For instance, the concept CHAMPION from his artificial example is internally complex in that it is responsive to several distinct aspects or “dimensions” of game-play—speed of attack, the spin placed on the ball, effectiveness in knocking over pins, etc. Likewise, Gallie maintains that the concept DEMOCRACY is responsive to a number of distinct features of political organizations (the presence of some form of popular election; equality of political opportunities of citizens, etc.). Gallie remarks that these features and elements that are appealed to here are to be understood as *descriptive*. For Gallie, the term “descriptive” appears to be interchangeable with “naturalistic”; he does not independently or explicitly specify what he understands by those two terms, although the relevant contrasting term, for Gallie, is “evaluative.”

Condition (3)—the condition that states that an essentially contested concept c must be “variously describable”—that there must be a number of different “Descriptions”<sup>3</sup> D<sub>1</sub>, D<sub>2</sub>, etc., which place different emphases or weights upon the distinct, descriptive “component parts or features” from condition (2).

To fix ideas, let us provide a simple example of some possible Descriptions, following up on the “component parts” Gallie discusses in his artificial example. Take the three elements of the sport mentioned above: speed, spin, and effectiveness at knocking over

---

<sup>3</sup>I will follow Gallie in using the capitalized “Descriptions” to signify these items.

pins (“score,” for ease of reference). We can then construct three very simple Gallie-style “Descriptions” of championship by assigning to each description a rank ordering of those three elements:

$$D_1 = \langle \text{speed, spin, score} \rangle$$
$$D_2 = \langle \text{spin, score, speed} \rangle$$
$$D_3 = \langle \text{score, speed, spin} \rangle$$

Thus Description  $D_1$  ranks speed over spin, and spin over score. Although Descriptions need not fit this pattern, for simplicity’s sake let us suppose (as Gallie does) that the orderings for each  $D_i$  are *lexical* rankings: For an adherent of description  $D_1$ , even marginal advantages in speed are sufficient to crown a team as “champions” over other teams who possess great advantages in spin and score; in case of ties in speed,  $D_1$  adherents turn to spin.

In order to meet condition (3), there must be at least two such descriptions that are at least superficially plausible ways of spelling out the criteria for application of the corresponding term.

(4) is by far the most recondite of Gallie’s conditions.<sup>4</sup> Gallie explains that this condition ensures that the “achievement [that  $c$  picks out] must be of a kind that admits of considerable modification in the light of changing circumstances; and such modification cannot be prescribed or predicted in advance” (p. 125). Explained in terms of his artificial example, Gallie says that this condition ensures that “*tomorrow’s* circumstances may bring out hitherto latent virtues in the play of *any* of the contestant teams” (p. 127, fn. 2, emphases in original). The example of “changing circumstances” Gallie introduces into his sporting example is that of bowling in new alleys that might propose special problems for a particular style of play. But his discussion here is not particularly illustrative: his remarks focus on the lowering or raising of standards, suggesting that adherents of a description like  $D_1$ , when they encounter a challenging new alley that poses difficulty for a speed

---

<sup>4</sup>It is also one of the most influential, as it appears to have directly informed H.L.A. Hart’s discussion of the “open texture” of general terms (see [Hart \(1998\)](#)).

bowling team, pointing to their team's performance as showing what "speed can do when everything seems against it" (p. 127).

... a proper use of it by  $P_1$  [an adherent to  $D_1$ ] in a situation  $S_1$  affords no sure guide to anyone else as to  $P_1$ 's next, and perhaps equally proper, use of it in some future situation  $S_2$ . (p. 125, fn. 1)

This is the condition that is supposed to secure for Gallie that the contested concepts in question are *essentially* and perpetually contested.

... even in a situation which conforms to conditions (1), (2) and (3) it is conceivable that experience should establish *one* style of description as again, for the purpose of moral or aesthetic persuasion, universally more acceptable than any other. This result could hardly be expected, however, if condition (4) be added ... (p. 127, fn. 2, emphasis in original)

I think that Gallie is on to something here, but it will take a little bit of digging to extract it. The gist of Gallie's contention is that the concept  $c$  is such that the conditions under which correct application are sufficiently subtle and variable that as new, unforeseen situations and settings emerge, these will pose problems and challenges for those who favor a particular Description  $D_i$ , who will then have to modify or refine their descriptions. And the way in which that Description will be modified will not be predictable in advance, and indeed it will be unpredictable along two dimensions. First, it will be in principle impossible to envisage all of the possible cases in which the concept might be called upon to perform its work. Some novel circumstances will always emerge, and they will put pressure and tension on the concept use. The second sort of unpredictability derives from the way in which those who have a particular  $D_i$  will accommodate these new features or combinations of features that emerge in this setting. There are branching possibilities available here—a concept-user might, for instance, virtually ignore the new feature altogether, or she might take it to present an occasion for a refinement or modification of the elements of their preferred Description. Thus, for instances, the introduction of balls of a variety of weights might lead some who had previously favored a speed attack to favor lighter balls and greater speed, while others who were proponents of the speed attack might favor whatever sized ball permits the greatest momentum in striking the pins.

A concept that meets conditions (1)–(4) is *ripe* for contest and dispute. Condition (5) requires that there be an actual dispute: community members must be actively engaged in disputing the question which of the various  $D_i$  is the correct one. For Gallie, in order for this condition to be satisfied, concept-users must view the employment of the concept as its own sort of competitive sport, seizing and remarking upon events that reflect favorably on their own preferred  $D_i$  and those events that reflect poorly on others'. As for the details of what this aggressive and defensive use of  $c$  consists in, Gallie does not give us much to go on, other than gesturing at exhortations and exclamations that the various partisans can be expected to offer in borderline situations. So, for instances, when a slowly pitched ball knocks over all of the pins, but in a rather weak and unimpressive fashion, the speed partisans should be expected to offer snide remarks along the lines of "...and you call *that* bowling?!" or "Some champions you lot are."

### 2.1.3 Critical discussion

I will turn in a moment to discussion of Gallie's remaining conditions, (6) and (7). For ease of exposition, however, I will orient the remainder of this section (§2.1) around critical scrutiny of Gallie's project. As we proceed, we should keep in mind that I am willing to accept that Gallie and I are discussing largely the same conceptual phenomenon: our explanatory targets are the same. Moreover, our explanatory projects are much the same: We are both interested in the question of how such disputes may properly be understood as turning on matters of evidence, argument, and justification, rather than on confusion, ambiguity, or merely partisan or rhetorical appeals. The critical questions I will be posing as we discuss Gallie's account, then, are (a) how well it specifies that explanatory target, and (b) whether it helps us to discharge that explanatory task.

**2.1.3.1 Exemplars and equivocation** Gallie's condition (6) claims that disputes over an essentially contested concept  $c$  will be anchored by a set of authoritative exemplars which all of the concept users recognize as instances where the concept correctly applies.

This condition plays at least two important roles in Gallie’s explanation of essentially contested concepts. First, the presence of an authoritative example is supposed to explain how arguments over contested concepts can appeal to considerations of evidence, argument and cogency. As we shall see in §2.1.3.3, Gallie will argue one of the features that disputants appeal to in the course of arguing for the correctness of their preferred Description  $D_i$  is the superiority of that Description in characterizing the exemplar’s achievement. I defer until that point consideration of whether authoritative exemplars are fit to play that role in the way Gallie hopes they can.

For now, let us consider the second role that exemplars are supposed to play in Gallie’s account of essentially contested concepts: aiding in answering a substantial objection. That objection holds that most—perhaps all—disputes over apparently contested concepts are simply cases in which the parties have distinct concepts and hence are not engaged in disagreement over the correct application of a single concept. Gallie is also concerned to answer this objection, which he nicely expresses in the voice of an objector:

“[T]he kind of situation you have described is indistinguishable from those situations in which people engage in apparently endless contests as to the right application of some epithet or slogan, which in fact serves simply to confuse two *different* concepts about whose proper application no one need have contested at all. The important question is how are these all-too-familiar cases to be distinguished from the artificial example which you have presented? To all appearances your concept of ‘the champions’ not only denotes consistently different sets of individuals (teams) according as it is used by different parties (supporters); it also connotes different achievements (in the way of different methods, strategies, and styles favored by the different teams) according as it is used by different groups of supporters. Is there, then, any real ground for maintaining that it has a *single* meaning, that *could* be contested?” (pp. 127–128, emphases in original)

My own response to this objection was to include as an explicit condition on the contestedness of a concept that it must be single, univocal concept. Once that condition was in place, of course, I owe an account—which I have only begun to sketch—of the differences that distinguish cases of genuine conceptual disagreement from cases of conceptual equivocation. But Gallie responds to the equivocation challenge differently. Rather than insisting on univocality as its own separate condition, Gallie contends that the presence



of authoritative exemplars—ensured by condition (6)—will provide him with basis for answering the equivocation objection.

Is Gallie's condition (6) true of all conceptual contests? Must such contests always involve examples that the whole conceptual community accepts as falling under the concept? Gallie does not provide an argument to that effect. What reason, then, do we have for accepting it? It seems to me that the plausibility of Gallie's condition derives from the difficulty we have in imagining a conceptual dispute between two parties where there is not even one example that they would both affirm falls under the concept. If two parties cannot find an actual case on which they agree, that fact would presumably go a long way to undermining their—indeed, anyone's—conviction that the two were engaged in a dispute over a single concept. Hence we are led to the plausible general claim that for any two parties to a conceptual contest there will be some example on which they agree. I do not wish at present to consider whether this plausible claim is true. But note that the truth of this plausible claim—for any two parties to a dispute there is an example on which they agree—would not entail that for any given dispute there is some example on which all parties agree. Suppose, as Gallie allows, that there are disputes to which there are more than two parties. Then even if the plausible claim were true, we might nevertheless have a kind of overlapping but imperfect consensus that would appear to be fit to ground a conceptual dispute even in the absence of an exemplar which all parties agree falls under the concept. For instance, suppose we have five parties to a conceptual dispute, each of whom favors a distinct, competing Description (one of  $D_1, D_2, \dots, D_5$ ) of the concept under dispute. Further suppose that we have five candidate examples  $E_1, \dots, E_5$ , which happen to generate as much consensus as we can manage to get from these disputants. Now suppose that for every  $D_i$ ,  $D_i$  bars exactly  $E_i$  from that set from counting as falling under the concept. Without filling in further details from a specific example, it is hard to make any sort of definitive claim on the matter; but the general structure of the case described surely seems coherent. And on its face it seems that when there is that sort of overlap and agreement

between the parties, there is nothing absurd or implausible in thinking that it is a single concept that is under dispute, even though there is no single example that all parties agree falls under the concept.

We should also note that Gallie's inclusion of this condition bars what seems to me to be a salient characteristic of contested concepts: that even when there are cases that are understood as central or canonical exemplars on which a community broadly agrees, later conceptual inquiry can sometimes lead a community to retract as mistaken their judgment that those cases were instance to which the concept applies.

If it is false that all conceptual disputes require agreement on at least one example, then *a fortiori* it is false that all conceptual disputes require agreement on at least one shared *authoritative* example. That much we can know without knowing anything about what it means to say that an example is authoritative. But—given that we are about to ask the question whether the presence of authoritative exemplars would equip Gallie with an answer to the equivocation objection—we should ask what Gallie means by saying that an example is an authoritative exemplar. An answer is not evident from the text. Other than the feature presently under consideration—that they anchor the univocity of a concept under dispute—all we are told about authoritative exemplars is (a) that they are cases to which all members of the community agree the concept applies and (b) that they are appealed to in the course of arguments in conceptual contests. Perhaps these two features are, for Gallie, all that are involved in the authoritativeness of an exemplar. Of course, each new independent feature we attempt to shoehorn into authoritativeness makes it harder for Gallie to demonstrate that condition (6) holds for every conceptual contest. So if the question is whether condition (6) is a necessary condition, then charity seems to demand that we stop with the minimal specification of authoritativeness. However, I propose to take it as read, for the reasons canvassed above, that the condition is false. Hence the remainder of our discussion of condition (6) will concern the question of whether authoritative exemplars, when they *are* present, are capable of fulfilling the roles that Gallie

intends for them. In considering that question, charity seems to require that we invest authoritativeness with as many compelling features as we can.

So here are some reasonable candidates for a robust account of the authoritativeness of these exemplars. Perhaps to say that an exemplar for a concept *c* is authoritative is to say that its status is unimpeachable and indefeasible: it is and forever shall be an instance of *c*, no matter what anyone says or does. Perhaps it is to say that no member of the community can reasonably deny that the exemplar is an instance of *c*, or can ever have any evidence that it is not an instance of *c*. Perhaps, if *c* is a gradable or comparative concept, it is to say that it is necessarily a perfect or superlative exemplar: that no other object could exceed or surpass the exemplar with respect to *C*-ness. Perhaps it is to say that the concept *c* was introduced through some manner of demonstrative, reference-fixing baptism exercised upon the exemplar. Or perhaps it is to say that resemblance or similarity to the exemplar is constitutive of what it is to be a *C*. Perhaps—and notice that this does not follow from the presence of the minimal features (a) and (b) above—it is even to say that an assertion that the exemplar is *C* is true or correct.

In order to try out these stronger versions of authoritativeness, let us turn to Gallie's preferred bowling example. Suppose that there is some team of yore,  $T_0$ , that is an authoritative exemplar of CHAMPIONSHIP: authoritative in all of the respects canvassed above. Suppose that the concept CHAMPION was introduced or coined with the exemplar in mind; suppose that no one can ever doubt or deny the championship of this team, that it is by all accounts the unsurpassed and unequaled greatest champion of all time. Would the presence of an exemplar with all of these features be sufficient to answer the equivocation challenge: that is, would the existence an exemplar with all of these features provide sufficient grounds for showing that the parties are engaged in a dispute over the same concept?

I think that the answer is pretty clearly no. To see why, we need go no further than to consider the different Descriptions  $D_1$ – $D_3$  we have sketched above, based on Gallie's own characterization of potential positions in the dispute. Now, suppose that adherents

of  $D_1$ ,  $D_2$ , and  $D_3$  (call them  $P_1$ ,  $P_2$ , and  $P_3$ , respectively) all accept  $T_0$  as an authoritative exemplar in the robust sense we are now considering. Such a supposition is only plausible provided that  $T_0$  was a team that exhibited, in a remarkable and salient way, each of the elements of the game that Descriptions  $D_1$ ,  $D_2$ , and  $D_3$  rate over all others: speed, spin, and score, respectively. For if  $T_0$  failed to exhibit any one these elements in a remarkable and unsurpassed way, then it would fail to serve as an authoritative exemplar for whichever party that favors the Description emphasizing that element to the virtual exclusion of the others.

Now consider: what follows from the fact that all three parties all correctly regard  $T_0$  as the authoritative exemplar of the concept CHAMPIONSHIP? Does this entail that all three have a single, shared, contested concept in common rather than three different concepts that each associates with the same linguistic term? For in contending that having an authoritative example in common is sufficient to block equivocation, that seems to be precisely the inference that Gallie wants us to make. He encourages us to reason as follows:

1.  $T_0$  is the authoritative exemplar for  $P_1$ 's concept CHAMPION
2.  $T_0$  is the authoritative exemplar for  $P_2$ 's concept CHAMPION
3.  $T_0$  is the authoritative exemplar for  $P_3$ 's concept CHAMPION
4. Therefore,  $P_1$ ,  $P_2$ ,  $P_3$  have the same concept CHAMPION

However, this inference is only acceptable provided that we suppose that  $T_0$ —or any other authoritative exemplar—can only be an authoritative exemplar for a single concept. But surely there is no reason why this must be so: given  $T_0$ 's remarkable performance in the elements of speed, spin, and score, there is nothing incoherent in supposing that the parties' different Descriptions  $D_1$ ,  $D_2$ , and  $D_3$  just correspond to three distinct concepts—SPEED-CHAMPION, SPIN-CHAMPION, and SCORE-CHAMPION—each of which is authoritatively exemplified in the performance of  $T_0$ , and each of which governs the use of their respective adherents' uses of the term "champion."

The conclusion to draw from the discussion of this section, I believe, is that Gallie's

condition (6) is not, in fact, a necessary condition for the contestedness of a concept. Moreover, if we hope to block the equivocation objection, it is not enough to appeal to the presence of exemplars on which the parties agree, even if those exemplars are in some extravagant sense “authoritative.”

This does not yet mean that Gallie cannot answer in some other way to the (quite plausible) equivocation diagnosis of his own artificial example. Indeed, I think that there is considerable merit in the “easy answer” Gallie initially offers to the challenge—which he subsequently finds insufficient and supplements by instituting condition (6). That “easy answer” is simply that all of the parties *believe* themselves to be employing a single concept: they believe themselves to be playing the same game and competing for the same “championship.” As Gallie points out, the objector may, just as simply, respond that this belief might be a “persistent delusion” (p. 128). Indeed, given Gallie’s description of his artificial case, it is difficult to see how the participants in that dispute could avoid reaching the conclusion that their earlier belief that they were playing the same game was erroneous. That is why I should prefer the following modification of the “easy answer”: All of the parties to the dispute are *committed* to employing a single concept in common. However, as we shall see, this version of the answer does not seem to be available on Gallie’s account.

What is missing from Gallie’s discussion is a feature that is central to the idea of univocality. One has to give weight to other parties’ convictions concerning championship: that a concept user’s own deliberation about which conception or Description is correct must draw not only on their own—or their own like-minded partisan group’s—prior use, but that it must involve some measure of deference or accommodation of prior patterns of use and judgments of others whom they take to be employing the same concept but with whom they disagree. We will have to make sure that the account of contested concepts developed here respects this demand.

**2.1.3.2 The descriptive versus the evaluative** In condition (1), Gallie explicitly restricts essentially contested concepts to evaluative domains. Perhaps this is not a surprise: the examples of contested concepts that come most readily to mind—including the “real life” examples Gallie introduces and my own target cases from §1.2 in Chapter 1—are all situated within value theory, broadly construed. But Gallie’s position is that this is a *necessary* feature of contested concepts: only evaluative or appraisive concepts could possibly qualify as contested. My own account of contested concepts to this point remains agnostic on this issue: nothing I have said thus far commits me to one position or the other. But by insisting upon condition (1), Gallie’s account has foreclosed on the possibility that this might have been coincidental or an accidental generalization rather than the reflection of some deeper characteristic of essentially contested concepts. We should consider whether he has reasons for doing so.

But before we address the more general question of whether contested concepts must be evaluative, we should briefly note that the gloss I quoted as part of Gallie’s condition (1) is even more stringent than that. Gallie contends that a contested concept is appraisive “in the sense that it signifies or accredits some kind of valued achievement” (p. 122). That is, any essentially contested concept must (a) classify only the results of agentive activities and endeavors—it must be directed toward taking stock of human achievements—and (b) that the characterization of these endeavors must be *positive*: Contested concepts must be understood to pick out activities, practices, or goals that the community’s members are prepared to praise in others or strive to achieve themselves. Moreover, condition (7) tells us that conceptual disputes must play a role in developing or sustaining that achievement. This derives from Gallie’s view about the distinctive social function that essentially contested concepts play, which is reflected in his condition (7). For Gallie, essentially contested concepts have special social significance because the competition over them provides community members additional incentive to strive for achievement, above and beyond the value found in the praiseworthy achievement itself: the thrill of competition through aggressive and defensive uses of a concept will provide further incentive to excel. So, for

instance, a community that contests the concept of democracy, Gallie thinks, is more likely to dedicate itself to realizing the various elements involved in the Description of that concept.

Let us grant to Gallie that *some* contested concepts are certainly concerned with the commendatory evaluation of human activity: After all, my own target case of the concept FAIRNESS seems to fit that bill. Must *all* essentially contested concepts be restricted to the commendatory? The example of CRUELTY from target case (2) in Chapter 1 shows that the disputes characteristic of contested concepts may emerge just as easily over negative “achievements”—over what counts as a result that community members are enjoined to condemn and avoid. Moreover, even if we suppose that all essentially contested concepts are evaluative, whether commendatory or condemnatory, that would not give us reason to think that all such concepts are concerned with the evaluation of agency and its results, unless we were provided with some reason to think that there are no evaluative concepts that apply outside of that domain.

But let us return to the more pressing general question of whether essentially contested concepts must be evaluative *at all*. Other than by stipulation or a hasty generalization based on his own preferred examples of essentially contested concepts, how can Gallie argue that this condition obtains?

Remarkably, Gallie’s arguments for the necessity of his conditions are mainly packed into a single footnote. In considering the status of condition (1), he writes:

There remains the possibility that the addition of condition (4) renders condition (1) superfluous. This could be maintained if, and only if, instances could be produced of a concept which conforms to my conditions (2), (3) and (4) and which is yet wholly non-appraisive in character. My suspicion is, however, that no *purely* naturalistic concept will be found conforming to my conditions (2), (3) and (4). (p. 127, fn. 2, emphasis in original)

Unless I am mistaken, Gallie has gotten himself turned around here. Surely he means to say that condition (1) will be superfluous—i.e., it will not further provide additional constriction of the set of essentially contested concepts—unless there is some “wholly non-

appraisive” concept that nevertheless conforms to conditions (2)–(4). But the upshot of his remarks for the question at hand is clear: He suspects that condition (1) correctly characterizes essentially contested concepts because he suspects that no “purely naturalistic”—i.e., no nonappraisive—concepts will satisfy the “internal complexity” conditions (2)–(4). These, the reader will recall, are the conditions that most closely approximate the conditions from my own preliminary account, and in this passage we find Gallie giving them pride of place over the (he thinks, likely superfluous) condition (1). In fact, he allows that if a concept were to satisfy conditions (2)–(4), he would be prepared to count it as contested even if it failed to meet condition (1). As necessary conditions go, this is about as weak as they come.<sup>5</sup>

Nevertheless, despite the failure of this argument, Gallie has presented us with a reasonable challenge: If we are curious about whether there could be non-evaluative concepts that exhibit the features that characterize contestedness, then we ought to try to find some examples and consider whether non-evaluative may exhibit the characteristics described in the specification above. This question will be taken up again later, in §4.2.

But before we prepare to undertake that challenge, we should take note of why Gallie suspects that it cannot be met. What are his reasons for thinking that every contested concept will have to be evaluative? His remarks, compressed within that same footnote, reveal certain key features of his views about descriptive (“naturalistic”, “non-appraisive”) classification and evaluative classification—features that explain why Gallie thinks no non-evaluative concept could exhibit contested behavior:

Given conditions (2) and (3) we have the sort of situation where a multi-dimensional description or classification of certain facts is possible. But in any such situation, specific evidential or methodological reasons apart, it would be absurd to prefer one style of possible description or classification to the others. But substitute achievements for facts, i.e., an appraisive concept or classification for a purely naturalistic one, and the absurdity disappears, since for the purposes of moral or aesthetic persuasion one style of description or classification may very definitely be preferable to another which is *logically* equipollent with it. (p. 127, fn. 2, emphasis in original)

---

<sup>5</sup>This concession presumably means that Gallie is likewise prepared to hedge on the necessity of condition (7), since it could only be necessary if (1) were.



The view that underlies these remarks seems to be something like the following. The activity of description is a naturalistic, non-appraisive attempt to report the facts—the way the world is. A description is the sort of utterance that can be either true or false, depending upon whether it reports the facts correctly. Descriptive concepts are either responsive to a simple sort of fact about the world, or else they are complex in a way that we might call “flatly conjunctive”: They are responsive to the co-presence of a number of distinct descriptive or naturalistic features of the world, each of which must be of equal weight. Disagreements over descriptive claims, or over the content of a descriptive concept, are, at least in principle, objectively resolvable into general agreement by means of general intersubjective rational procedures that appeal to evidence and argumentation.

By contrast, the activity of valuation, whatever it might be, is not an attempt to report any facts. As such, an evaluation is not the sort of utterance that can be true or false. Evaluative concepts are, at least in some respects, responsive to the way that the world is—indeed, they are anchored in descriptive features of the world, as we can see from the example of the various Descriptions that attempt to encapsulate the evaluative concept CHAMPION. However, although evaluative concepts may be complex and responsive to a variety of different descriptive features of the world, they need not be flatly conjunctive: they can be responsive to these descriptive or naturalistic features in a way that reflects different weight or influence among the descriptive features. Hence a single descriptive concept can correspond to a variety of different evaluative concepts that are responsive to the same naturalistic features, where each of these evaluative concepts may be differentiated by the distribution of weights accorded to the various descriptive features. However, disagreements over evaluative claims, or over the content of an evaluative concept, are not even in principle objectively resolvable into general agreement by means of general intersubjective rational procedures that appeal to evidence and argumentation.

This is a lot to take in, even if it is rather sketchy and schematic. I will want to return to this account of the contrast between the evaluative and the non-evaluative once we start to consider whether, and how, non-evaluative concepts might be contested.

In the next section, our discussion will turn to the vindicatory burden: the challenge both Gallie and I hope to address by explaining how disputes over contested concepts can be genuine, rational disputes. As we proceed, it will be worth noting the serious constraints under which Gallie must labor to meet this burden given his view of the distinction between descriptive and evaluative concepts.

**2.1.3.3 Rational and genuine disputes: The vindicatory burden** Both Gallie and I are concerned to show that disputes over contested concepts may be *genuine* and *principled* disputes—and we understand this burden in roughly the same way. Each of us wants to consider whether these disputes may be conducted in a way that is subject to rational standards of evidence, cogency, and justification. We both accept that such disputes certainly *may* be carried out in ways that blatantly disregard these standards, that rely instead upon appeals to emotion or greed, or what we might generally regard as attempts to provide arational psychological causes that may bring about “conversion” of other parties to the dispute. But is such the fate of every such dispute? How seriously can we take the claims of the parties that there are *reasons* for thinking that a particular conception or Description is the best, correct or proper way of spelling out the content of the concept?<sup>6</sup>

In broad terms, Gallie and I agree that these disputes may be rationally grounded. Ultimately, however, I think that Gallie’s attempt to show that essentially contested concepts meet the rationality burden is deficient. In particular, I think that, because of his overly restrictive view of the rational constraints on evaluative disputes, Gallie is in a position to show that these contests are rational and genuine only in an attenuated and inadequate sense. But in order to see what I think he gets right and where I think he goes wrong, and in order to learn from his attempt, let us consider Gallie’s strategy for demonstrating the rationality of disputes over contested concepts.

---

<sup>6</sup>In this respect at least, Gallie surely gets off on the wrong foot with the choice of his artificial example. In order to illustrate the liability disputes over contested concepts might pose to standards of evidence, reason, and cogency, he could scarcely have done worse than to choose the concept CHAMPION, which aligns these two sorts of competitions—contests within the game (where team allegiance, blind partisanship, and rationalizing, bad faith arguments are par for the course) and contests over the correct way of spelling out contested concepts.

The first stage of Gallie's approach to this question is defensive, amounting to an attempt to shift the burden of proof. Why, he asks, would we suspect that disputes over contested concepts *couldn't* be genuine or rational? The reason, he thinks, is that we presume that *a dispute can be genuine or rational only if it is possible, through rational means,<sup>7</sup> to achieve a universally acceptable resolution of that dispute.* However, as we have already seen, by Gallie's own stipulation, disputes over essentially contested concepts are such that they "are not resolvable by arguments of any kind" (see the quote on my page 20) and such that there is no "general principle for deciding which of two contestant uses of an essentially contested concept really 'uses it best'" (p. 139). Together with the presumption that the rationality of a dispute requires the possibility of rational consensus, these concessions would entail that disputes over contested concepts cannot be genuine or rational. However, Gallie says, there is little reason to credit the presumption, since (a) "universal agreement does not figure among the familiarly recognized criteria of rational justification" (p. 138), and since (b) there is no convincing defense of this requirement on rationality. That presumption, Gallie tells us, simply sets too demanding a standard for the genuineness and rationality of a dispute.

Instead, Gallie hopes to demonstrate that a weaker standard will suffice. Even though, he contends, it is not possible in the case of essentially contested concepts to find principles or arguments that exert a *general* rational force, and even though, he contends, it is not possible to come up with a general assessment of the rationality of the parties' positions, Gallie nevertheless will venture to claim that there *is* another way of showing that a dispute is rational and genuine. We can do so, Gallie claims, by satisfying a more modest standard: by showing that *individual disputants'* conduct and positions within the dispute—their "continued use (or in the more dramatic case of conversion . . . change of use) of the concept in question" (p. 139)—is liable to rational assessment.

Equipped with this new standard for the rationality and genuineness of a dispute,

---

<sup>7</sup>Gallie does not explicitly include the "through rational means" qualification here, but it seems to fit his position, since he countenances the possibility that universal agreement regarding a contested concept might be achieved through nefarious arational means like bribery, bullying, or extermination.

Gallie sets out to demonstrate that it can be met by disputes over essentially contested concepts, and in particular that it can be met by his artificial example of the concept CHAMPION. In demonstrating this, he aims to convince us, by means of careful consideration of a hypothetical case of the rationally driven conversion of one of the participants in the dispute, that the standard is sufficient. Let us note that, whatever its deficiencies, this will be the first detailed attempt presented thus far in the thesis that attempts to show, by means of example, how a dispute over a contested concept can meet the vindicatory burden and exhibit liability to standards of cogency, evidence, and the like. Since that vindicatory burden is so central to our discussion, and since the details of Gallie's attempt are going to prove important to our later discussion, I will quote from Gallie's presentation at considerable length.

Gallie begins by stipulating the configuration of various parties within his artificial example:

To show how [individual conversion that is subject to rational demands] is possible let me revert, yet once again, to my artificial example and consider the supporters of three contestant teams  $T_1$ ,  $T_2$ , and  $T_3$ . And for simplicity let us assume that the style of play of  $T_2$  can be said to stand midway between the styles of  $T_1$  and  $T_3$ . Let us recall, too, that in each of these groups of supporters there will always be wavering or marginal individuals, who are more than usually aware of the appeals—the characteristic excellences—of teams other than that which at the moment they favor and support. Let us concentrate on an individual  $I_2$ , at present a marginal supporter of  $T_2$ . (p. 139)

The case, as recounted here, is rather maddeningly unspecific. We are left to guess what sort of specifications of styles of play might fit the pattern that Gallie describes, and what it might mean to say that team  $T_2$  has a style of play (and presumably, that  $I_2$  subscribes to a Description  $D_2$  of the concept CHAMPION) that stands midway between the others.

And here is Gallie's recounting of  $I_2$ 's conversion:

A particular performance of Team  $T_1$ , or some shrewd appraisive comment from one of  $T_1$ 's supporters, suddenly makes him realize much more completely than heretofore the justice of  $T_1$ 's claim to be sustaining and advancing the original exemplar team's style of play in "the best possible way." This tips the scale for him and he is converted to being a supporter of  $T_1$ . (p. 139)

The example continues to be terribly under-described. Nevertheless, it is difficult to see how we could fill in the details so that any piece of this transaction could count as providing an argument or evidence or a reason that would support or compel  $I_2$ 's rejection of Description  $D_2$  in favor of  $D_1$ . What could a performance by  $T_1$  be like in order for that mere performance to give  $I_2$  a reason for her conversion? And what sort of "shrewd appraisive remark" should we expect from  $I_1$ ? Can we construct an explicit line of reasoning—an argument—that traces the line of rational force that leads  $I_2$  to convert? What form would an argument take if it were to draw upon that performance or that remark as evidence?

Gallie continues:

But now we may assume that the same particular performance (or shrewd appraisive comment) has had a comparable—though not so dramatically effective—influence on other supporters of  $T_2$ . It has slightly shaken them, we might say. At least it has made them aware that, in comparable circumstances  $T_2$  must make a comparable effective adaptation of *its* style of play if it is to keep their unwavering support. Further, we may assume that although supporters of  $T_3$  are less shaken by the particular performance, they have at least been made to "sit up and take notice"; and similarly, with decreasing degrees of force for supporters of other teams whose styles of play are still remoter from that of  $T_1$ . (p. 139)

Certainly we may assume or suppose that  $T_1$ 's performance or  $I_1$ 's shrewd remark has had this pattern of influence, if we like. And if we do so, the conversion and revision we have stipulated will have roughly the sort of ripple pattern we might expect a piece of evidence or a rational appeal might have.

While insisting that there may be this much objectivity in the grounds of any particular conversion, we may nevertheless agree ... that fundamental differences in attitude, of a kind for which no logical justification can be given, must lie at the back of the kind of situation which we have just discussed. Why should one style of play (as in our artificial example) appeal to one group of supporters and another style to a second group? Why should one facet of Democracy or of the Christian Message appeal so strongly to one type or group or communion, another to a second? (p. 141)

Remember Gallie's contention in support of his more modest standard for the rationality of a dispute: that in order to show that a dispute is rational or genuine, it is sufficient to show that an individual's position within that dispute is liable to rational assessment

or revision. We have just seen that Gallie's own particular exposition of that standard is deeply troubling. However, I think that there are general reasons to be suspicious of the standard's sufficiency. Indeed, I think that there are disputes that are neither rational nor genuine, but where individual parties' views are rationally assessable.

Take, as an example, the following dispute. Suppose you and I are arguing over whether rattlesnake meat is tasty: You claim that it is, I claim that it is not. The "dispute" in which we are engaged surely should not count as a dispute over a contested concept. Surely if any dispute is arational, if any dispute fails to be genuine, if any dispute is a case of interlocutors talking past one another, this would be one. There is simply nothing that you might say that would provide me with a reason for thinking that rattlesnake meat is tasty; there is simply nothing that I could say that would provide you with a reason for thinking that it is not.<sup>8</sup> However, individuals' stances and conduct within the discussion nevertheless remain liable to some form of rational assessment. To see that this is so, suppose that at various points during the course of our discussion, I make each of the following three assertions:

- (A) Rattlesnake meat is not tasty.
- (B) Rattlesnake meat tastes exactly like chicken meat.
- (C) Chicken meat is tasty.

The position that I have thereby staked out in the dispute is rationally untenable. No matter how permissive and relativized the semantics of the predicate "... is tasty," there is no plausible understanding of it on which I can coherently remain committed to all three of these assertions at the same time. When you pointed this out to me, you presented me with a reason for changing my stated position (although you have not given me a reason to give up any particular one of the claims), and you have done so by exploiting my position's liability to argument and demands of cogency and coherence. But despite that, it does not follow that our dispute over whether rattlesnake is tasty is a genuine or rational dispute. Thus, the fact that an individual position within a dispute is subject to

---

<sup>8</sup>This is not to deny that there might be some contested concepts in culinary settings.

rational force or assessment does not entail that the dispute itself—the apparent *interpersonal* disagreement—is genuine or rational. Thus, Gallie’s proposed modest standard for meeting the vindicatory burden fails.

But *why* has it failed? What has gone wrong? Why has this standard for the rationality and genuineness of a dispute failed, and might there be a way to correct it or at least learn from the error? Here is a tentative suggestion. Notice that Gallie considers only two extreme positions concerning the scope or force of argumentative, evidential or rational appeal. The first position—the one that Gallie apparently thinks is the appropriate position for non-evaluative disputes—supposes that the scope of those considerations will be universal, such as might in principle or under ideal conditions lead to complete consensus among all reasoners. However, when he considers the prospects for this standard in the case of evaluative disputes, he (perhaps quite properly) rejects it as too stringent. In response, Gallie switches instead to the other extreme: Rather than considering rational standards or evidence that has the broadest possible, universal scope, he considers an incredibly weak and narrow, strictly individualistic or solipsistic standard: one according to which only the individual’s values, convictions, judgments, or assertions can count as exerting rational force for that individual.

## 2.2 PEACOCKE, HIGGINBOTHAM, AND “IMPLICIT CONCEPTIONS”

In a set of articles a number of years ago, Christopher Peacocke and James Higginbotham articulated an account of conceptions and their relationship to concepts.<sup>9</sup> While there are points of contact between their views and my own, there are marked differences as well.

Peacocke’s and Higginbotham’s accounts have roughly the same view about the basic structure of the relationship between concepts and conceptions as my own account. All three of us agree that there are some concepts the contents of which establish some condi-

---

<sup>9</sup>See [Peacocke \(1998b,a\)](#); [Higginbotham \(1998a,b\)](#)

tions on our employment of the concepts, but which are indeterminate across a range of cases. We agree that a conception of a concept is essentially a more determinate version of that same concept, one whose content provides conditions or criteria that are more specific and more determinate than those provided by the corresponding concept.

The grand and overarching difference between my view and Peacocke's and Higginbotham's—the one that is responsible for the remainder of the differences—lies in the different kind of explanatory projects in which we're engaged. Let me briefly sketch the role that concepts and conceptions play in Peacocke's and Higginbotham's projects. But first: certain key notions recur with different labels in Peacocke's and Higginbotham's articles. For the sake of clarity, let's stipulate some terminology. Say that a thinker has mastered a concept *c* just in case she is consistently able to employ *c* correctly across a wide range of cases: That is, she is consistently able to make very fine-grained, specific, and correct applications and inferences involving *c*. Next, say that a thinker's formulation or explicit grasp of a concept *c* is the linguistic explication of the content of *c* that she is able explicitly to frame at a given time. Finally, say that a thinker is able to correctly formulate *c* just in case she can produce a detailed explication of the content of the concept that would correctly guide the employment of *c* across a wide range of cases.

Peacocke (Peacocke, 1998b) is struck by the following phenomenon: In some cases a thinker who has mastered a concept *c* is nevertheless unable correctly to formulate it. That is, thinkers are sometimes able to employ a concept in a way that outstrips the guidance provided by any formulation of the concept she can muster. A beginning logic student, for instance, might be unable to provide an explication of the concept of conjunction, but might nevertheless be able to judge and infer perfectly well about conjunction, almost as if she had, say, a written explication before her to use as a guide. Isaac Newton might not have been able to formulate the concept *LIMIT*, but he was nevertheless able to reason about convergent series almost as if he had an explication of it before him to use as a guide. Moreover, Peacocke suggests, a thinker who is unable to formulate a concept correctly will sometimes be able to put herself in a position to do so by careful scrutiny of her own



employment of the concept across a suitable range of hypothetical (“simulated”) cases.<sup>10</sup>

For Peacocke, this ability demands explanation. Unless we are willing to count such a thinker’s persistent correct employment of a concept as a string of happy accidents (and hence as unprincipled or unjustified), there must be some account of how a thinker’s understanding of that concept could somehow be more determinate than her conscious grasp of that concept. Peacocke suggests the best way of accounting for that success is to posit some sub-personal contentful state in that thinker, where (a) the content of this state is the (determinate) explication of the content of the concept and (b) this state is causally responsible for guiding the thinker’s correct employment of the concept. A thinker who has such a state is said to possess an implicit conception of the concept.

Higginbotham (Higginbotham, 1998a) also has as a central example the case of a thinker who has mastered a concept but cannot formulate it. But while Peacocke focuses on the transition from implicit to explicit mastery, Higginbotham pivots in the other direction: he contrasts such a thinker with one who merely possesses a concept she has not (yet) mastered. (The notion of possessing a concept that is in play here is much less demanding than the one developed in Peacocke (1992), which involves something closer to what is called “concept mastery” in these articles.) The criteria for possessing a concept do not emerge crisply in the Higginbotham articles. Perhaps it involves the capacity to have thoughts that involve the concept. Or perhaps it involves something akin to minimal competence—the ability to employ the concept in a range of central cases. Nevertheless, Higginbotham’s project is to explain what it is that a concept master has and that the conceptual novice lacks and therefore must gradually develop on the way to achieving mastery. And Higginbotham’s answer is that the master has, while the novice lacks, a *conception* of the concept in question. As in Peacocke’s discussion, a conception for Higginbotham is to be understood as a state of the thinker, the content of which more

---

<sup>10</sup>A note about motivation: Peacocke claims that the knowledge produced by such reasoning is a priori, since it is not grounded in empirical inquiry. Peacocke’s work in this area is connected with his recent project of rehabilitating some of the central tenets of classical Rationalism.

determinate than the content of the corresponding concept, and which accounts for the master's expertise in employment of the concept.

Thus, Peacocke and Higginbotham are each intrigued by cases in which a thinker's employment of a concept is more determinate than her explicit grasp of that concept. Each claims that the phenomenon is to be explained by the presence of these more determinate, submerged, subpersonal states. As such, to a rough approximation, Peacocke and Higginbotham are both concerned to use conceptions as a device for explaining an individual's grasp and exercise of concepts, and, as a result, their accounts are closely tied to the idea of conceptions as contentful (though not always consciously available) mental states of individuals, the contents of which are to be revealed through judgments about hypothetical cases.

We should also notice that the cases that interest Peacocke and Higginbotham all are supposed to be ones in which there is a unique and settled correct conception of the concept in question: Their discussion of each of their examples proceed as if there were a unique, uncontroversial correct explication of the concept. Some commentators in [Villanueva \(1998\)](#) press Peacocke or Higginbotham on precisely this point regarding some of the examples discussed in their articles, either suggesting alternative conceptions of a concept under discussion or challenging them to defend the correctness of the candidate conception that they suggest. (For example, see [Jacob \(1998, pp. 173-4\)](#).)

It is worth noting here a parallel between Peacocke and Higginbotham's research project and a use of the concept/conception distinction that fits squarely within the tradition in which I am interested. I am thinking here of John Rawls's tentative claim in ([Rawls, 1971, p. 47](#)) that the thought experiments he describes concerning justice might be understood as making explicit a body of implicit moral principles that guide and explain our moral judgments, on the model of Chomsky's implicit grammar. Rawls backed off of this suggestion in later work, but it serves as an illustration of the way in which the uses of the concept/conception distinction that descend from W.B. Gallie are related to the projects that are under discussion in [Villanueva \(1998\)](#).

Now, Peacocke and Higginbotham might be wrong about whether there is in fact a single, uncontroversial conception of the particular concepts they consider. But given the aims of their projects they are entitled not to be interested in the possibility that there might be a number of distinct candidate conceptions. They are each happy to presume for their discussion of examples that there is a single correct conception of a concept, or to suggest that the reader may transpose the discussion to the case of a concept that she thinks actually does have a unique correct explication. Sticking to concepts that have an uncontroversial unique correct conception is, in light of their aims, an appropriate simplification rather than an evasion or an oversight. Doing so merely removes a complication that would not enrich their discussions, and allows them to focus on the questions that matter for them.

We are now finally in a position to see the differences between my view and the view that emerges in Higginbotham's and Peacocke's programs.

The first difference is that I treat concepts and conceptions strictly as *abstracta*, as units of content or meaning, and I am not concerned with the question whether there are any corresponding states within a thinker. I do not wish to reject Peacocke and Higginbotham's claim that there are such contentful states. Moreover, I am not interested in explaining concept possession, concept acquisition, concept attribution, or concept mastery. I will also not be concerned about implicit conceptions, subpersonal content, or inarticulate mastery: as I have mentioned above, I will for the most part simply assume in the examples I discuss that all the thinkers involved are able to formulate the concepts and conceptions without much difficulty.

Second, the phenomenon that I *am* interested in—the phenomenon of contested concepts—comes into view only once we relax Peacocke's and Higginbotham's simplifying supposition that there is a single uncontroversial candidate conception of a concept. In his commentary on Peacocke's article, Georges Rey ([Rey, 1998](#)) presses hard on the improbability of the claim that Newton or Leibniz have embedded within them a determinate sub-personal conception that guided their employment of the concept LIMIT. One

might even urge that there are a number of different conceptions of limit that would fit the bill. Peacocke allows that some concepts might amount to what I will call contested concepts, but that his own research program requires “only the recognition that *not all* cases of theoretical development are the resolutions of indeterminacies” (Peacocke, 1998a, p. 63, emphasis mine).

Notice, however, that the point of having a distinction between concepts and conceptions becomes much clearer once we turn our attention to the possibility of distinct rival candidate conceptions of a concept. If we focus exclusively on those concepts that have a single, uncontroversial, determinate conception that is implicit in any thinker who is competent with respect to that concept, then it becomes hard to see the importance of a distinction between concepts and conceptions. Why not just say that the content of the concept CHAIR is that content provided in the full, correct explication of the concept, and then say that thinkers like Newton have a grasp only of part of the (settled and determinate content of the) concept?

### 3.0 WHAT CONCEPTS (AND CONCEPTIONS) JUST MIGHT BE

As its title advertises, this is a thesis about a particular type of concept. However, like many fairly ordinary everyday terms that are made to bear a lot of philosophical weight, “concept” has been appropriated and pressed into service in very different philosophical projects. So in order to avoid confusion, I aim in this chapter to make it clear what I am talking about when I talk about concepts. A thorough treatment of the general topic of concepts is beyond the scope of a work of this size, so the work developed in this chapter will of necessity raise many questions and beg a few others. But I am not terribly interested in engaging in squabbles about whether this or that account is correct or fits best with our pretheoretic notion of “concept.” To put this in the terminology of the dissertation: I do not want to offer a competing conception of the concept CONCEPT. Instead, I have a relatively modest aim: I hope merely to carve out my general topic area and make the notion of concept I plan to use clear, plausible, and defensible, and sufficiently well connected to some interesting philosophical problems before I go on to employ it in the succeeding chapters. Let others use the term as they may, so long as all those who use it are clear in advance about what they mean.

Once I have sketched the sense in which I have been using the term “concept,” I will be in a position in (§3.3) to give a quick parallel treatment to the cognate term that also plays a role in the project: “conception.”

One thing on which most accounts of concepts seem to agree is that concepts are ingredients in, or constituents of, content-laden entities (e.g., thoughts, representations, or propositions), and that the content of these ingredient concepts contribute to the content

of the entities containing them. Most accounts also accept the general methodological position that it makes sense to investigate concepts by considering those which correspond to lexical entries in natural language: If you want to investigate a concept, on this presupposition, you might as well start by investigating the concept that corresponds to the word “cat” or the one that corresponds to “red.”

One angle of approach comes at the topic of concepts with the aim of understanding them in the context of psychology or cognitive science, and thus have started by understanding concepts as particular mental representations or as mental representation-types, perhaps sub-personal states. In this thesis, I neither suppose nor preclude the possibility that there are these sort of discrete distinguishable processes or states, at either the level of token or type. If there are such states, then the connections between work on concepts in psychology and the cognitive sciences and my own work will turn out to be unproblematic: the states or state-types that *they* call concepts will then turn out to be those states or state-types that play a central role in a psychological explanation of how humans are the kind of creatures who employ (what I call) concepts. But, again, nothing in my own view will depend upon there being any such states.

### 3.1 CONCEPTS AS CLUSTERS OF NORMS

Once we approach concepts from the point of view of contents or meanings, we can see that concepts are normative: They provide standards for the correct employment of linguistic terms.

I propose that we can build a fecund account of concepts by taking the normative character of concepts to be their definitive characteristic. On the account I propose, a concept F simply is a cluster of norms providing standards for the correct employment of the corresponding linguistic term “F.”<sup>1</sup> More specifically, for the type of concepts stud-

---

<sup>1</sup>This allows me to sidestep controversies about whether there is non-conceptual content, i.e., content that doesn't have a corresponding lexical entry, so to speak.

ied in this thesis—those concepts corresponding to unary natural language predicates—I propose that the concept *F* is a cluster of norms that provide standards for the correct classification of objects as *F*.<sup>2</sup> Call a norm that is a member of that cluster—one of the sort that we are sometimes given to describe as “part of the concept of *F*”—a *conceptual* norm.

Note that the formulation I have suggested is framed in terms of *a* cluster or collection of norms, rather than the cluster or collection that governs the employment of the term. We are going to have to come up with a way of narrowing down the various clusters to find the concept. Intuitively, not all of the norms that provide standards for employment of that term will be conceptual norms that make up the concept *F*. For instance, norms of etiquette and prudence (e.g., those that counsel against the explicit application of the term to the utterances of a superior) and norms of implicature (e.g., those that enjoin the use of the term when a stronger term may correctly be applied instead) do not count as conceptual norms. We want to be able to say which norms are “part of the concept” and which ones are not. Again, speaking loosely, what we mean when we say that something is “part of the concept” is that it is part of an answer to the question what it is to be an *F*.

In fact, not even all of the norms that provide standards for the correctness of classification will be suitable candidates for membership of the cluster that constitutes the concept. For the sake of concreteness, I’ll try to bring out the difficulties here by working through a particular example.

Consider the concept of a lie, in the sense of an act of mendacity. The concept *LIE*, then, is a cluster of norms that provides standards for the employment of the corresponding general term “lie”. In particular, the concept *LIE* provides standards for classifying items (canonically, spoken or written assertions) as “lies.” As was mentioned above, not all of the norms that govern the application of the term will count as conceptual. Some, for instance, will concern whether an overt assertion of one’s classificatory judgment of an assertion as a lie is correct by certain other standards, e.g., standards of politeness, of

---

<sup>2</sup>“Objects” here should be understood very broadly, so as not to preclude the inclusion of events, states of affairs, or hypothetical objects—we will want to keep the door open to just about any sort of “thing” we might classify using a unary natural language predicate.

prudence, of legal burdens of proof, and so on.

And surely there are other norms that concern the application of a term in cases where thinkers have incomplete evidence. For instance, there might be norms that defeasibly license and encourage the application of the term “cat” to fuzzy four-legged critters who present a certain sort of appearance. But presumably it isn’t part of the concept CAT that cats have four legs. Similarly, there are likely norms that defeasibly license or encourage the application of the term “lie” when a speaker is sweating profusely and looking evasive. Such norms concern the application of the terms under partial or imperfect information about the case in question, and hence such a norm doesn’t constitute part of the answer to the question of what it is to be a lie.

Conceptual norms, then, are those norms that are involved in answering the question, “What is it to be an F?” So, then: how are we to determine which norms are part of the concept, or “analytic” or “part of the meaning of a term,” and which ones are not? Is there any way of taking the general formulations expressed above and turning them into more concrete tests?

We might have a hard time selecting the features that distinguish the conceptual norms from the non-conceptual ones, but as a first, provisional pass, one well suited to the kind of project undertaken in the other chapters of the thesis I hazard the following: To be a conceptual norm is to be a norm that governs the application of the corresponding linguistic term across those cases where a thinker takes herself to have answered these unresolved questions in the case—in the case of the concept LIE, once it is settled by evidence or stipulation that the utterance is false, that the speaker knew it to be so, and so on. Conceptual norms, then, apply across the range of cases—both actual and counterfactual—where deficiencies of relevant information about the objects are taken to be resolved, either by stipulation or investigation.

This account of the distinction between conceptual and non-conceptual norms does not provide the sort of sharp, hard and fast boundaries for our concepts that previous generations have hoped for or feared. Examples that debunk the idea that there is a hard-



and-fast boundary of that sort, including the discussion in [Putnam's \(1962\)](#) of the question of whether the norm #apply term t only to animals# is a part of the concept CAT.

No doubt alarm bells are going off in the readers' minds. What I have written thus far might be thought to signal a return to the bad old days, where concepts were definitions, or necessary and sufficient conditions. Am I urging a return to the bad old days where conceptual analysis and intuition-mongering held sway? Not at all. But I think that construing concepts as norms provides us with ways around some of the most common and devastating concerns about concepts, and allows us to view the philosophical predilections for conceptual analysis and intuition-pumping in a new light.

### 3.2 NORM-LIKE FEATURES OF CONCEPTS

This account of concepts rests squarely on the back of the notion of the normative. Unfortunately, in the space of this thesis, I cannot really even begin to address the whole host of intriguing and crucial philosophical questions that this dependence will raise. I will not develop a general account of what norms are, or where they come from. Nor will I be able to address the various types of skeptical challenges to their existence or even their possibility. For the purposes of this thesis, I remain neutral on most questions about the origins or the metaphysics of the normative, e.g., whether, or how, the existence of norms might be accounted for in a materialist or physicalist world view, and whether norms are fundamentally social or conventional or whether there are "transcendental" norms that derive their authority from something other than a community's social practices. I also take no stand on the question of whether norms must be public or private, although my discussion of conceptual norms proceeds as if these norms are in fact shared and sometimes capable of joint investigation and scrutiny. I do suppose that there can be norms that no community has yet adopted: One can consider a set of rules for a game that no one has yet played.

Rather than mounting a general account and defense of the normative, I will have

to rely on the plausibility of the position that there are norms, whatever their origin or make-up. Surely there are some rules, principles, standards for correctness, at least in some areas of human activity—wherever these come from, whatever they must depend upon, however skeptical challenges to their existence may be rebutted.

In addition, my account of concepts will draw upon some of the plausible features of norms which it seems that just about any account of the normative should have to accommodate. I highlight some of these features of norms, to point out that, once concepts are understood as clusters of norms, these features can now reasonably be thought to carry over to concepts. This, I claim, is a good thing, since these features comport nicely with the features that concepts are generally understood to have—including some features that have been marshaled to strike down alternative accounts of concepts that were unable to accommodate them.

### **3.2.1 Norms may be of various types and origins**

Norms may be introduced into a community in myriad ways. Some are introduced by stipulation or fiat by specific community members who are authorized to do so, others by the community's explicit joint adoption of a convention, others by a tacit integration into a community's activities. As I said above, I take no stand on the question of whether norms may or must have an extra-normative foundation. Nevertheless, any plausible view of the normative will also have to accommodate the fact that norms may surely incorporate or confer normative significance upon facts that seem clearly not to be in themselves socially generated or normative. (Consider, for instance, a nomadic tribe that has adopted a rule that enjoins them to travel each day whichever direction the wind is blowing at a certain hour of the morning.) In fact, a norm may even confer status and set standards of conduct in ways that outstrip a community's epistemic reach. There is nothing incoherent in supposing that a community has to rely upon evidence or guesswork in trying to meet a norm's demands, or in supposing that a community might be governed by a norm without ever having any way of knowing whether their actions were in conformity with it.

Consider a rule within the tribe that confers special status on the eldest member of a tribe, despite the fact that no tribe member knows which community member was born first.

Concepts seem to enjoy a parallel variety of origins and types, as we should expect if concepts are clusters of norms. A concept might be introduced in one fell swoop by a particular act by an individual or sub-community, or it might have a more obscure origin that no community member can recall. Moreover, a concept might require deference to extra-social facts. Suppose, for instance, that Kripkean theories about the semantics of natural kind terms are correct. Then the concept WATER dictates that what it is for a substance to be water—what it is for the term “water” correctly to apply to a substance—is for that substance to have the same underlying microphysical composition as the stuff that was originally dubbed “water”—even when a community is unable to settle questions about underlying microphysical composition. The account of concepts as clusters of norms, indeed, is compatible with a variety of competing views in the philosophy of language. Rather than a competitor in among them, then, this theory is perhaps best understood as a higher-level framework from which we can view these disputes. In particular, we can view many disputes within the philosophy of language as disputes over how best to limn the conceptual norms that govern the employment of a particular type of linguistic term: If we understand those norms as Kripke proposes we should understand them, how well does that capture the proprieties of our employment of natural kind terms?

### **3.2.2 Not all norms admit of crisp and comprehensive verbal formulations**

Some norms—e.g., the rules that constitute and govern the game of “Tic Tac Toe,” —are susceptible of rather clear, specific, and exhaustive verbal formulation. Such also appears to be the case with at least some concepts, including perhaps that hoary philosophical favorite, the concept bachelor. What philosophers have come to call an analysis of a concept, then, may be understood as an attempt to provide a linguistic formulation of the norms that constitute a concept; an analysis of this sort is correct or successful to the extent that it completely and correctly expresses those norms. Of course, even formulations that

are initially plausible and widely accepted might prove to be unstable, in that they might prove susceptible to challenges in the form of cases where the norms direct us to one judgment while adherence to the accepted formulation of those norms appears to direct us to the opposite verdict.

But for other norms, such formulation is at least hard to provide, and may in fact be in principle impossible. As an example of this outside the domain of concepts, consider pronominal anaphora. Evidently there are norms that govern the selection of the anaphoric referent of a pronoun—in unambiguous grammatical sentences there are uncontroversial correct answers to the question of which noun phrase a pronoun refers to. And native speakers of a language can in fact wend their way through complicated constructions and without much effort select that correct noun phrase. But speakers who can easily perform these tasks cannot provide a formulation of the standards that distinguish between the correct noun phrase and the incorrect alternatives; indeed even the best attempts of professional linguists to render a clear, complete, and specific formulation of rules that govern this selection have long met with a rather discouraging success rate (see, for instance, [Lappin and Leass \(1994\)](#)). So it appears that humans are subject to, and capable of following, norms that they cannot formulate or express.

A parallel claim appears to hold for concepts as well. We are capable of using concepts for which we are incapable of constructing such a linguistic formulation—indeed, for which we may have little reason to suppose such a formulation would be possible. The concepts *GAME* and *ORGANISM*, for example, might well have no formulation; so too with the concept *RED*. This is the pitfall that famously was the undoing of various earlier, discredited views of concepts—e.g., those views that, generalizing from their putative successes in some cases (like *BACHELOR*), just identified concepts with their definitions or verbal formulations. But understanding concepts as clusters of norms provides us with a general and uniform account of concepts, as well as a way of explaining the tight connection between a concept and its putative analysis, while nevertheless freeing us from being committed to the claim that every concept must have such an analysis.

Concepts, then, can be made a subject of shared inquiry and scrutiny. Sometimes a concept that cannot be given a full formulation may be given a partial formulation.

### 3.2.3 Norms can be indeterminate

Norms need not, and generally will not, cover all of the possible cases that emerge under their scope. A set of rules might simply be silent in some cases, neither requiring nor forbidding nor permitting a particular action. For instance, suppose that the rules for eligibility for membership in a Philosophy Graduate Student organization expressly declare that any Ph.D. student in the department is eligible for membership, and suppose that these rules also declare that those who are not graduate students in the department are ineligible for membership. Further suppose that these are the only provisions in the rulebook that concern eligibility for membership. These rules provide us with no verdict concerning the eligibility of Masters students, if any are enrolled. Alternatively, a set of norms might present us with standards of correctness that provide conflicting verdicts in a particular case. For instance, our graduate student organization might have a membership rule that dictates that no Ph.D. student shall be declared ineligible for membership, and another rule that bars convicted felons from eligibility. In the case of a Ph.D. student with a serious criminal record, each of these rules, considered singly, requires a verdict that is incompatible with the verdict of the other. And so the set of rules, considered as an aggregate, does not provide us with determinate guidance in such a case.<sup>3</sup> Of course, carefully crafted or amended sets of rules will be less likely to run into such difficulties. Moreover, sets of rules may be modified to accommodate some specific troubling cases that emerge, or they may be made to incorporate general interpretive principles that provide instruction for how to handle cases of mute or conflicting norms—a closure clause, for instance, indicating that only those explicitly listed as eligible are eligible. Or they may contain rules that tell us which rules' verdicts have priority over others'. But the general point remains: norms may have areas of indeterminacy.

---

<sup>3</sup>Tappenden (1993, p. 243), discusses similar examples, for rather different purposes.

Once we understand them as clusters of norms, concepts can be understood to admit of the same varieties of indeterminacy. It might be the case that for some concept *F* and some object *o*, the norms that constitute *F* fail to render a classificatory verdict: The concept might neither require, permit, nor forbid a given classification, or might appear both to require and forbid a particular classification. Hence the concept might, with respect to some objects, just render no classificatory verdict at all. And despite various philosophers' demands that concepts must be understood to have sharp, determinate boundaries, this indeterminacy seems a better fit with the way that we normally think about concepts. For instance, for a borderline case on the color spectrum, the concept RED might not entitle or require us to classify the object as red, and also might not entitle or require us to reject a classification of it as red.

The plausibility of each of these claims about norms, and the parallel claims about concepts, constitutes evidence, albeit weak evidence, that the normative theory of concepts is correct. My reason for bringing them up is that they are features that I think concepts must have in order to admit of the type that I discuss in the thesis. I hope that even if the reader rejects the account of the concepts that I am sketching, she will recognize the importance of the notion of a contested concept and that this recognition will be reflected in her own account of concepts.

### **3.2.4 Concept individuation and identification**

At various points throughout the thesis, I, and the various fictitious concept-users I introduce to illustrate the phenomena I wish to investigate, will make claims about the identity or distinctness of concepts. As a result it will be helpful to have at least a rudimentary view about how concepts are supposed to be identified and individuated. What distinguishes between a case where two different concepts are in play within a discussion and a case in which one and the same concept is in play? Part of the difficulty that emerges from this will derive from the difficulties in individuating or identifying norms. But I aim to show

that there are (i) straightforward cases where we do a decent enough job distinguishing and identifying concepts, and (ii) other cases where raising the question is intelligible.

Straightforward cases indicate that we often do a fairly good job of managing our conceptual kennels. I suppose that it is possible in principle to individuate and distinguish concepts from one another, and that in fact we do so in a loose and colloquial manner all the time. Sometimes, for instance, clearly the right thing to say is that there are two concepts associated with one and the same natural language term; that is, there can be instances of homonymy. This is what occurs, for instance, when there is one concept affiliated with the term “bank” that concerns the classification of financial institutions, and another that concerns the classification of terrain around bodies of water. I also suppose that it is reasonable to say that two distinct natural language terms—perhaps terms from different natural languages—may be governed by the same concept. These are cases of synonymy. Across different languages, we suppose that we have enough information about the classificatory norms of, say, German speakers in applying the term “Katze” across actual and hypothetical cases, to hazard a judgment that the concept that bears on their classification is the same as that which bears on our classifications employing the term “cat.” The same applies to cases of synonymy within a language, as is presumably the case for the concept(s) associated with synonymous English words like “couch” and “sofa.” Our claims of homonymy and synonymy stem from, and are grounded in, a judgment that the norms that govern their application are the same.

But are there any general identity conditions for concepts that we can develop that will help us to adjudicate harder cases? Let us consider the options.

Once we say that a concept is a cluster of norms, we would appear to have an easy option for stating identity conditions. Why not just say that a concept is to be individuated by the norms that compose it: That is, why not say that numerical identity of concepts—being one and the same concept—is to be understood as compositional identity—consisting of exactly the same norms. (Again, *modulo* puzzles about how to individuate norms.) But this easy option appears to impose an identity condition that is

too strict to make sense of actual judgments of conceptual identity: sometimes concept-users will assert numerical identity of concept—being one and the same concept—despite acknowledging compositional difference—i.e., despite differences in the norms that constitute the concept.

Perhaps sufficient similarity will serve as an appropriate standard for identity: two concepts are one and the same if they have enough overlap, or if their norms are sufficiently similar. But there are problems with this account, too. There are the general sorts of puzzles about similarity. But, more specifically and more troubling, there are extremely similar concepts that communities do recognize as distinct—concepts which are perhaps more similar than other concepts that communities regard as distinct.

Both of the standards that we have rejected to this point have been rejected on the grounds that they conflict with the ways in which concept-using communities make judgments regarding concept identity and difference. So perhaps our standard should just be that concept identity or difference is settled by whether a community that uses it explicitly regards two concepts as distinct or the same. The problem with that approach is that communities will sometimes, by their very own retrospective lights, make mistakes on this count, distinguishing between concepts that are later understood to be the same, or running together concepts that are later adjudged to be distinct.

The lesson, I think, is that while communities of concept-users do approach and resolve questions of concept identity and distinctness, and while outsiders can make similar decisions, there is no general, hard-and-fast, set of criteria. The judgments that are right to make will sometimes, it seems, rest upon sufficient or relevant similarity, or compositional identity. In other cases, judgments will have to do with historical features regarding the employment of a concept—whether present usage is profitably thought of as continuous with prior usage. Sometimes the claims of identity or difference will rest upon an intention to use the term in common with other communities or sub-communities, but this seems to be neither necessary nor sufficient for these claims. And to some extent questions of concept identity or difference will be sensitive to concept or purpose, or to the range of



conceptual questions that are being asked.

I think that the negative lesson we should draw from this discussion is this: it seems that there are no hard-and-fast criteria for identifying and individuating concepts. We should not expect to find ourselves, as philosophers, drawing upon them; nor should we expect to find our toy case example interlocutors drawing upon them. But there are two key positive observations: (a) Conceptual communities sometimes take very seriously the question of whether two concepts are the same or different, and (b) when they do take these questions seriously, there are matters for them to argue about and material for them to draw upon, with the result that it gets argued about in a principled way.

### 3.3 CONCEPTIONS

With my view of concepts now more fully sketched, I am now in a position to quickly indicate how this view will accommodate the other item that features in the thesis: that of “conception.”

On this account, conceptions are, like concepts, clusters of conceptual norms. A cluster of norms is properly regarded as a conception only relative to some other, less determinate cluster of conceptual norms: a conception of F takes a concept F and offers a resolution of some area of indeterminacy or conflict in the cluster of norms. That is, something is a conception of F in virtue of its making the “What is it to be an F?” norms—the conceptual norms—more precise, coherent, and determinate.

A conception might be a cluster of norms that no community has yet adopted: a particular conception of F might be a proposal for a novel or innovative manner of more determinate use. This is in line with my assumption (on page 52) that it makes sense to talk about norms that no community has yet endorsed.

### 3.3.1 Conception-relative assessment

Assessment of the truth of utterances or beliefs involving contested concepts, and assessments of the integrity of inferences involving them may be made *relative to a conception*. (Strictly speaking, I think the point applies not only to contested concepts but to any abstract concept that admits of a variety of different conceptions, even if those conceptions are not understood as competing.) At this stage in the project, I do not yet have a formal apparatus for working this out, but what I have in mind can be approximated in English by a pseudo-formal operator such as

On conception  $C_k$  of concept  $c$ , ...

such that, when this operator has scope over a sentence  $S$ , it fixes the semantic assessment of the corresponding linguistic term “ $C$ ” in  $S$  to a particular conception  $C_k$  of that concept.

This allows us to assess as true certain intuitively true sentences such as “On Goofus’s conception of cruelty, his caricaturing of Gus was not cruel” and “On Gallant’s conception of cruelty, Goofus’s caricature was cruel.” This construction allows a conceptual community—and outsiders who observe their conceptual behavior—to discuss and explore different conceptions of a concept and assess applications of them and inferences involving them without endorsing the particular conception in question.

But now we can also use this formulation to express the idea that underwrites the Competition Condition: that there is a unique correct conception of that concept. For now, let us use that sort of pseudo-formal operator in the following construction:

On the correct conception  $C_c$  of  $C$ , ...

which is understood as fixing the assessment of the sentence to the correct conception of the concept.

Given that in the case of contested concepts community members are committed to deferring to a unique correct conception of the concept, we should normally understand

utterances and beliefs as prefixed in this way, i.e., as liable to assessment relative to the correct conception of the concept.

#### 4.0 SOME EXAMPLES: CONCEPTUAL CONTESTS IN ACTION

A central claim of the thesis is that in order to understand the character of contested concepts—in order to understand their structure, their development, their semantics—we will have to understand the contests that are distinctive of them. From our discussion of the target cases, we have a view of how those contests *start*: We have seen the initial disagreement over a case, followed by a declaration of divergent conceptions of the relevant concept. But in order to address the desiderata, we are going to have to consider the way in which these disputes play out beyond that initial disagreement.

One of the first things that we learned from the discussion of Gallie’s attempt to exhibit “the logic of conversion” through the discussion of an extended example is that we need more careful attention to the sort of reasons that are offered in support of a particular conception or against one. It won’t do to, as Gallie does, simply to stipulate that one party makes a “shrewd” remark that triggers a change of view. We want to know what sort of remarks might be made, how they might be thought to provide a reason for a change of view, and what sort of reasons those are.

So in the following section I propose to develop two extended examples that trace the contours of another artificial dispute, but to trace them more closely rather than Gallie has. These will serve a number of purposes, the primary of which will be to exhibit an extended discussion. But it will also provide us with us with a way into discussion of two issues that have lingered in the background.

The first issue is the distinction between what we might call (following here H.L.A. Hart’s terminology from [Hart \(1998\)](#) in discussing social rules or conventions) an inter-

nal perspective and an external perspective on the dispute. The examples that we have discussed up to this point concern concepts that are contested within the community of English speakers. That is, the examples that we are discussing to this point concern cases where the reader herself will likely have intuitions about the concept's use. The examples that we are going to pursue will not have that obvious connection—they will govern terms that we do not actually use, so there will be no question of the reader's tacitly inserting herself in the dispute.

What we will discover is that when we view concepts of this sort—from this sort of external perspective, where the concepts in question govern the employment of terms that we do not use, or involve concepts with which we do not have an extended history of use—the temptation will be strong to suppose that the concept is nonsensical or empty.

The second issue that will be drawn out of the discussion of these examples is the question of whether contested concepts must only be “evaluative”—and by this I mean appraisive or conduct-guiding—in character, or whether some contested concepts might be found that derive from other domains.

#### **4.1 THE CONCEPT BORBITY**

Let us imagine a remote village with a population whose linguistic and conceptual behavior is for the most part similar with our own, with one area of exception: They are wont to proclaim some of the acts that they and their fellow village-folk perform “borb.” Until recently, their employment of the term has largely been coincident and uncontroversial. For example, when Motta twisted Abner's arm until he screamed, all of the neighbors agreed that they would not have thought Motta capable of such a borb deed. And when Shep tied poor Vic to his own horse and dragged him through the town square, the village newspaper's editorial page used “borb” or cognate terms no fewer than six times to describe what Shep had done, and the locals thought it quite apt. Naturally, there have been

some controversies. Everyone agreed that Letta shouldn't have punched Hector, but some weren't sure that she had hit him hard enough for the act really to count as *borb*. Others weren't even sure that she had meant to hit him—which, everyone granted, would have meant the act could not have been a *borb* one.

Let us suppose that “*borb*” is a term that features prominently and weighs heavily in the affairs in our little village. For instance, let us stipulate that these villagers are for the most part so docile and peaceful that they've never bothered to develop a political or penal system. Instead, the residents care deeply about “*borbity*,” and (most) try to avoid actions they judge to be *borb*, censure those who commit *borb* acts, and so on. In addition, our villagers have done a bit of what we might call primitive theorizing about *borbity*: They have a number of platitudes that they express and which they take as guides to their actions (“Better *borb* to yourself than *borb* to others,” “Only time heals the sting of a *borb* act”) and fables, to boot (“The Boy Who Cried ‘*Borb!*’”).

But now let us suppose that a recent incident has thrown the village into turmoil. One of the locals—a widower named Poppy—has taken to painting a large scarlet “R” (for “RapsCALLION”) on the foreheads of his children when their misbehavior displeases him.

The controversy concerns not whether Poppy is in the wrong; everyone agrees that he is. Nor is the dispute over whether he should desist; all agree that he should. Rather, the issue is whether what he is doing is *borb*. Let us suppose that primitive polling data from the local gossips indicate that the villagers are divided on the question: 48% think the punishment is *borb*; 48% think the punishment is not *borb*; and 4% just report being unsure. Everyone agrees that Poppy's children are not physically pained by being painted; and therein lies the heart of the dispute. Some villagers claim that *because* there's no pain associated with it, the act cannot be *borb*; others claim that cases like Poppy's acts of brandings can be *borb* even though they cause no physical pain.

In order for us—the readers—to have a proxy in the discussion, we'll introduce a stranger from another village—a village where the denizens, like us readers—have never

ever heard of the term “borb”. Call this proxy—this visitor to the village—Smith. Smith spends an afternoon in the village square listening to villagers talking about old Poppy and his poor children, and decides to wade into the fracas and help them sort it out.

“Well,” says Smith, “in the village where I live, we had a big fight over whether the Big Old Tree in our village square was over a hundred years old. We cut it down and counted the rings, and then we all agreed it was over a hundred years old—or had been, anyway, before we cut it down. Maybe you villagers just need to spend a bit more time and figure out the facts before you argue any more. Perhaps you can do some measuring, or maybe you could build a sort of ‘borb detector’—a device you could hook up to tell you whether an act is borb.”

“Mister,” says one of the village folk, “if we could have cut something open to resolve this dispute, we’d have done it days ago. But unlike interlopers such as yourself, we all know borb acts aren’t like that. You can’t measure something, weigh it, or count something to tell whether it is borb. In fact, there could never be a contraption you could build that could settle this dispute. And anyone who thinks there could be, is either just plain crazy or doesn’t know what borbity is.”

“And what’s more,” the villager continues, “we all pretty much agree on all the facts we can hope to find out. This isn’t like that time with Letta, when we weren’t sure whether she meant to punch Hector. We all know it didn’t cause the kids pain when they were painted, and we all agree that they were terribly embarrassed and ashamed.” The rest of the villagers in the square solemnly nod in agreement.

“You know,” says Smith, “I remember another time, back in my village, when we got into a big argument over whether this old-timer in our village was bald. We knew all the facts, there, too. We counted and found out that the fellow had exactly 88 hairs growing out of his scalp. And then we had another couple of fellows double-check the count. But that didn’t settle the matter for us. Some of us said that having only 88 hairs made him bald; others said that no man with 88 hairs on his head could count as bald. But then some other very smart fellow showed us that we could get ourselves all tangled up trying

to say whether someone was bald or wasn't, since we all agree that taking away one hair couldn't make someone bald and adding one more hair couldn't make someone *stop* being bald. Perhaps this problem of yours with borbity is like that kind of disagreement."

"We already had a case like that, Mister," a villager intones, "back when we weren't sure whether Letta hit Hector hard enough for the punch to be borb. We went through the same thing, and it didn't take very long for us to figure out that we were getting into the kind of trouble you describe. But this is different. It's not an issue like that, where we all agree about what matters in terms of being bald is the amount of hair someone has on their head, but disagree about where to draw the line on the question of how many hairs you have to have. The question with Poppy is kind of an all or nothing question: whether something that doesn't cause any pain at all can still be borb."

Smith ponders this for a moment. "Okay, then," he finally says, "I guess I just don't have any idea what kind of thing borbity could be, then. Maybe you should just quit arguing, shake hands, and say that there isn't any such thing as an act's being borb. Maybe it is just a local superstition. Or maybe it's just a word that you made up to use to express your disapproval of acts you don't like, or stuff you don't think people should do."

At this, cries go up in the crowd, and a villager hisses, "Fellow, we know that there is a world of difference between stuff that we disapprove of and stuff that's borb. Most of us here would approve of putting you over someone's knee and whupping you hard, if only it wouldn't be borb to do that. The borbity of that *explains* why we ultimately disapprove of it. And as for whether there is such a thing as an act's being borb, well, the only thing stopping us from doing just that is that everyone here knows that would be borb." Again: solemn—and now vaguely menacing—nods.

Smith shifts uneasily, and quickly decides upon another approach. "Look, the way this stands now, it seems you all agree on a whole bunch of cases of 'borbity' where pain was involved. I've always suspected that a word couldn't mean something unless at least a majority of folks who use it agree on it. I reckon that's true. And even if it isn't true now, it sounds like a sensible policy to adopt. That would settle this once and for all—'borb'



will only apply to acts that cause pain, and Poppy's painting won't be borb. And it would prevent this kind of dispute from emerging again with any similar issue."

This suggestion is greeted with happy nods from 48% of those gathered, but their glee is short-lived. For a voice from the crowd says, "Stranger, maybe we could do that. I suppose there's nothing *incoherent* about it. But it's not the way we've always done things around these parts, and there's good reason for us not to start now. There have been lots of times where that policy just would have gotten things wrong. For instance, time was when almost all of us agreed that only the whole numbers we used for counting were numbers, and that those new-fangled 'so-called' negative numbers weren't really numbers. But it turned out that we were wrong: Those really were numbers, all along. If we'd have adopted your policy before that, we would have been in a mess. We'd either have to have said that we had changed the meaning of the word 'number'—and it didn't seem to us that we had—or we would have had to keep 'number' to pick out what we now call 'positive numbers', and we'd have had to come up with a new word for the stuff that included the positive numbers and the negative numbers. I think that just struck most of us as a heck of a long way to go just to avoid saying we had made a mistake about numbers. We were all comfortable saying that we had just been wrong about what numbers were before. Now, in that case, it worked out really nicely for us not to follow the majority rule method that you're suggesting. So I see no reason why we should try this out now with 'borb'." The other villagers, remembering the 'number' incident, slowly nod in agreement.

Undaunted, Smith offers, "Well, where I come from, we had a tussle a while back over whether a thing you sit on was a chair even if it had no back to the seat. Half of us thought it was, half of us thought it wasn't. So what we did was strike an agreement. We agreed that we would just call ones with backs 'chairs<sub>1</sub>' and ones without backs 'chairs<sub>2</sub>', and since then we haven't had any problems over 'chairs'. Maybe you should do the same thing. Just break up your language so that you have *two* words: 'borb<sub>1</sub>'—which you will apply only to stuff that involves physical pain—and 'borb<sub>2</sub>'—which you'll apply to the other kinds of stuff the rest of you are worried about."

A voice pipes up from the back of the square: “Maybe that worked out for you guys only because that was a such a pointless thing to have a tussle over in the first place! Why would anyone care whether a thing was a *chair*? I can’t see how that would possibly matter! But this question about borbity matters a lot to us, Stranger, and no one here is going to be made happy by splitting things up like that. What words we use don’t matter to us: we’re still going to be arguing about which of these two picks out acts that are really borb.

“See, those of us Poppy supporters who think it has to involve pain would discount ‘borb<sub>2</sub>’ as a weak sibling of the real ‘borb’, the one that we use in all those sayings and stories. The rest of us—the ones who think that borbity doesn’t have to involve pain—are going to think that ‘borb<sub>1</sub>’ is too narrow to pick out all of the genuinely borb acts, and they’re going to be upset when we refuse to apply our old sayings and stories to cases like Poppy’s. See, having different words won’t solve anything: we’re arguing about what borbity *really is*. Did you guys argue afterward about which word really picked out chairs?” Smith had to allow as they had not.

“And please don’t suggest that we hold a vote to decide whether what Poppy did was borb,” a last villager muttered with disgust. “You should know from what we said before that none of us is going to be happy with that suggestion. Sure, maybe some of Poppy’s friends could do some politicking, handing out candy to get people to vote that borb acts have to involve pain. Or maybe some sweet talker could tell us all about Poppy’s kids’ plight in a way that made us all so welled up with tears that we decided that his act just had to be borb. But none of us thinks that this is the kind of question where people are entitled to take that kind of approach. That stuff is beside the point, and it would be a mistake to fall for it! We are supposed to be arguing about what borbity is! What it *really is*! What it is that we’ve been talking about, and scolding people for, all this time! People have got to argue about this, and use reasons to get us to see the other side. Maybe that will end our troubles, or maybe it won’t and we’ll just be arguing until we drop dead. But nobody here thinks that a vote or a pretty speech is going to change what is already really borb. That is a cheap way out.”

“Well,” says Smith, “what is left, then? What kind of reasons are you talking about? What kind of arguments could you possibly have in mind to do the job?”

“Well, for instance,” says one villager named Jones, “I am one of the folks who thinks that an act can be borb just by being especially humiliating. For instance, that time when Shep dragged Vic all around the town square with his own horse, we thought that was an especially borb deed. And I think the reason it struck as all that way is that it was so especially humiliating.”

“Come on, now, Jones,” says another, “that hurt Vic pretty badly. He was hollering in pain, and he has scars up and down his back to this day. I don’t see how that helps your side; that’s a case where there was a lot of pain present.”

“Well,” said Jones, “think about it this way. Borbity isn’t an all-or-nothing kind of thing. Some acts are more borb than others, right? I mean, we do sometimes sit around and talk about whether one act was more borb than another.”—General agreement on this point.

“All right, then, you all remember that time back a few years ago when Shep dragged Clem around the square? If I am remembering aright, we all pretty much agreed that this time with Vic was more borb than that time with Clem.”—“Well, maybe, go on.”

“Well, think about it now. Clem didn’t suffer any less pain than what Vic suffered this time around. So there must have been something else about it that made Vic’s dragging more borb. And it wasn’t that Clem had done more to get Shep all riled up, or anything else like that. We all agree that these were both pretty much unprovoked incidents. Now, if you think back to that time, you might recall that the one big difference that we all remarked on was how Vic got it worse than Clem, because at least Clem wasn’t dragged around by his own horse. We didn’t connect those two things at the time and say that was what made Vic’s dragging borbier than Clem’s, but I’m starting to think that just is what made the difference.”

“Well, Jones,” said one his opposite numbers, “you’ve given us something I guess we need to go think about for a while. But maybe we just got confused about borbitity this

time around. Maybe we just got more upset at Shep this time because he was a repeat offender, and in our anger we just mistakenly slipped into saying that this time around the dragging was borbier.”

“Well, then, how about this,” says Jones. “We all know that Shep has got a thing about dragging folks around the square here. But suppose Shep had taken Vic out into the woods, where no one could have seen him get dragged. Wouldn’t we say then that Shep’s treatment of Vic was less borb?”

“Now wait just one minute,” says the stranger Smith. “That didn’t really happen, did it? Shep only dragged Vic the one time. How can making up an imaginary story help tell you anything about what borbity is?”

“Well, these imaginary cases get our reckoning going and help us to make sure that we know what we’re all talking about. We all think we’ve got a pretty good handle on what ‘borbity’ means, and that we know how to ‘get along’ using it in new cases. Since we pretty much know how to think about borbity and whether to count acts as borb when we really see them happen, we all think that we should all be able to think about it in these imaginary cases. Now, if I start telling such a tale, I’m opening myself up to have to provide more details as the discussion goes on about it, so people can ask about other features of this out-of-town dragging that they might think were relevant. But for the moment, I am just going to ask the question again: If Shep had dragged Vic outside of town, where no one else could see it, wouldn’t that have been less borb?”

“I guess so,” “Yeah, sure,” “Sounds about right, I think,” come the sheepish replies. Jones lets the words settle in, the glances be exchanged, the eyebrows arch. “See now, Smith,” says Jones, “that’s the kind of arguing we’re talking about. I lay out an example like that, where it seems to suggest that there’s something more to the borbity in Shep’s dragging Vic than just what resulted from the pain. It looks as if folks are inclined to say that it would have been less borb if Shep had dragged Vic someplace not so public. But it wouldn’t necessarily have been more painful if he had done that. So if you say that, well then, it sure looks hard to say that borbity depends only with the amount of pain involved.

And if that's true, well, then, that looks a bit rosier for those of us who have been saying that Poppy's painting is borb."

"So, I guess you win the argument, then," says Smith.

"Whoa; hold up there, stranger," says Jones, "that doesn't end this thing. There are lots of different paths things could take from here. One thing these folks might do is decide that their reckoning is just plain wrong."

"That's pretty much how I feel about it," says a voice from the middle of the crowd. "I think something's just throwing me off track."

"Well," says Jones, "perhaps there's something else you have in the back of your mind that is messing around with the way you are reckoning about the story."

"Well, there's a lot more cement in the town square. Maybe part of what affected my reckoning about your story is that, without really noticing it, I was imagining that Vic was being dragged on nice soft pine needles and mud, and that's why I was inclined to say that it was going to be less borb than being dragged around the square."

"Well, then, I guess maybe to get this to line up with the other case you need to think about him being dragged around up on the mountain peak, where it's pretty much as hard and bumpy as the square," counters Jones.

"Or better yet, how about supposing I was dragged around the square when there was nobody around. And make sure you're supposing I got hurt just as much and in the same way as when that rascal dragged me here in broad daylight. What would you reckon then?" asks Vic.

"See, stranger," says Jones to Smith, "we're all thinking about it, and thinking hard and thinking honestly. Now, maybe someone might come up with a way of changing the example where our reckonings seem to get pulled in the other direction. Or maybe someone's going to trot out other kinds of examples that seem to get our reckoning toward thinking that there's got to be pain present before something counts as borb. They might trot out old cases where there was humiliation present but we declined to call the act borb, and try to show that there are lots of humiliating acts that we never called out as

borb before. That's a pretty serious challenge, and I'm not sure I could answer it right now, other than to say that we were just neglecting part of borbity that some of us see pretty clear now. Another thing they could try would be taking out a bunch of our sayings about borbity that are near and dear to our hearts, like 'Only time can heal the sting of a borb act,' and try to convince us that these near and dear sayings don't fit with what we're saying about borbity. Or they might try to push folks with my view for some kind of a connection between humiliation and pain; after all, if I'm saying that either one of them can make an act borb, then I should probably have something to say about why it is that these two pretty different things could make something borb. It'd be good if I had something to say about that, too. But it doesn't mean that the dispute is over just because one side or the other can't think of something to say at the moment. We're a careful and ponderous folk, and we don't take to forcing one another to hurry along in our thinking. We want to get things right. Maybe a few of us are disingenuous or have axes to grind, whether we know it or not. But, like we said before, for the most part we just want to find out what borbity really is."

## 4.2 THE CONCEPT SNORKER

Smith takes a while to carefully ponder this himself. At last he says, "I still don't see why you fellows should think that there's any such thing as borbity, or any way of getting things right or wrong with it. It looks to me, as an outsider, as if you guys have this word, and about one half of you take it to mean one thing, and the other half take it to mean another. Sure, you have a lot of cases where you both apply that term, but that doesn't mean you're using the word with the same meaning, or that there's any right answer to the question what it really means.

"Now, since you guys like the imaginary stories so much, let me tell one. Let's say that at the next village I pass through, half the people in some village use the word 'snorker'

to mean a horse and the other half used the word ‘snorker’ to mean a big, four-legged mammal. They may get along a good long time in calling all the horses ‘snorkers,’ but that peace is gonna break down real fast when the first elk wanders into town and half the folks say it’s a snorker and half the folks say that it isn’t. And it just doesn’t make any sense to ask who’s right about whether that elk’s *really* a snorker. There’s nothing about the way the world is that could tell you whether the thing is really a snorker, or which half of town is using the word right. Half of them use it one way, half of them use it the other. And there just isn’t any right or wrong about it, either way.”

“And I have to say,” Smith continues, “right about now you guys look as silly to me as those folks arguing about snorkers. And it isn’t any use in threatening again to beat me up and saying that the only thing stopping you is that it would be borb to cause me that much pain. That’d be just like these other villagers trying to argue me into accepting that there must be such a thing as snorkers because, after all, I rode into town on one.”

At this the villagers are slightly abashed. Jones steps forward again. “Stranger, now I have to ask you some questions about this little town you’re imagining. It sure sounds, from the way you tell the tale, as if these folks were already pretty set in their ways about how to use ‘snorker’ before that elk wandered into town. You said that half of them use it to mean horses and the other half used it to mean big four-legged mammals in general. What do you mean by that? What does that tell us about the state of things before that elk wandered into town? Does it mean these folks were already sitting around arguing about what the word meant?”

“No, I guess not. Not before they saw that elk,” replies Smith. “I guess that as I’m telling the story, they have never seen another big, four-legged mammal before.”

“Okay then,” Jones asks, “I guess we have to wonder how half of them could mean ‘big, four-legged mammal’ and not to mean horse. Saying that they meant different things by snorker—does that mean that each of them has a little picture or a little rulebook in his head that they drew on in using the word, and half of them had a horsey picture or a voice in their heads that says, ‘Only call horses ‘snorkers,’ and the other half had, what,

*mammally* pictures in their heads, or a rule that says ‘It’s okay to call any big four-legged mammal a ‘snorker?’”

“Well, that makes it sound kind of silly,” Smith replies. “But I guess that I do mean something like half of them have something like a definition in their heads, and that definition of ‘snorker’ tells them that snorkers have got to be horses. And then the other half, well they just got a different definition. Or maybe I just want to say that half of them are disposed to call only horses snorkers and half of them are disposed to call any big, four-legged mammal a snorker.”

“Now wait a minute,” says Jones, “it isn’t just that they’re disposed to use ‘snorker’ this way—it’s also that they’re disposed to get riled up when other folks in their village use it the different way.” —That’s right.

“And what do they say when they get riled up like that? Don’t they say to the other guys, ‘Hey, you’ve got it wrong!’”—Yeah, they do say that.

“So,” says Jones, “it sounds like these folks aren’t willing to let people get away with their dispositions, or just chalk them up as different somehow. And they don’t decide the matter by telling the other side, ‘You can just go mean by “snorker” whatever you like.’” —“Right,” Smith concurs.

“Well then, I guess they’ll probably not decide to settle it by subscribing their snorker talk.”—Nope.

“You know, when we stick to our guns and refuse to subscript some word, we usually have pretty good reasons for it. We think it really matters that we keep the word around without breaking it down that way. We only do it when we think the word picks out something special, something that’s worthy of our attention. Is that what’s going on here?”—Yeah, sure. They think it’s important.

“So it’s not like that time you told us about when your community got all riled up about ‘chair’, right?”—Right.

“Let’s set that aside for a minute, mister,” says another villager to Smith. “I’ve got a different question for you now. Just how do these imaginary guys go about conducting



this argument about whether their elks are snorkers?”

“Oh,” says Smith, gazing at the sky, “I guess by your rules I opened myself up for these kinds of questions by telling an imaginary tale, didn’t I? Let me see . . . I guess they get all terribly riled up, and they go back and forth, calling one another names, saying the other side doesn’t know anything about snorkers. Maybe some fellows switch their opinions around, but for the most part they stand their ground and refuse to budge.”

“Well now, see, that does sound pretty strange.”—How do you mean?

“See, you said just a bit ago that these guys care a great deal about whether an elk is a snorker. But they’re not willing or able really to argue about that question. When it comes time to dig in and try to work the thing out, all they do is sit around and trade insults. And if I wandered into that town, I guess I’d figure that it was full of pretty angry, stubborn folks, who don’t really care about whether an elk is really a snorker. I mean, maybe they’re just so angry at one another that they’re temporarily out of sorts and unable to muster a decent argument. Or maybe they’re just a bit shocked for the moment that it has turned out that there was any room for disagreement about snorkers at all. But I’d think they’d have some kinds of reasons why they care about whether an elk is a snorker, and that once things cooled down, the more reflective ones would be able to trot out those considerations.”

“I’m just not following you,” says Smith.

“It’s like this. Either there are some reasons available for thinking one side or the other is right about snorkers, or there aren’t any reasons. If there are reasons for that, then I’d like to know what kind of reasons they are before I say that these folks in your imaginary village are just being stubborn. But if there isn’t anything that they can say in support of their side, then it looks as if you’re right. We should laugh at these folks and think that they’re being stubborn and ridiculous. Because they’d be disagreeing as if there was some reason to think that they’re right and the other side’s wrong, but there just aren’t any reasons like that.”

“Well, then there you go!” says Smith. “Now we know that these folks are being silly,

because I sure can't think of any reason for deciding that an elk is a snorker! In fact, I can't imagine anything anybody could say on either side of that dispute that would give me a good reason to think that an elk is, or isn't, a snorker. And I'll bet that none of you can either, can you?" he concludes, a smug smile playing at the corners of his mouth. The crowd of villagers falls into a perplexed silence for half a minute or so, and most slowly shake their heads. "See," says Smith, "so I guess that there just aren't any reasons for or against thinking elks are snorkers. And now you finally see just how silly you guys look to me. I can't imagine a single thing you guys could say that would give me a good reason for thinking that Poppy's painting was borb, or that it was not."

After a long pause, Jones replies, "You might be right, Mister, that there's nothing that we could say that would give you a reason to think one way or the other about borbity, just like there isn't anything that these other villagers could say that would give any of us a reason to think one way or the other about snorkers. But it just doesn't follow that there's nothing anyone could say that would give the members of the community who know how to talk and think about snorkers reasons for thinking one way or the other about whether elks are snorkers. As outsiders who just aren't used to talking and thinking that way, who don't have any history with it, we just aren't receptive to the kind of things that they can say that should count as reasons for those insiders who do have that background. Maybe if we lived among them for a while and started talking 'snorker'-talk with them, we'd gradually be able to come to have a reasoned view on the matter. Until we get that kind of handle on it, nothing really should count as a reason for us one way or the other, and these guys are going to look silly to people like us who can't see why these folks are making the kinds of moves they are."

"So now, fellow, I want you to try to think about what these imaginary villagers might be saying to one another supposing they aren't just trading insults. If they really care about snorkers, what kinds of things would they argue about that might display why they care?"

"Okay, then," says Smith. "I think I see how this would go. Maybe they have a bunch of sayings about snorkers, and these sayings mention things like 'the mane of a snorker'."

Elks don't have manes, so I guess maybe that counts a bit in favor of discounting them as elks. In order to be consistent, they either have to tinker with those sayings, get rid of them, or maybe even, if it isn't too far-fetched, to think again about what counts as a mane. Or maybe they have some religious beliefs about snorkers, claiming that the mountain gods gave snorkers to the village because they're strong and just the best creatures to use for dragging logs around. Elks are even stronger and better at dragging logs than horses are, so maybe that counts in favor of elks being snorkers."

"Right," says Jones. "Now you're seeing how it might go, there, Stranger."

### 4.3 DISCUSSION

At this point, let's leave our intrepid interlocutors and turn to the question of what we can learn from their labors.

Let's draw out the lesson that is supposed to be thrown into relief by the example's use of a concept that is alien both to the reader and to the proxy Smith. The lesson is this: When the concept under discussion is alien in this way, it is extremely tempting to judge that the dispute is equivocal or vacuous, and to tie that judgment to the idea that the term under consideration is ambiguous or that the concept itself is vacuous. Where one is not a user of the concept in question, one will often take the approaches that Smith takes in turn: that the community members are simply "talking past one another", since they mean different things by "borb," or that they are grasping at the wind, since there just is no such thing as borbity.

For during the course of the discussion we see Smith attempt to use a number of distinct strategies to defuse or resolve the dispute, and in each case the villagers had good reason for rejecting those strategies. This is not to say that the villagers might not be wrong. For instance, it might turn out that their concept is confused, in the sense of (Camp, 2002): that is, as we have seen before, it might turn out that from the point of

view of the community itself, their own usage has run together two concepts that ought properly to be distinguished and made to do their work separately. But the fact that some such disputes can be resolved in this way does not entail that all can.

Based on what we have seen in these cases, I want to suggest the following preliminary explanation of how contested concepts might emerge within a community, and what a community might gain by having them. In doing so, I hope to satisfy, at least in part, desideratum (A).

First, it is often the case that some of a community's concepts are introduced in a way that leaves their conditions of application unsettled and underdetermined. For instance, a term "C" might be coined in a suggestive setting and the linguistic community might have little in the way of explicit guidance for carrying on in applying the term.<sup>1</sup> For some concepts, this underdetermination might be intended: the concept is introduced as abstract—that is the conceptual norms governing the application of the corresponding term are introduced as attenuated or underdeveloped—on purpose, in order that the details might be filled in and developed through future activity—through further application and theorizing—or because of a conviction that there is *something* onto which the concept fastens. For other concepts, the underdetermination might be inadvertent and unrecognized, the result of concept-users' consideration of only a narrower range of candidate cases, or from a mistaken conviction that the concept is already coherent and stable.

Second, a community's employment of abstract concepts will often proceed apace despite the weak and attenuated standards those provide. For a community will often (and often correctly) have the sense that an abstract concept that they employ is important or noteworthy—that their employment of it is responsive to some features of the world which are worth noting, or that it points toward some feature of the world or our practices that place demands upon our conduct. In such cases, a community of concept users will

---

<sup>1</sup>Of course, we can maintain this while still allowing that other among a community's concepts might be introduced as fully-fledged, stable, explicit concepts, with relatively sharp, crisp boundaries of application, perhaps, by means of stipulation or some Kripke-style act of reference-fixing.

often feel their way forward, deploying the concept and reasoning with it as best they can. A community's conceptual reach is thereby permitted to exceed its conceptual grasp, and in doing so their concept use might generate a history or pattern of employment of that concept that will contain lacunae or conflicts.

#### 4.3.1 Conceptual disputes without conceptions

Notice first that the dispute over BORB does not consist of an argument over which of two (or more) competing, full-blown conceptions is the correct one: Neither side in the dispute has attempted to offer a general or complete formulation of the concept BORB. Rather, the dispute is framed in terms of one rather specific conceptual question: whether borbity may involve humiliation or other forms of emotional pain, or whether only physical pain contributes to the borbity of an act. However, even though neither party is prepared at this point to articulate a particular conception of BORB and put it forward as the best, proper, or correct one, it seems from their remarks that the participants are committed to the idea that there *is* such a correct conception— witness their remarks about their aim to discover what borbity really is, what it is that they have been talking about all along when they have condemned acts as borb, and so on.

What are we to say about a dispute like this, in which a conceptual dispute takes place without the articulation of specific competing conceptions? It would be a problem if my view cannot accommodate disagreements that have this form, since many such discussions proceed piecemeal at this level of abstraction, dealing with crucial conceptual disagreements before the disputants have well worked out broad views about the content of the concept as a whole. I suggest that the way to fit these sorts of discussion into the framework I have proposed here is to say that such arguments involve *general* claims about the best, proper, or correct conception. We may understand the disputants as in effect claiming, say, that no conception of the concept could be correct unless it were responsiveness to humiliation, or that whatever the correct conception of the concept BORB is, it must be responsive to humiliation. Indeed, many conceptual contests may proceed this way

without significant impairment, beginning with the piecemeal consideration of a specific criterion of correctness on candidate conceptions, with the aim of winnowing down the field of candidates in the hope that the correct one eventually will be revealed. Although this sort of claim is more narrow and guarded than a claim that articulates a full-blown conception, it nevertheless stakes out a substantive position in a conceptual contest.

#### **4.3.2 The benefits of articulating conceptions**

Now, with just a little work, we could certainly try to extract more specific conceptions from the participants' remarks. And there can be little doubt that even if such formulation of conceptions is a difficult and tentative task, it enriches and deepens the discussion considerably by revealing a web of interconnected issues the disputants would do well (eventually) to confront. For instance, it seems that both sides hold in common that borbity involves the infliction of some sort of pain, and that said infliction must result from some intentional act. So given these areas of agreement and the point of disagreement highlighted above, we might try to frame the parties competing candidate conceptions of BORB in the following way:

**Poppy's Supporter's Conception of BORB:** An act is borb to the extent that it involves the intentional, unwarranted infliction of some degree of physical pain.

**Poppy's Opponent's Conception of BORB:** An act is borb to the extent that it involves the intentional, unwarranted infliction of some degree of physical or emotional pain.

In fact, as far as it is described in the narration above, the dispute over the borbity of Poppy's act might be understood in at least two different ways. The first, and perhaps more plausible way, is that sketched immediately above: the parties are engaged in a dispute over the concept of borbity, where something like these two conceptions are under consideration. But there is another possibility that might emerge as a more apt characterization, depending the state of the community's other concepts. For it might be, and we might discover from the contours of continued discussion, that the dispute over the bor-

bity of Poppy's act instead turns on the contested concept PAIN. For perhaps it would be apt to characterize the community as having in common the following shared conception of BORB:

**A Shared Conception of BORB** An act is borb to the extent that it involves the intentional, unwarranted infliction of some degree of pain.

The dispute, then, would turn upon a conceptual question about PAIN: whether humiliation (and perhaps other forms of emotional distress) can properly count as pain.

Either way of characterizing the dispute—as involving a contested concept of borbity or as involving a contested concept of pain—will account for the bare fact of the villager's disagreement over the borbity of Poppy's act, since each way of characterizing the dispute equivalently highlights the local point of disagreement: whether an act can count as borb in virtue of the humiliation it causes. But it would be a mistake to infer from that fact that the two ways of characterizing the dispute are *generally* equivalent or that it is a matter of indifference which way the dispute is characterized. For there will be wider conceptual differences and repercussions depending upon which characterization is the more apt. If the community has a contested concept of borbity, then we should expect the discussion to remain limited to actual and hypothetical putative instances of borbity, maxims and sayings about borbity, and so on. But if the concept PAIN is contested within this community, we should then expect the discussion (eventually) to take a turn that is not (yet) evident in the narrative above, highlighting cases and claims about pain that stretch far beyond the scope of discussions of borbity.

### 4.3.3 Non-evaluative contested concepts

The “snorkers” case points the way toward showing that contested concepts might not derive from evaluative or normative domains—that they can be, broadly speaking, descriptive. We do not have to have examples that derive from value theory: they do not have to be evaluative or normative. All that is needed is that the question of conceptual

application is normative. Once we have that, as we have seen, there can be a bona fide conceptual contest. It will turn on the issue of what sort of concept is in question. Is SNORKER a species concept, one that is responsive to natural or historical facts? Is it a functional concept, or perhaps a technical concept, one that involves the role that the animals play in some profession? The discussion of the concept *snorker* suggests that these are the sorts of issues that can be raised in a genuine dispute over a non-evaluative concept.



## 5.0 THEORETICAL NOVELTY

Now that we have a better sense of the character of contested concepts and the way in which disputes over them might be waged, I propose to return to the question raised in my discussion of Ronald Dworkin's remark at the very beginning of the thesis: whether philosophers' theories of content or meaning already have a place for contested concepts. I have already argued (in §2.1) that the one sustained attempt to develop this framework (W.B. Gallie's) is not up to the task. In this chapter, I turn to the question of whether other, better-understood tools from the philosophy of language or mind might provide an adequate explanation of the phenomenon without running into the same difficulties.

In order to carry out this task, I consider seven alternative explanations of the phenomena that contested concepts are supposed to explain. For each alternative diagnosis, I briefly discuss the general strategy and show how the strategy under consideration works for some cases that bear at least a superficial resemblance to Goofus and Gallant's predicaments in cases (1)–(3) and to the villagers' dispute in Chapter 4. But as we shall see, each of these fails to account for the central cases of contested concepts or to provide a general explanation of the phenomenon. Along the way, we'll be learning lessons from these failures that will put us in a better position to understand the nature of contested concepts and to understand their connection to other theories of mental and linguistic content, addressing desideratum (B) from §1.4.

## 5.1 THE EQUIVOCATION/DISAMBIGUATION STRATEGY

### 5.1.1 An explanation of the strategy

The first alternative explanation of the phenomenon exhibited in the target cases would be simply to deny that the univocality condition is met, thereby rejecting Goofus and Gallant's conviction that they are reasoning univocally about fairness, cruelty, and apology. Perhaps instead they are merely "talking past one another"—perhaps they have different concepts, or have radically different meanings associated with their respective uses of the terms. This diagnosis is made all the more tempting by the fact that in the examples above the disputants have managed to articulate different conceptions of the relevant concept. This opens space to raise the question whether some act would be cruel on Goofus's conception of cruelty, and then ask, separately, whether that same act would be cruel on Gallant's conception of cruelty. In fact, the disputants' answers to such questions might well coincide: that is, Gallant might well agree that on Goofus's conception of cruelty, Goofus's taking half the medicine is fair, and Goofus might accept that on Gallant's conception of fairness, it is not.

Once these conceptions are distinguished, it is easy to mount a charge of equivocation and propose a strategy for disambiguating the term in question. One could simply approach the parties in case (2), in which the concept CRUELTY is under dispute, and say,

Look, this dispute is silly: You each just mean something different by the word "cruel". This dispute can easily be dissolved if we just take care to mark the difference between what each of you is asserting. From now on, let's just use the word "cruel<sub>O</sub>" when we want to discuss cruelty as Goofus understood it, and then we'll just use the word "cruel<sub>A</sub>" when we want to discuss cruelty as Gallant understood it. So, Goofus, you have just been (correctly!) asserting that your caricature was not cruel<sub>O</sub>, and Gallant, you have just been (correctly!) asserting that his caricature was cruel<sub>A</sub>. And so you see: the two of you never really disagreed after all!

Note that this is precisely the sort of charge that Smith raises—and which the villagers reject—in the hypothetical example of the conceptual contest over BORBITY. (See page 68.)

Now, surely there are disputes for which this diagnosis is apt. As an instance, consider the following example:

**Case 4 (RESPONSIBILITY)**

Goofus's and Gallant's cars have collided. Gallant disavows any responsibility, since only Goofus had violated any traffic codes. Goofus replies that Gallant must be partially responsible for the crash, since, after all, it would not have occurred if Gallant had not driven his car into the intersection just as Goofus ran the red light.

Most likely the right thing to say about such a case is that there are two different concepts of responsibility in play in the dispute—one a concept concerning liability or blame, the other a concept concerning cause—and that, even though they might be interconnected in interesting ways, each is important in its own right and distinct from the other: Failing to distinguish them can only end in confusion, tears, and curbside rage.<sup>1</sup>

However, even though this explanatory strategy works in some disputes that resemble the target cases (1)–(3), it is hard to stretch it to accommodate all such cases. In particular, it cannot comfortably accommodate cases (1)–(3) themselves.

To start, let us note that there will be cases where the disputants reject the equivocation diagnosis and continue to insist that they are reasoning univocally, all the while maintaining that the other party is employing the concept of cruelty without fully grasping it. It is true the parties might be mistaken about their own dispute, stubbornly or clumsily refusing to accept an equivocation diagnosis when they should. Note too that it also can happen that disputants accept an equivocation diagnosis when they should not. Thus it seems that the acceptance or rejection of a disambiguation is an imperfectly reliable indicator of whether there is genuine ambiguity at hand.

Here's a reasonable test for the aptness of the disambiguation strategy in a particular

---

<sup>1</sup>An exception: Suppose Goofus goes on to contend that Gallant should get a traffic ticket, or to argue that Gallant should have to pay for the full damages to his own vehicle, and so on. In such a case we should probably understand Goofus to be articulating a particular conception—something in the neighborhood of a strict liability conception—of the liability-related concept of responsibility.

case—a test that is independent of the disputants’ explicit views about whether the strategy is apt. Consider the stability of the resolution it brings to the dispute and other, relevantly similar ones. If an appreciation of the disambiguated alternatives enables members of the linguistic community to harmoniously go about the business of reasoning about responsibility, without bogging down in related disputes, that would redound to the credit of the equivocation diagnosis. For instance, suppose that when Goofus refuses to pay his elevated auto insurance premiums, his agent brings him around by reminding him that Gallant was only responsible<sub>O</sub> for their crash. Or suppose that Gus recognizes that he might be responsible<sub>A</sub> for his employees’ malfeasance even though nothing that he did caused them to take money from the till. Such results are good signs that an equivocation diagnosis was the right approach.

However, the disambiguation strategy is unlikely to bring a stable resolution to the disputes in the target cases, even if both parties were to accept the strategy initially. To see that this is so, suppose that in case (3) Goofus and Gallant accept the counsel to take the disambiguation approach to “apology.” Each pleads guilty to the charge of equivocation and they agree to monitor their “apology”-talk carefully, the better to scrupulously discriminate between these two “concepts” of apology. To mark the difference, they resolve to use “apology<sub>A</sub>” to pick out speech-acts that meet the constraints that Gallant proposed, and to use “apology<sub>O</sub>” to pick out speech-acts that meet Goofus’s constraints.

So far, so good. However, if we bring on board some reasonable assumptions about the ways in which Goofus and Gallant’s community’s concept of apology resembles our own, the apparent resolution of the dispute will almost inevitably turn out to have achieved nothing more than moving the bump in the rug. For essentially the same dispute will emerge once more, centered upon the question which of these two “concepts” picks out “genuine” apologies—apologies of the sort that have been the subject of the community’s earlier discussions of apology.

Why? The community’s previous, undifferentiated “apology”-discourse will likely have provided them with ample reasons to care deeply about apologies: They will have a wealth

of commitments about the status and consequences of apologies. For simplicity's sake, suppose that Goofus and Gallant's community has long embraced the following principles about apologies, which they have heretofore understood as univocal:

- ◇ To refuse to accept an apology is contemptibly rude.
- ◇ An apology goes a long way toward making amends for a harmful deed.
- ◇ 'Tis saintly to forgive an offense for which no apology has been offered.

Suppose, as is plausible, that members of the community remain committed to these principles even after the disambiguation strategy is in place. As a result, the members of the community consider themselves to be bound to employ these principles to make judgments—judgments of rather great import—about the conduct and status of community members. But now, in light of their acceptance of the disambiguation strategy, they are obligated to consider the possibility that their set of principles equivocates about apology. Surely there are some cases—likely including the “responsibility” dispute above—in which the diagnosis of equivocation might help to disentangle some rather nasty (perhaps latent or unnoticed) incoherencies in their principles (it is not too difficult to imagine the incoherent principles that might have been endorsed by a community that has heretofore failed to distinguish between the two sorts of responsibility discussed in Case (4) on page 86). And recall from the villager's discussion of the concept *BORBITY* Smith's report that such a disambiguation strategy was successful in his own linguistic community to resolve a dispute about the meaning of the term “chair” (see my §4.1 above at page 68.)

But what about a case for which disambiguation does not provide such aid? The case of apology probably fits the bill here: the principles cited above do not appear to be rendered co-tenable by reading “apology” as “apology<sub>A</sub>” in some of the principles while reading it as “apology<sub>O</sub>” in others. In such a circumstance, there is no reason to think that the set of relevant principles concerning apology are in fact equivocal, and the community members will presumably be within their rights to return to their prior conviction that these crucial principles are univocal. And now the dispute will quite naturally be framed in terms of the question of which (if either) of the two proposed “disambiguations” of apology is the

right one—the one that picks out the kind of speech-act, for instance, about which the community has these long-standing commitments. Each side of the dispute, of course, will come away thinking that their particular disambiguated version is the genuine one, the important one, and the one that is entitled to inherit all the special statuses conferred by those shared principles.

This, of course, brings APOLOGY squarely back into line with the account of contested concepts: There is a single concept in play, which admits of more than one plausible way of spelling-out its content, and these spellings-out are properly understood as competing for the status of the correct or right or proper conception of that concept. The disambiguation strategy has proved unstable, and the equivocation diagnosis has faltered with it.

Of course, even if the argument of the preceding paragraphs is sound, I have only demonstrated that one specific attempted disambiguation fails to explain one candidate contested concept. But this particular discussion carries with it an important lesson that might well generalize to reveal a common characteristic of contested concepts that makes them resistant to the disambiguation move. It appears that a contested concept can meet the Univocality Condition in virtue of the sort of role that the concept plays within a community, including its incorporation in a set of general principles, where

- (a) the community is strongly committed to those principles; and where
- (b) those principles explain in part the significance of that concept, e.g., why it matters whether an object falls under that concept.

Surely part of the reason Goofus and Gallant would not have been content to accept the disambiguation strategy is that they recognize what hangs in the balance. For instance, if his speech-act does qualify as an apology, then Goofus has begun to make amends for the harm his remarks have caused, and Gallant must, on pain of being rude, accept his apology.

This observation also serves the purpose of explaining why such disputes persist when they do. It also serves in part to explain the earlier observation why so many central instances of contested concepts are moral and legal concepts, where holding principles in

common is crucial.

## 5.2 DISJUNCTIVE CONCEPTS

### 5.2.1 Explanation of the strategy

This diagnostic strategy takes its start from the observation that some of our concepts are such that there are several distinct ways in which an object or event can fall under them. So, in a relatively simple case, a concept F might just have two-pronged disjunctive criteria of application: an object or event can count as F in virtue either of being G or in virtue of being H. In more complicated cases, F might be responsive to a wider set of criteria of application: meeting some sufficient number or some weighted combination of those criteria, in different variations, might make an object count as F. Call such concepts—concepts which can be understood as involving more than one distinct way of falling under it—“many ways” concepts.<sup>2</sup>

### 5.2.2 How the strategy might work

It is fairly easy to see how concepts of this sort might be thought to generate disputes that resemble those in cases (1)–(3). First, take some “many ways” concept F. Then come up with a case in which at least one party to the dispute recognizes one (or more) of these ways in which an object can count as F, but fails to recognize one (or more) of the other ways. Call errors of this sort “one way out of many” errors.

---

<sup>2</sup>A number of other conceptual relations might fit in this category. Perhaps items fall under a generic concept in virtue of falling under a subordinate species concept. Perhaps items fall under a determinable concept in virtue of falling under a concept that is determinate with respect to that determinable. It might be possible to gerrymander disjunctive criteria of application for any concept F such that F turns out to be a “many ways” concept. I don’t think that this possibility matters for the project at hand, but if we like we could say that a concept is a “one way” or “many ways” concept only relative to particular ways of individuating criteria of application.

To see this approach in action, consider the following (rather outlandish) modification to Case (3):

**Case 5 (*Apology Redux*)**

After their exchange at the post-recital reception, Goofus and Gallant disagree over whether Goofus has apologized. Why? Gallant contends that an apology must involve the production of written text, whereas according to Goofus an apology is a vocal performance.

A reasonable diagnosis of such a dispute would be that Goofus's and Gallant's "conceptions" of apology are incomplete: Each captures one of the several ways in which one might produce an apology, but each errs insofar as it fails to recognize (or explicitly precludes) other ways in which an apology might be manifest. In short, each is guilty of the dreaded "one way out of many" error. On the basis of such a diagnosis, then, the right approach to the dispute would seem to be to attempt to consolidate these two "conceptions" of apology, with the understanding that the concept apology is sufficiently broad to comprise both vocal and written performances.

Will the same approach work for characterizing the disputes in the target cases? In particular, can the disputes in those cases be understood as "one way out of many" errors, without making recourse to the contested concept/competing conception framework? Here's how we might try to bring it off. We start by considering the claim that in each of these disputes Goofus and Gallant have each fastened onto one among several sets of criteria of application for the concepts they consider, and that each has mistaken his favored set of criteria for the whole story about how an act could count as an instance of fairness, cruelty, or apology. For example: we might attempt to characterize the dispute in case (2) by saying that Goofus and Gallant have gotten mired in a dispute in which each (correctly) presents one of the ways in which an act can properly count as cruel, but in which she also (incorrectly) denies the legitimacy of the other ways. That is, we can try to make the case that an act can fall under CRUEL either in virtue of causing physical



pain or in virtue of compromising dignity. Or, to characterize case (3), we might use a slightly more complicated scheme and suggest that the concept of apology is responsive both to the acceptance of blame and to the degree of regret for the harm one's actions has caused. More specifically, we might suggest that in most cases an act will only qualify as an apology if it involves the recognition of the wrongness of the act (this is "one way" in which an act can count as an apology), but, in cases where the agent's regret for the harm her act has caused is sufficiently deep and sincere, that regret can compensate for her failure to appreciate the wrongness of her act (this is another way which an act can count as an apology). Hence, we might say, each of Goofus's and Gallant's "conceptions" addresses only one aspect of the concept of apology and ignores that aspect's interaction with other relevant considerations.

### **5.2.3 Why the strategy fails**

So, does this approach provide us with a way of making sense of the interesting features of cases without adverting to the notions of contested concept or competing conceptions? Let's start with the good news. Clearly the feature of concepts that gets central play here is important and interesting; one of the central achievements of philosophies of language and mind was getting us explicitly to recognize that our concepts can involve such complexities. And just as clearly, we can make use of this bit of theory to sharpen up the presentation of our critique of the "conceptions" on offer in case p5.

But now the "bad" news. The "one way out of many" error approach is not going to supplant the contested concept/competing conception framework. In fact, quite often the "one way"/"many way" approach will turn out to depend on that framework.

How could that be? We are quite sure, when we address case (5), that we have correctly assessed the state of the debate: Goofus and Gallant are simply mistaken about the concept APOLOGY, each having committed a "one way out of many" error. And it seems that we have been able to develop this assessment without making use of the notions of contested concepts or competing conceptions. If we have been able to do so, however, I think it

is because Goofus's and Gallant's positions in case (5) are so obviously defective that it is hard to take seriously the idea that anyone in the dispute has offered a plausible, competing conception of apology. But now suppose that we were asked to provide reasons to back up our assessment of Goofus and Gallant's troubles. That is, suppose that, having offered the "one way out of many" error account of case (5), we are asked, Why should anyone think that Goofus and Gallant have each captured only one of several different ways in which apology can be instantiated?

The strongest argument we could muster, I think, would take the following shape. We would be able, presumably without much trouble, to point to a number of prior instances where the community has counted the production of written text as an apology, as well as a number of other cases where the community has counted spoken words as apology. Moreover, we might well expect that we would discover near-consensus in native speaker informants' responses to various posed hypothetical scenarios: Apologies can either be written or spoken.

With this support on board, we could press our case in the following way:

Goofus [/Gallant] must admit that his own conception of apology is mistaken, or else he must endorse and defend the following position: No written [/spoken] performance properly counts as an apology, and thus all of the earlier cases where written [/spoken] performances were counted as apologies, as well as the present disposition of native speakers to count them as such, must be understood as errors.

And of course it is hard to suppose that anyone would be able to sincerely, plausibly, or reasonably endorse or defend those positions. This, I think, is what underwrites our strong conviction that each of these boys has only got part of the story about apology: to get the whole story right, we see, they'd almost certainly have to adopt the "many ways" account of apology in order to relieve themselves of the obligation to defend the position about the community's employment of apology.

The same strategy is presumably what we would turn to for support of our "one way out of many" diagnoses of cases (2) and (3). To demonstrate that Goofus and Gallant each focus on only one of the "many ways" in which an act can be cruel or a performative

can be an apology, we should try to show that our “many ways” approach squares with the community’s employment of each of these concepts, and that Goofus’s and Gallant’s commitments to their respective “one way” positions puts them in less tenable positions when they attempt to account for their community’s employment of the term in question. For the time being, I won’t go into the details of how our arguments might go here. But we should recognize that our “one way out of many” diagnosis will be rather more difficult to maintain here than it was in case (5), largely because what Goofus and Gallant have on offer in these cases will presumably square better with the community’s employment of the corresponding term and will put each of them in a stronger position to defend the position that the community has been wrong about apology in the past or to respond to our contention that they fail to recognize other ways in which an act could qualify as an apology. So far, so good. It surely looks like we’re holding our own against Goofus and Gallant.

But wait a minute. Our objective was not to enter into the dispute with Goofus and Gallant; rather, we were trying to provide a diagnosis of their dispute that didn’t rely on the notion of contested concepts. But now that it’s clearer what’s involved in making a “one way”/“many ways” diagnosis plausible, we are in a position to see why this diagnosis will in many cases depend upon the notion of contested concepts rather than supplanting or obviating it.

In the course of the discussion, I have been happy to let the reader carry on with the initial impression that “we” have solely occupied the position of detached observers: philosophers of language or mind who are studying some hypothetical disputes and trying to figure out how to fully explain those disputes using the “one way”/“many ways” distinction. No doubt that is a reasonable way of understanding “our” relationship to the dispute in those cases where we take ourselves to have fixed in place—by means of stipulation, assumption, or even on the basis of sound reasons—a commitment to the view that the concept in question must be understood as a “many ways” concept. For instance, this seemed to be our situation when we were trying to explain case (5). But not all in-

teresting cases will rest on that fixed point. When we consider more subtle disputes like those in the target cases—or even when we are pressed to defend that commitment in simpler cases like (5)—we will often find that in order to offer a “one way”/“many ways” explanation of a dispute, we must slip into the position of becoming a participant in the dispute. We will find ourselves, that is, having to argue that the best way of spelling out the concept in the dispute involves understanding it as a “many ways” concept—and having to do this as a precursor to offering the “one way”/“many ways” explanation of the dispute. And of course what we’re doing under those circumstances—when we are trying to meet our obligation to argue that a particular concept is best understood as a “many ways” concept—sure seems to fit nicely with the contested concept/competing conception framework laid out in §1.3.

Thus a second alternative explanation has fallen by the wayside. But once again, in demonstrating its inadequacy we have learned a bit more about the way in which disputes over contested concepts work.

## 5.3 CONTEXT SENSITIVITY

### 5.3.1 Explanation of the strategy

Perhaps the disputes between Goofus and Gallant are to be explained by appeal to way in which the concepts involved are sensitive to the context in which they are employed. This strategy has been used with a fair bit of success in explaining other apparently vacuous or irresolvable disputes. Consider contextualist accounts of knowledge, such as those in [DeRose \(1992\)](#) and [Lewis \(1996\)](#). On these accounts, the concept of knowledge is sensitive to whether a putative knower is in a position to rule out all of the relevant alternative possibilities, where the relevance of various possibilities varies from context of ascription to context of ascription. Thus a detective who has found solid alibis for the maid, the valet, and the chauffeur might well count as knowing that the butler murdered Mr. Jenk-

ins when the question is posed within the epistemic context in play down at the station house. But that same detective might not count as knowing that the butler did it when the question is posed during a discussion of skepticism in the philosophy seminar room, since she cannot rule out the (there-relevant) possibility that the whole investigation was just a hallucination.

### 5.3.2 How the strategy might work

Here's how context sensitivity might be thought to generate the curious features of cases (1)–(3) without recourse to the notion of contested concepts or competing conceptions.

Suppose that an epistemologist and a beat cop meet right after work at their favorite local bar to discuss one of the detective's cases over a beer. We should not be terribly surprised if they emphatically disagree on the question of whether the detective knew that the butler did it. We might even find that each of them is able, Goofus and Gallant-style, to offer a "conception" of knowledge that they take to ground their assessment of the detective's epistemic state. (The Beat Cop: "If all the other explanations are unreasonable, you know the one you're left with is right." The Epistemologist: "No, knowledge requires being able to rule out every logically possible alternative!") Thus it might appear that each disputant is defending a distinct, competing conception of the concept of knowledge. Nevertheless, viewing the matter through contextualist lenses, we can now see that the best explanation of the dispute might well be that each disputant has continued to carry on the conversation as if he were still in the epistemic context he left behind at work. Perhaps all disputes over contested concepts similarly mistake contextually varying standards for competing, incompatible conceptions.

No doubt context sensitivity has a role to play in a number of disputes concerning the concepts from the target cases (1)–(3): Some of the concepts under dispute in those cases appear to be context sensitive in several respects. For example, it seems quite clear that cruelty is context sensitive: the same action-type might well be cruel if *performed* in one context—at, say, a christening—but not if performed in another—at a Friars Club

roast. So too it seems right to say that one and the same act—say, Goofus’s caricaturing of Guy—would count as cruel when *assessed* according to the standards in effect in the christening, but again, might fail to count as cruel when assessed according to the standards in place when Slappy White is on the dais at the Friars Club. And surely it would be a mistake to treat these differences as constituting distinct conceptions of cruelty (the christening conception of cruelty and the Friars conception of cruelty?) that compete with one another for wholesale adoption by those who are considering the question what cruelty is. One of the advantages of contextualism is that it accounts for our intuition that our assessments of knowledge, moral status, etc. can and should shift across settings; without a contextualist explanation of the legitimacy of these shifts, they might otherwise appear irrational or erratic. Goofus and Gallant would do well to learn this lesson and to steer clear of conceptions that are so finely individuated that they fix in place the standards of cruelty from a particular context. Each might even be able to incorporate such flexibility into his own position in the dispute: for instance Gallant’s claim that cruelty is the compromise of dignity becomes more plausible—and better able to accommodate our intuitions about the christening and Friars club—if we allow the standards of dignity to vary with context.

### 5.3.3 Why the strategy fails

But can context sensitivity be the explanation of the particular disputes considered in cases (1)–(3)? I think not, and that instead the right conclusion to draw is that contextualism complements the contested concept view rather than subsuming or supplanting it. To see why, think about the point at which DeRose and Lewis join the argument over the concept of knowledge. The sophistication of their contextualism shouldn’t lead us to lose sight of the fact that they start from a specific and rather contentious understanding of what knowledge consists in, something akin to the following: to know that  $p$  is to be in an epistemic position to rule out relevant alternative possibilities to  $p$ . This starting point is one among many competing theories of knowledge on offer, and it should probably be viewed as slugging it out with reliabilist conceptions of knowledge, causal conceptions of

knowledge, epistemic virtue theories, and the like. That is, contextualist theories should be understood as offering and fleshing out one particular competing conception of the contested concept KNOWLEDGE. We still must use that contested concept/competing conception framework in order to make sense of contextualist theories' role in the broader debate about the concept of knowledge.

Here is another line of thought that leads to the same result. The contextualist theories we have considered so far have worked by suggesting that a target concept is sensitive to standards (of relevance, of dignity) that are supplied by context. Variation of standard is what accounted for the different "conceptions" of knowledge offered by the beat cop and the epistemologist. But what sort of variation of standard might be thought to explain the different conceptions offered by Goofus and Gallant in the target cases? What accounts for the different standards that Goofus and Gallant employ in their disagreement in those cases? The sort of contextualist approach we are considering here will account for their dispute only if we can find some common ground regarding cruelty—something that can occupy the same role for cruelty that was occupied in the detective story by the ability to rule out relevant alternatives—and then explain what sort of standard that common ground relies upon, and, ideally, explain why Goofus and Gallant disagree about which standards are operative. The prospects for getting this done appear quite dim for the case of the concept of cruelty.

It might be thought that we should now consider another, broader sort of contextualist approach. Perhaps what contexts supply in cases like (1)–(3) is not a standard, but rather a conception. On this approach, we might say that the concept cruelty somehow varies widely in its content from context to context: in one setting, cruelty is the infliction of physical or emotional suffering, in another it is the compromise of dignity. This strategy might work, but it is rather ad hoc and smacks of desperation. (Think about how unsatisfying a contextualist theory of knowledge would be if it were to report that in some contexts knowledge is the ability to rule out all logically possible alternatives, while in other contexts knowledge is a reliably veridical mental representation of some state of

affairs, and so on.) Surely it would be better, if we could, to find some common conception of cruelty that provides a substantive account that explains what unifies or holds together our judgments and inferences about cruelty. (And if we can't find something that does that job, perhaps we should check to see whether those judgments involve equivocation!) Note, too, that to take such a broad contextualist approach would be to offer up a particular (and not altogether satisfying) conception of cruelty, which would properly be viewed as in competition with other conceptions, including Goofus's and Gallant's.

It doesn't appear that contextualism will do all the work we hoped our theory of contested concepts would do. None of this should be taken to minimize the achievements of contextualist theorists, or even to suggest that their arguments have no bearing on arguments over contested concepts. The works of DeRose, Lewis, and others have made a contribution to the broader investigation of knowledge by showing that a certain conception of knowledge can elegantly address some persistent problems. That's not only an achievement in its own right, but it also gives their conception of knowledge a leg up over other conceptions that cannot adequately address those problems.

## 5.4 SORITES VAGUENESS

### 5.4.1 Explanation of the strategy

I have said that contested concepts are indeterminate. Perhaps, then, contested concepts might be assimilated to another, well-studied sort of indeterminacy: vagueness. Of course, there is much disagreement about how to understand disputes over vagueness, so such assimilation would not resolve all the problems from the target cases. But there would still be some obvious benefits to this strategy, if we could make it work. It would provide us with a menu of well-studied approaches to the target cases, and we would free ourselves from the obligation to provide a distinct philosophical account of contested concepts.

Our hopes here might be buoyed by the fact that there are cases of disputes over



vagueness that have at least the same superficial structure as the target cases. Even better, we can construct such a dispute using one of the putatively contested concepts discussed in the target cases: CRUELTY.

#### **Case 6 (CRUELTY REDUX)**

Gus is a willing participant in a psychological experiment in Professor Z's lab. He is strapped into an apparatus that is primed to send 50 volts AC coursing through him at various stages of the experiment. Graduate Assistant Gallant claims that such a jolt would cause Gus so much physical pain that it would be cruel to carry out the experiment. Graduate Assistant Goofus, however, claims that the experiment would not be cruel: 50 volts will cause a fair bit of physical pain, he allows, but not enough that the experiment is cruel.

As with the target cases, we may assume that the dispute emerges in spite of the parties' agreement on a wide range of cases in which the concept is brought to bear (e.g., each agrees that to strike Gus with a tire iron would be cruel and that to offer him his choice of beverage would not be cruel) and that the dispute does not rest on a disagreement over other facts of the case (each is roughly familiar with the amount of physical pain 50 volts usually induces, each knows, from prior misadventures, a fair bit about Gus's capacity to tolerate pain, and so on). Moreover, the parties to the dispute stake out different positions on the question of how the concept CRUEL should be understood as more determinate, and these positions conflict or compete with one another. But for all that, case (6) is fairly clearly a classic dispute over vagueness. Doesn't this demonstrate that the concept CRUEL—and other apparently contested concepts—is simply vague?

#### **5.4.2 Why the strategy doesn't work for the target cases**

Before answering that question, we should be careful to consider what vagueness consists in. If by "vague" we just mean that a concept is unspecific, or indeterminate, or that it is capable of being made more precise, then surely contested concepts are vague. But the

sort of vagueness that philosophers have studied extensively—the kind displayed in the concepts HEAP, BALD, GREEN, and TALL—involves a particular pattern of indeterminacy, one with a distinctive sort of structure.<sup>3</sup>

A concept *F* is vague in this philosophers' sense if and only if its under-determined cases can be ordered into a sorites-style array. That is, *F* is vague just in case its underdetermined cases can—in a way on which almost all those who grasp the concept *F* would agree—be put into in a sequence such that the following principle appears to hold: For each member of the sequence, if that member of the sequence is *F*, then the next member of the sequence is *F*. So, for instance, BALD is vague in virtue of the fact that it is possible to arrange borderline cases of baldness in a sequence based on the number of hairs present on a head, where it seems correct to say that if a head with  $n$  hairs on it is bald, then a head with  $n + 1$  hairs is bald.

So this is the sort of indeterminacy—the kind with “fuzzy boundaries” and susceptibility to sorites paradoxes—that philosophers have studied under the heading of vagueness. And while some of the perplexities that arise concerning vague concepts stem from general problems related to indeterminacy, most of what philosophers have come to understand about vague concepts instead concern the question of how to address their curious features that depend upon their having this particular structure. In order to figure out whether the phenomenon of contested concepts can be explained within the framework of vagueness, we need now to ask whether contested concepts can be shown to involve indeterminacy that has this particular structure.

So, then, let's return to the concept CRUEL. I think that Case (6) does succeed in showing that the concept CRUEL is vague in the sense discussed above: That is, it shows that cruel does involve borderline cases that allow for the sorites-style ordering distinctive of vague concepts. Is 50 volts of alternating current a cruel jolt to administer to Gus? It sure looks like a borderline case, and Goofus and Gallant differ in their answers. How can philosophical theories of vagueness contribute to our understanding of the dispute

---

<sup>3</sup>See Keefe (2000) for a nice survey of the various attempts to say precisely what this structure consists in.

between them? Presumably Goofus and Gallant can find some related cases—past misadventures or hypothetical instances—on which they concur. Say that Goofus and Gallant agree that it would not be cruel during the course of an experiment to administer 1 volt of alternating current, but to administer 100 volts would be cruel. Now the dispute between Goofus and Gallant is primed to lapse into sorites paradox. Principles that express an appropriate sorites sequence for the case are easy to come by:

- (A) If administering  $n$  volts would be cruel, then administering  $n + .01$  volts would be cruel; and
- (B) If administering  $n$  volts would not be cruel, then administering  $n - .01$  volts would not be cruel.

Noting these facts does help us to understand certain features of Goofus and Gallant's dispute in Case (6). For instance, in Case (6) neither Goofus nor Gallant offers a "conception" of cruelty that marks a sharp boundary at which a jolt becomes cruel. The vagueness of CRUEL explains this, by showing why Goofus and Gallant would have difficulty drawing or defending a sharply-defined general answer to the question of what level of voltage does the trick. It also explains why Goofus and Gallant will be wise not to allow their agreement on related cases and on principles (A) and (B) above to get them to agree to a common judgment about the 50 volts case: The same principles that could be used to argue Goofus down from 100 volts to 50 volts could be used to argue him down to the position that 1 volt was cruel (and vice versa for Gallant). Thus, not only is the argument suspect because it leads us to conclusions we suppose are false, but because an equally plausible argument leads us to a contradictory conclusion.

So it appears that CRUEL is a vague concept. I will not have anything to say here about the more widely-studied question of how to understand vagueness in general, or about how to address it as it bears on the concept of cruelty. What I do want to emphasize is that the discovery that a concept is vague does not preclude the possibility that it is contested in addition. It is uncontroversial that one of the ways in which an act can be cruel is by the infliction of physical pain. It appears that all we have done in case (6),

then, is to isolate one feature or dimension that is uncontroversially connected with a concept—an aspect that is not contested—and then locate a dispute where the only issue that is contested is the intensity or degree to which that feature must be present in order for the concept to apply or fail to apply. At best this shows that the concept is vague with respect to that feature. But the target case concerning Goofus's caricature appeared to show that there can be disputes over the concept CRUEL may also turn on other sorts of considerations, considerations that are independent of the degree of pain required for an act to count as cruel. In particular, the dispute turns upon whether physical or emotional harm is necessary for cruelty, and whether affronts to dignity must be present in cases of cruelty. It remains to be shown that the phenomenon of vagueness can handle these sorts of complications.

Now, of course, vagueness can be more complicated than our initial discussion suggested. In particular, a concept might be vague in a way that involves borderline cases that are arrayed along a number of distinct dimensions. For instance, color concepts—which are thought to be vague if any concepts are—are best understood as arrayed not along a single dimension but at least three: hue, saturation, and brightness. For simplicity's sake, philosophical discussion of vagueness tends to proceed in the manner described in the last paragraph: We ignore the sorites sequence that arrays candidate bald heads in virtue of the thickness of their hairs, and we focus instead upon the sequence that is ordered by number of hairs. But there's no reason to think that the consideration of these new dimensions adds any in-principle or theoretical difficulties. Loosely speaking, we just move from thinking of a vague concept as inducing a fuzzy border upon a one-dimensional array of cases to thinking of that concept as tracing a fuzzy-bordered region in an  $n$ -dimensional array of cases. Perhaps contested concepts are just vague in this more complicated way.

Nevertheless, not even multidimensional vagueness can capture the full variety of indeterminacy involved in CRUEL. Consider again the details of the target case concerning CRUEL. Remember that the indeterminacy there turned on the question of whether physical or emotional harm was necessary for cruelty (Goofus's view) or whether cruelty need

only involve compromise of dignity (a consequence of Gallant's view). This dispute seems not to involve variation along any specific dimension(s) that admits of sorites sequencing. Instead, it seems to turn upon questions that concern whether and how a given feature or dimension is connected with a concept: whether physical or emotional harm, or compromise of dignity, is necessary at all for cruelty, in any measure or degree. To be sure, these issues set the stage for further questions that are obviously connected to vagueness. If it were agreed that cruelty is the compromise of dignity, we might be able to arrange a sorites array that poses the challenge of saying what measure of compromise is required for cruelty. But Goofus and Gallant are not here engaged in a dispute over fuzzy boundary lines among some such dimension(s): they disagree instead over whether the concept is responsive to those features or dimensions at all. A dispute over this sort of indeterminacy, with its call for a sharp yes-or-no answer rather than boundary-drawing, doesn't fit the structure we find in vagueness cases.

Vague concepts and contested concepts have challenges in common. Both must, for instance, withstand objections to the coherence of indeterminacy. And contested concepts can have areas of vagueness, as cruel does. But not all of the indeterminacy involved in contested concepts has the same structure that characterizes cases of vagueness.

## 5.5 MERELY EPISTEMIC INDETERMINACY

### 5.5.1 Explanation of the strategy

In the Indeterminacy Condition—condition (2) from the specification of contested concepts in §1.3 of Chapter 1—I contended that the disputes in the target cases arise due to the indeterminacy of the contested concepts involved. This alternative approach to the target cases urges us to reject that contention, and instead understand apparently contested concepts as instances of merely epistemic indeterminacy: That is, we could accept that there is a fact of the matter whether what Goofus has offered is an apology, but that

it is not possible for Goofus and Gallant to achieve a certain sort of privileged epistemic standing (knowledge, knowledge of their own knowledge, certainty) with respect to that fact.

There is a history of applying this strategy to instances of sorites vagueness, arguing, e.g., that there are sharp boundaries for our concepts like BALD that outstrip even our potential grasp of them, in the sense that we cannot even fathom how we might go about inquiring where those sharp boundaries are drawn. What makes such epistemic accounts implausible, I suppose, is that it's hard to see how our concepts could outstrip our epistemic or conceptual grasp. I propose to consider two different sort of cases where this sort "outstripping" is plausible, and to show why they don't provide reasonable hope that something similar is driving the phenomena exhibited in the target cases.

### 5.5.2 A case where the strategy works

There are at least two sorts of example we should consider here. Let us start with an example that explicitly involves an empirical concept—the concept SLEEP—and then consider a case that has no such clear empirical ties.

#### Case 7 (SLEEP)

Herb and Flo are early 20th century physiologists who specialize in the study of sleep. A comatose patient is delivered to their laboratory, and the question is posed: Is the patient is asleep? Herb argues that the patient is asleep: Sleep, he contends, is a state of unconsciousness that results from diminished brain activity, which is probably what is present in cases of coma. Flo counters that the comatose patient is not sleeping: What is characteristic of sleep, she says, is merely a different, rather than a diminished, pattern of brain activity. Since the comatose patient has a diminished pattern of brain activity (as Flo and Herb conjecture in common), that patient is not asleep.

This case is superficially similar to the target cases. Each disputant has marshaled a particular "conception" of sleep, a different way of "spelling out" what sleep consists in. One

might worry that Herb and Flo are odd physiologists for their day, since their theories of sleep tie the phenomenon to a physiological feature (brain activity) they cannot hope to measure. They might not, if pressed, even be able to say much about what “brain activity” consists in, nor, presumably, could they begin to anticipate how to go about testing claims about patterns of brain activity. But someone who takes contested concepts seriously cannot insist that these features differentiate the present case from the target cases: Gallant might not have been terribly articulate on the topic of dignity if he had been pressed in the target case (2) concerning cruelty.

In light of the paucity of evidence and conceptual material for them to work with, it seems likely that the question of whether the comatose patient is asleep is epistemically indeterminate. But it also appears that Herb and Flo are not confronting a genuinely indeterminate question. There is, we think, a determinate fact of the matter at the time of the dispute about which (if either) of these types of state sleep is, and a determinate answer to the question whether to be comatose is to be asleep.

Why does that question have a determinate answer? Here’s one familiar sort of account, due to Putnam and Kripke. “Sleep” is a natural kind term, and through the appropriate sort of reference-fixing activity it came to pick out a specific physiological state-type. In the wake of that reference-fixing activity, standard uses of the term within the language community have picked out exactly that physiological state. Subsequent neuroscience has revealed that this state is a distinctive pattern of brain activity. It has also revealed that this pattern is not found in comatose patients. But a community need not be in a position to know any neuroscience or indeed to know anything about the underlying makeup of the natural kind before it gets to use the term with referential success. The term “sleep” picked out exactly that state even before anyone could have known anything about the nature of that state, just as standard uses of “water” referred to H<sub>2</sub>O long before anyone could have known anything about the nature of water’s chemical composition. Of course, the scientists in that community still had much heavy lifting to do to reveal the underlying nature of the kind (its composition, its “essence”), or even to develop the concepts in

which an account of that nature could be framed. Nevertheless, referring to exactly that kind was a snap: You just had to be a member of a community that had the right sort of reference-fixing activity to get you started, and Presto! your words would refer.

That is one plausible and widely accepted story about how our concepts can “outstrip” our grasp of them. So then the challenge is this: Why not think that the same sort of determinacy could be present in the target cases? As we have seen, many of the concepts we would want to count as contested are moral or evaluative concepts. Perhaps the only reason why we find disputes over these concepts perplexing—and the only reason why we are tempted to characterize these concepts as indeterminate—is that we stubbornly refuse to be realists about the domain under consideration. If only we were realists about (in this case) moral properties, this objection runs, we could then explain the target cases as cases of merely epistemic indeterminacy. An act can be fair, cruel, or an apology even in cases where Goofus and Gallant cannot have any way of knowing whether it is. Granted, it’s hard to say what sort of kinds moral kinds might turn out to be. It would be presumptuous to claim without argument that moral concepts pick out acts in a way that is driven by underlying microphysical properties. Presumably cruelty is not the sort of property we should expect neurophysiologists to tell us anything about. But perhaps there are moral kinds despite our inability to frame coherent views about their underpinnings.

However, I mentioned above that there was another avenue one might take to make the argument that the target cases involve merely epistemic indeterminacy. Here’s another case in which epistemic indeterminacy is plausible, but where appeal to an “underlying kind” doesn’t involve appeal to underlying microphysical composition.

#### **Case 8 (FOLLY-WAGER)**

Consider a late 16th century English tavern where the patrons are gambling before the development of rudimentary probability theory. A term to conjure with among these folks is “Folly-Wager”. One evening Nigel wagers 10 pence for a chance at winning 20 pence if a face card is drawn from the deck: “Folly-Wager!”, his friends exclaim.



(“Should not have taken that one, Nigel!” they cry: “Wish I could get some poor mind to stake that wager with me!”) Discussion of whether a bet is a “Folly-Wager” focuses on matters that have to do only with the structure of the bet itself rather than on features that are particular to the bettor, such as his financial state or the number of mouths to feed waiting at home. (For instance, standards for Folly-Wager talk wouldn’t countenance “It is a Folly-Wager to stake your last pence”; but they would countenance a claim of the form “It’s mad to put your last pence on a Folly-Wager”.) The concept also is independent of actual outcomes: If Nigel were to win his 10-for-20 bet, his peers wouldn’t rescind their assessment that he had accepted a Folly-Wager, but would instead say that he was lucky to have ended up profiting from it. All parties accept that Nigel’s 10-for-20 wager was a Folly-Wager. But the cases get trickier. Nigel later wagers 7 pence to win 20 if he draws a face card. Disputes follow around the table over whether Nigel has made a Folly-Wager. Although they’re a bit lost in these waters, Nigel gamely tries to defend himself by pointing out that 7 pence is not very much to stake in order to win 20. Reginald, the most ardent advocate of the Folly-Wager charge, argues it is altogether too rare to draw a face card, and that hence this must be a Folly-Wager.

Again we have a case that superficially resembles the target cases. Nigel and Reginald disagree over whether the bet is a Folly-Wager, and each attempts (in especially crude fashion) to explain what it is that qualifies it as a Folly-Wager. And again the kicker is that, although our merry gamblers don’t have the conceptual apparatus required to recognize or to articulate it, it nevertheless appears that there is an objective thing that they mean by “Folly-Wager”: A Folly-Wager is a bet that has negative expected value. Later developments of probability theory, we think, has revealed exactly what these fellows were talking about when they were talking about Folly-Wagers.

The earlier challenge recurs, then, and this time without any need to appeal to natural kinds or underlying microphysical constitution. Since there is (apparently) a correct way of understanding Folly-Wager that outstrips the epistemic and conceptual capacities

of the tavern gamblers, what grounds do we have for thinking that there is no objective conception of apology, cruelty, etc. that lies beyond our grasp, at least for the time being? (And of course the objector can urge that we need not ever come to grasp them.) Suppose that civilization had been wiped out before we came to understand brain activity or probability theory. Wouldn't the concepts still have had determinate boundaries of application fixed by the (now eternally unknown) facts? If so, then it would turn out that contested concepts are merely epistemically indeterminate and do not place new demands upon our theories.

### 5.5.3 Why the strategy fails for the target cases

In my response to this challenge, I will neither attack nor defend the direct reference view. Nor do I plan to consider the prospects for, or plausibility of, the sort of moral realism I sketched above. For the sake of the present argument, I will just grant that some version of both positions is true. My aim instead will be to show that even with those positions in place, the sleep and Folly-Wager disputes are substantively different from those in the target cases—different in a way that makes the epistemic indeterminacy explanation of the target cases implausible.

Let us first get a sense of why it works when it does; then we will come back through and see why the target cases lack those attributes. The advocate of this view will want to contend that the sleep and Folly-Wager cases are examples which have the following features: (a) two distinct conceptions have been offered and (b) there is a fact of the matter which (if either) is correct. I think instead that features (a) and (b) are in tension with one another. In particular, I contend that the reason that we think that there is a(n epistemically inaccessible) determinacy to the concepts in the sleep and Folly-Wager cases is that we suppose that the parties to the dispute actually already tacitly share a conception of the relevant concept, and hence they are not involved in a *conceptual* dispute. If we give up that supposition and try to shape the dispute into a genuine conceptual disagreement—a contest between competing conceptions of the relevant concepts—then we are also led

to give up the idea that there is determinate fact of the matter whether the concept applies in the cases at hand.

To get a sense of what I mean, let us return to the sleep case. Given the description of the events in that case, it appears that despite their disagreement about the case of the comatose patient Herb and Flo actually agree upon a conception of sleep. Each of them appears, from the description of the case, to agree that sleep is to be spelled out as a natural kind concept, for which the (then as yet unknown) underlying biological nature fixes the boundaries of correct application.

Though Herb and Flo differ in their “theories of sleep”—in their views about “what it is to be asleep,” it appears to me that there is a difference between the sort of theories they are offering and what we would properly call, in the idiom of the thesis, conceptions of sleep. Why? Because in our thinking about how to assess their dispute, we treat their claims in the sleep case as grounded upon (I) their (shared) view about how to understand the concept sleep (as a natural kind concept, with its boundaries of correct application fixed by the facts of biology), in combination with (II) their (divergent) favored guesses, conjectures, or informed surmises about which biological features that concept will turn out actually picks out. Their divergent guesses in (II) lead them to different claims about what sleep “really is”. But only (I) is the conceptual rather than the empirical part: (I) is a conceptual commitment, involving taking the “constitutive stance” to underlying neurophysical constitution: It takes the position that what it is to be asleep is to be in some particular (as yet unknown) neurological state-type. The only contribution that (II) makes is empirical rather than conceptual.

How are we to distinguish the two different “conceptions” of sleep Herb and Flo offer? What we have in this case is not a difference in the content of their commitments: rather it’s a difference in the force with which they are held. Here’s how we can tell. Here’s why we think that (I) is (strictly) conceptual, whereas (II) is only indirectly conceptual, depending on an admixture of empirical claim.

Much the same can be said for the concept FOLLY-WAGER in Case (8). In thinking

about that case, we suppose that there are determinate boundaries of application of the concept only because we understand Nigel and Reginald to share the same core conceptual commitments about Folly-Wager, even though they are unable at the moment to explicitly grasp or articulate them, and that those core conceptual commitments picked out a determinate phenomenon from probability theory. To see that you expected this, you can ask yourself what you would have anticipated Nigel and Reginald to have said if you had led them through a number of less computationally demanding exercises in probability theory. You also anticipate that they would say that the following game did not involve a Folly-Wager for either party: “Two bettors (‘Red’ and ‘Black’) each stake 10 pence. A checker is blindly drawn from an urn containing 500 Red and 500 Black checkers.” You also anticipate that both Nigel and Reginald would think that the same game would be a Folly-Wager for Black if the urn contained 500 Black and 501 Red checkers. You also expect that if we were to sit them down for a quick course in the rudiments of probability theory, Nigel and Reginald would both accept this spelling-out of the concept Folly-Wager, or that they would lack sensible grounds for objecting. That is, that their present commitments about the nature of Folly-Wagers already commit them to the expected utility conception of FOLLY-WAGER.

## 5.6 CONFUSION

### 5.6.1 An explanation of the strategy

It might be thought that the disputes in the target cases are due to the disputants employing confused concepts of fairness, cruelty, or apology. Some cases of confusion—most noticeably, those in which the participants are able to spell out what they take the content of the concept under discussion to be—are in principle indistinguishable from within the context of the dispute from disputes over contested concepts. In such a case of confusion, the participants may well take themselves to be engaged in a dispute over a single con-

tested concept: they might well understand themselves to be arguing about which of two ways of spelling out or understanding a certain concept is correct or better. But in cases of confusion, unlike cases of contested concepts, the rival “conceptions” will turn out to be independently important and neither will have a greater claim to correctness. The only way to determine whether a dispute rests upon confusion or upon a contested concepts, perhaps, is to consider it from outside the context of the dispute, either retrospectively, once we determine whether these conditions are met, or from the point of view of an observer who is conceptually and epistemically equipped to diagnose the confusion.

### 5.6.2 How the strategy might work

Let me illustrate what I mean by describing a toy case of confusion which is viewed from within as a dispute over a contested concept. (This example is decidedly not intended to mirror the actual history of the development of the concepts involved.)

#### Case 9 (ACCURACY)

Smith and Jones live in the early days of rigorous measurement, and that they are jointly designing some measuring device. In their conversations about the desired features of this instrument, they assure one another that they want to make the device as “accurate” as they reasonably can. Accuracy, they and other measurement aficionados of the day agree, is a cardinal virtue in the measurement business. When prompted to explain what the concept of accuracy involves, Smith, Jones, and their cohort have a standard, textbook answer: the accuracy in a measuring device is the closeness with which the device’s readings track the measured quantity.

But suppose that to this point in their theories of measurement, the community has not yet distinguished

the fine-grainedness of a measuring device’s readings (what we outside of the dispute would call its precision)

from

the extent to which correctly measures the quantity in question (what we outside of the dispute would call its accuracy).

Now, suppose all of Smith's and Jones's assertions and inferences about accuracy have been responsive to both of the features of their devices, without having occasion to recognize that the features are distinct. Perhaps this confusion has been made possible because all prior developments in instrumentation have improved devices along one of these dimensions without sacrificing progress along the other. In the past, say, when they have modified the device in a way that improves the things fine-grainedness of measurement, Smith and Jones have exclaimed in unison, "Terrific! An improvement in accuracy! Now our measurements will track the quantity more closely!" And when they have modified the device in a way that produces an improvement in the correctness with which the device tracks the measured quantity, they have cried out, "Yet another improvement in accuracy! Now our measurements will track the quantity even more closely."

But now suppose that Smith and Jones enter into a stage of development of instrumentation at which trade-offs in development or comparisons between machines must be made in ways that are responsive to these two different features of instruments. At such a stage, Smith and Jones may well begin to disagree in our application of our concept of accuracy. Smith, for instance, might judge that a particular instrument is an extremely accurate device, since its readings are extremely fine-grained; whereas Jones might judge that it is not, since its readings are moderately unreliable. Upon reflection, if they have their wits about them, it will become clear to them that there are two distinct ways of spelling out the concept of accuracy that has been in play in their discussions—two ways of spelling out what is meant by the closeness with which the device tracks the quantity in question. Once they have distinguished these two ways of spelling out the concept accuracy, it might well be that Smith takes the fine-grainedness reading to be the (uniquely) proper spelling out of the concept accuracy, and Jones might similarly plump for spelling it out as correctness. If so, they might

well take themselves to be involved in a dispute over the contested concept of accuracy, and take themselves to be offering competing conceptions of that concept.

### 5.6.3 Why the strategy fails for the target cases

But I think that this case is properly understood as a case of confusion, and that such cases are distinct from cases of contested conceptions. Here's why.

Once we have distinguished the two distinct plausible conceptions, Smith and Jones will likely come to appreciate that each of these stands on its own, independently of the other, as important and worthy of a place within our conceptual framework. Moreover, their previous employment of the concept—due to its imprecision and roughly equal balance between the two conceptions—does not favor with either of these two conceptions. Properly understood, these two “conceptions” are not competing for survival or value—the recognition of the merits of one does not threaten our recognition of the merits of the other. The proper move for Smith and Jones to make is to distinguish the two and preserve both of these conceptions as *distinct concepts*. Perhaps they would be wise to decide to confer upon each concept a distinct term to aid in avoiding future confusion. But in any event, the dispute is distinct from a dispute over a contested concept in that it fails in this way to meet the Competition Condition from the account of contested concepts above: the conceptions turn out not to be competitive with one another.

Moreover, given the development of instrumentation as I have described it above, the dispute fails to meet the Competition Condition in another way as well. For it doesn't seem that either of these “conceptions” (or any other candidate precisification) could possibly be said to be a proper or correct or best way of spelling out what the concept of accuracy. For instance, each precisification seems to fit nicely under the slogan that we used to present the content of the concept of accuracy—the closeness of the measurement in tracking the quantity. So there would seem to be no grounds for a preference between the two readings on those grounds; Smith and Jones have just come to understand that

there are two distinct senses of “closeness” that our earlier judgments and inferences were running together. And since those judgments and inferences involving the term “accuracy” up to now were equally responsive to both of these senses of “closeness to the quantity”, our past usage and commitments seem not to provide us with grounds for declaring that one or the other was better or more correct. The measurement community would better served by going back over those judgments and inferences and recasting our earlier proclamations and inferences in light of our conceptual enrichment, determining which of the two distinct concepts properly applies in that instance.

## 5.7 TEMPORAL EXTERNALISM

### 5.7.1 An explanation of the strategy

The view I wish to consider here stems from the work of Mark Wilson and Henry Jackman. (Wilson disavows the label “temporal externalism” and some of the consequences that Jackman draws out of his work.) The analysis in these articles draws from examples where a community is imagined using a term “locally”—in a range of familiar circumstances. We are asked to imagine cases where a new item is introduced into the community, and the community seamlessly and without deliberation accommodates this new addition into its conceptual use (say, by applying the term). But then we are asked to consider an alternative course of events, in which the community encounters the new type or kind in a different setting or context and seamlessly, naturally, and without deliberation accommodates the new addition by making a different semantic move—perhaps even precisely the opposite semantic move (accepting cases that were banished in the earlier example and banishing cases that were accepted in the earlier version of the example).



### 5.7.2 How the strategy might work

Let me try here to relate the point to the sort of indeterminacy discussed in and by tweaking one of the examples originally presented in [Wilson \(1982\)](#) and discussed in [Jackman \(1999\)](#): the case of Grant's zebra.

In that historical example, the term "Grant's zebra" is introduced upon a first encounter with a group of zebras. Unbeknownst to Grant and his party, there is a morphologically distinct population of the same species dwelling nearby. At this point, it looks as if there is no fact of the matter whether these beasts fall under the term. That is, it appears that there are no facts that could determine at this stage which group the term "Grant's zebra" picks out: the morphologically distinct subspecies or the species. Looking at this concept-introducing activity from the outside, and in light of the zoological facts, we can tell that there are (at least) two incompatible ways in which the indeterminacy of the term could be resolved by future developments. If Grant's expedition were to travel next to an area where the morphologically distinct members of the species interbreed with those originally dubbed "Grant's zebra", then Grant would naturally, without reflection, extend "Grant's zebra" to cover these little beasts. However, if Grant were to travel next to an area where the subspecies do not intermingle, the term would likely not be extended in this way.

To translate this into concept-talk: Grant introduces a concept: GRANT'S ZEBRA. This concept is simply the cluster of norms that govern the employment of the linguistic term "Grant's zebra". At the outset, these norms are rather puny and underdeveloped. In fact, at such an early stage it is hard even to produce a linguistic formulation of those norms. Perhaps: #Apply term t to the members of the population that Grant has discovered#. Or: #Apply term t to these sorts of animals#. However formulated, it seems quite clear that the concept is indeterminate, in a way that comes out clearly when we consider the particular linguistic formulations that are attempts to formulate these norms (Which population did Grant discover—the species, or the subspecies? Which sort of animal are these?) Certain objects—those individuals Grant originally happened upon—fall under the concept

Grant's zebra. Other items—ham sandwiches—do not. But some objects—the undiscovered, morphologically distinct, interbreeding zebras—are neither included nor excluded by the norms that constitute Grant's zebra.

In his discussion of Wilson's example, [Jackman \(1999\)](#) quickly (and, I think, correctly) dismisses the following avenue of response:

Had Grant been presented with the full set of zoological facts immediately after his introduction of the term, he *would have* reached such-and-so a decision about the scope of the term.

Jackman doesn't deny the *truth* of such a subjunctive claim in the kind of case at hand. Rather, he denies its *relevance* for the sort of questions about meaning that he is addressing. The response considered above works only by describing a different sort of development of the facts of the case. But it is precisely the point of temporal externalism that different ways the case might develop would resolve the indeterminacy differently without "changing the meaning" of "Grant's zebra". (One of the most salutary results of temporal externalism project for my own efforts here is its suggestion that resolutions of indeterminacy of this sort do not involve changes in meaning.)

### **5.7.3 Why the strategy fails for the target cases**

Jackman and Wilson explicitly restrict their accounts from covering the sort of cases where there is explicit deliberation about which extension of a predicate is the correct one. The phenomena on which they focus their attentions are ones where different historical developments result in different extensions of a predicate because those differing developments result in different patterns of use, with all of this playing out absent any deliberation or attention from the linguistic community. Moreover, both Jackman and Wilson are concerned not with cases where the semantic issue concerns whether an entire community seamlessly and unreflectively expands the extension of a particular predicate or refrains from expanding it, depending upon those historical developments.

Nevertheless, I want to suggest that tweaking the development of Wilson's original case can point us in the direction of the sort of phenomenon that is exhibited in contested concepts.

Let us suppose that instead of moving his encampment, Grant stays in place and develops a research program around the beasts he encounters there. He develops and publishes a mass of observations about "Grant's zebras," building up a little cottage industry around them and intriguing the natural history community.

Now suppose that after a decent spell of his research, Grant is presented with the news that there is a population of morphologically distinct zebras interbreeding with the population he has studied. The question is posed: Are these newly discovered animals Grant's zebras or not? To translate it into the idiom of the thesis: We can ask, which conception of Grant's zebra is correct—the one whose norms require inclusion of the new animals or the one that requires exclusion of them? It seems right to say at this point that the question might well not be settled—the extension of the term might still be indeterminate, even if Grant happens to have a view on this off the cuff. It also seems right to say that at this stage it might not just be up to Grant to stipulate that these new beasts are, or are not, Grant's zebras. Nevertheless, it also seems right to claim that there can be principled arguments about which is the best or correct way of resolving this indeterminacy. In saying that these arguments are "principled," I think we mean that they amount to something more than mere attempts to stipulate a correct answer or to appeal to a bare intuition. These disputes will often be quite fervent, and might be carried out by theorists who think that there already is a settled, natural, and correct answer to the question and therefore that the term is not, in fact, indeterminate.

## 6.0 CONTESTED CONCEPTS AND COMPETING CONCEPTIONS: A POSITIVE ACCOUNT

In this concluding chapter, I shall attempt to tie together the lessons from the previous chapter and draw from them a more general positive account of the phenomena of contested concepts.

Before I do so, I should make it clear that the work I develop in this chapter is quite speculative and tentative. Although I believe that contested concepts are an important and interesting phenomenon, and that they have been given insufficient theoretical attention, I harbor no illusions that providing such an account will be easy or fully manageable within the space of a thesis. Ultimately, I stand by the account that I offer here and think that something like it is ultimately correct. But even if the reader is not similarly convinced by the positive view, I hope that the work up to this point is compelling at least as an advertisement for further work in this area.

I hope that the work developed up to this point has made some significant progress in understanding contested concepts. Indeed, I think that the references that I have made throughout the text to points of connection between our discussion and the desiderata go a long way toward helping us to understand the general phenomenon. However, there are two remaining major issues presented to which we must attend in order to develop a more complete positive account.

One of these remaining issues is meeting desideratum (D): providing the resolution of the paradox that emerged in Chapter 1 between the Indeterminacy and Competition conditions. This issue will be the final piece of the account in the thesis, presented in §6.3

below. However, in order to have the tools required to resolve the paradox, we must try to address the other chief remaining issue, which is a holdover from desideratum C and our discussion of W.B. Gallie's work, from §2.1. The conceptual disputes discussed to this point in the thesis have provided us with some examples of conceptual contests and has provided us with some sense of the connection between the character of those disputes and the content of contested concepts. However, I propose now to draw together those lessons and assay a more general treatment of this issue. In particular, my aim will be to try to meet the vindicatory burden—I raised in Chapter 1 and reiterated in the discussion in §2.1 of Gallie—of showing that disputes over contested concepts may be rational and genuine disputes.

## 6.1 RATIONAL AND GENUINE DISPUTES

I have contended that disputes over contested concepts, when they are conducted in an appropriate way, are principled, rationally grounded exercises in justification and reasoning, more closely akin to truth-seeking and rational inquiry than they are akin to attempts at politicking, stipulative definition, or bare conflicting reports of idiosyncratic ways of understanding a general term.

To say that disputes about contested concepts are subject to rational, impartial standards is not to deny that disputes over contested concepts will sometimes—indeed, perhaps often—fail to meet that standard. Surely some such disputes are in fact conducted with an eye toward self-interest, or conducted in bad faith or in a self-deceived way in order to shape the development of the concept in a way that suits the preferences or situations of the disputant. For instance, in target case (2) in Chapter 1 concerning the concept CRUELTY, it might be that Goofus's position in the dispute is motivated and shaped (whether consciously or not) by a reluctance to view his own mocking behavior as cruel. An account of contested concepts cannot plausibly deny that this often happens. But an account of

contested concepts does not have to show that there *aren't* any political or bad-faith or ad hoc considerations appealed to in conceptual disputes. It merely has to show (i) that there are grounds that might provide grist for the sort of rational inquiry described above, and (ii) that disputants may hold themselves and other disputants to those standards, rather than being forced to appeal strictly to selfish, prudential, ad hoc, or other generally arational or idiosyncratic considerations.

What sort of rational grounds might there be for conducting and adjudicating these disputes? Now that we have seen (in Chapters 4 and 5) some longer and more detailed presentations of these disputes, we can mine that earlier discussion for those elements upon which the participants may draw in the course of their arguments. What we have learned, I claim, is that disputes of this sort will regularly involve (i) appeals to the history of the community's employment of the concept in prior cases, (ii) consideration of hypothetical cases—either counterfactual tweakings of actual cases, or the wholesale generation of novel, imaginary cases that are understood to have certain stipulated features that are appealed to during the course of the dispute, and (iii) general principles concerning the concept, including tentative general formulations of the concept, shared sayings and general precepts and commonly held convictions involving the concept. In addition, the univocity condition on contested concepts places additional rational pressures on the disputes, in that it places a burden on the disputants to take into account and accommodate their fellow concept-users' appeals to their own prior judgments and commitments concerning the concept.

The following subsections trace through these elements in greater detail and offer an account of how they might be appealed to in conceptual contests in ways that underwrite the claim that these disputes may be rational and genuine. Then in §6.2, I argue that these patterns of argumentation we observe in disputes over contested concepts are properly understood as instances of the method of reflective equilibrium. I close the thesis by considering the consequences of this connection for the way that we understand contested concepts in general.

### 6.1.1 The rational force of prior agreement

Argumentation over contested concepts will often draw upon a rich and broad history of previous judgments involving the concept. To how see this works, let us consider again the example of the hypothetical conceptual contest over BORB from Chapter 4. The villagers in that earlier hypothetical example are in a position where there are actual instances of previous coincident judgments involving the concept BORB: there are cases where the community (largely) agrees that an act was borb, and other cases where the community (largely) agrees that an act was *not* borb. When a conceptual dispute arises out of a setting where that sort of agreement is present, as is the case for the concept BORB, argumentation over controversial cases may effectively appeal to these prior cases of coincident judgment.

But some careful attention must be paid to the *way* in which these particular cases are incorporated into the argument. How can a case on which the parties already agree provide a reason for one or the other side in the dispute to change their position regarding a case over which they disagree? If we have nothing more than bare agreement and disagreement over cases, then it is hard to see how cases of agreement can exhibit any sort of rational force in making progress in a dispute. However, one way in which these cases of agreement may become fit to play that rational role derives from one of the customary features of conceptual contests.

To see how, let us start with perhaps the simplest sort of case, one so simple it is not explicitly canvassed in the BORB example in Chapter 4, but which we may easily and quickly sketch here.

Suppose that one of the prior cases on which there was wide community agreement was one where community members agreed that an act was borb, but where it is also generally understood that the act did not involve physical pain. If there is such a case, then Poppy's detractors have an obvious line of reasoning with which they can confront those who defend the narrower conception of BORB:

You claim that (1) in order to be borb an act must inflict physical pain. However, in an earlier case, we all agreed in our judgment that (2) act  $A_1$  was borb. Surely you still accept

that (3) act  $A_1$  did not cause physical pain. (1), (2), and (3), however, are inconsistent: You have contradicted yourself! So you should give up (1) and accept instead  $\neg(1)$ : an act need not inflict physical pain in order to be borb.<sup>1</sup>

Such a line of reasoning clearly must exhibit devastating rational force upon a disputant who is committed to (1), (2) and (3), by showing that her commitments regarding a particular case are not co-tenable with the general conceptual claim she has advanced. Of course, a disputant who finds herself in that position has a number of different ways she can extract herself from it.

First, she might simply follow the resolution that her opponent proposed, upholding her earlier commitments concerning act  $A_1$  while renouncing (1) her conception of BORB. Such a move would amount to capitulation and concession: The conceptual dispute would be brought to a close, at least on this particular issue and for the time being.

Second, she might retract her earlier judgment that (2): She can claim that her earlier judgment that act  $A_1$  was borb was mistaken. In doing so, she will endorse the judgment that the act in fact was not borb, and that other community members are mistaken to think otherwise. This response leaves the conceptual dispute open and alive: The judgments regarding particular cases have been brought in line and made to cohere with the general pronouncements about the content of the concept in question, and the disputants are prepared for further argument or challenge.

Finally, the disputant might reject just (3) and reject her earlier commitment to the idea that the act caused physical pain. Of course, lest this maneuver be blatantly *ad hoc*, it will have to be accompanied by further considerations adduced to demonstrate that the act *had* inflicted physical pain. In its simplest version, these further considerations might rely on revisiting certain questions of fact, shifting the dispute (at least temporarily) from a conceptual one to an empirical one. In more complicated versions, her rejection of (3) might rely upon articulating a view about what properly counts as physical pain. In such

---

<sup>1</sup>The same strategy might be employed in the other direction, in order to defend the narrower conception of BORB, if there are cases where an act intentionally inflicted a great deal of unwarranted humiliation but where (in explicit judgments or perhaps by omission of a judgment that the act was borb) the community members appear to be committed to the claim that the act was *not* borb.



a case, the original conceptual dispute over BORB turns into a conceptual dispute over the concept PHYSICAL PAIN.

When they are available, such arguments involving straightforward judgments about earlier cases are among the most obvious and easily presented: no doubt one of the first issues that a conceptual disputant will have to address is the presence of any such straightforward incoherence between their general position in the dispute and their prior judgments about particular cases.

But as the details of the BORB dispute show, there are other, rather more baroque ways in which judgments about past cases can feature in a conceptual dispute. Recall that in the villagers' discussion, one disputant reminded the others of their general agreement that (4) borbity comes in degrees, and then reminded them of their common comparative judgment that (5) one act  $A_2$  (Shep's dragging Vic behind Vic's own horse) was more borb than another  $A_3$  (Shep's dragging Clem with Shep's horse), even though (6) acts  $A_2$  and  $A_3$  did not differ with respect to the amount of physical pain involved. This is presented as a challenge or a puzzle: What is it that explains our shared commitment to (5), given that we are also committed to (6)? The proponents of the broader conception of BORB, then, may contend that their own conception is superior to the narrower one, since its incorporation of humiliation into the concept provides an extremely plausible *explanation* of our commitment to (5) and *grounds* for our continued endorsement of it: it is the humiliation involved in being dragged behind one's own horse that makes act  $A_2$  more borb than  $A_3$ . And to the extent that the broader conception can account for this pattern of judgments while the narrower conception cannot, we have a reason to think that the broader conception is better, closer to correct, or more proper than the narrower conception.

Of course, as in the initial simple example above, there are a number of available responses to this line of argument: indeed, given the comparative complexity of the argument, there are even more potential responses, not all of which I can consider here. For instance, one might renounce the community's earlier commitment to (5), rejecting it as

a mistake. This would have the effect of undercutting the putative explanatory advantage of the broader conception by arguing for the elimination of the phenomenon that only it is supposed to explain. Alternatively, one might disavow (6) and attempt to show that the two acts did differ in the physical pain they caused, and that this is what explains the community's differential judgment about  $A_2$  and  $A_3$ . This strategy amounts to trying to argue that the narrower conception does not suffer a comparative disadvantage since it, too, can explain and justify the pattern of judgments in the cases at hand. Or one might argue that there is some other difference in the two acts—something other than a difference in physical pain or humiliation—that explains and justifies the differential judgment. Such a maneuver might well involve offering, and arguing for, a new candidate conception of BORB: one that differs from either of the two originally under consideration.

In a conceptual dispute, these earlier cases are presented in a number of different ways: in the borbity dispute, they are adduced in order to show that a commitment to applying the concept in one case leads to a commitment to applying the concept in another case.

The preceding discussion makes clear, I think, at least two ways in which agreement on previous cases can play an argumentative role in a dispute over contested concepts: either by highlighting inconsistencies between a particular conception and earlier judgments, or by providing grounds for a claim of explanatory and justificatory superiority for one conception over another. Arguments of either can exhibit a sort of rational force in such a dispute. However, we have also seen that the sort of rational force that they impose may be resisted or answered in ways that also exhibit rational force. One person's *modus ponens*, as it is sometimes said, is another's *modus tollens*. Nevertheless, it is also apparent that the choice between these two options will have further rational repercussions, in the form of forcing the thinker to adjust her other commitments in accordance with her choice. The question that remains on the table, then, is whether there is any method or vantage point from which we might ground the whole enterprise and say that one or the other of these choices is superior or correct. Formulating and addressing this question will form the heart of 6.2, later in this chapter.

### 6.1.2 The role of hypothetical cases

One curious, and likely controversial, feature of the fictitious borbity dispute concerns its participants' appeal, in the course of their argument, to judgments about hypothetical cases. Where the supply of actual cases runs out or requires further supplementation of detail, the villagers are quite content to sketch in a non-actual case and invite one another to judge whether the concept applies or fails to apply, and they welcome and encourage their fellows to do the same in return. There are some parallels between the villagers use of these examples and the use we are used to seeing in philosophical arguments. However, note that although developing and offering these hypothetical cases requires some conceptual sophistication (as any subtle argument will), it does not require any philosophical training to appreciate that they exhibit that a commitment to a particular judgment about a hypothetical case may be consistent or inconsistent with other judgments to which one is committed. Even disputants who have little experience with conceptual argument can appreciate that hypothetical examples seem to exhibit this sort of rational force in arguments of this sort.

Nevertheless, there are a number of philosophers who have raised objections to the introduction of judgments about hypothetical cases into philosophical arguments. (See, *inter alia*, [Fodor \(1964\)](#).) Given our interest in showing that disputes over contested concepts may be rational and genuine, we must turn our attention to this issue.

The objections to appeals to hypothetical cases can typically and usefully be decomposed into three parts. In the first part, the objector describes the features of hypothetical cases themselves. In the second part, she describes features of the judgments that thinkers tend to make about hypothetical cases (perhaps explaining these features in terms of the cases themselves). In the third part, the objector explains why these features render judgments about hypothetical cases unfit to play the roles in arguments for which they are pressed into service.

So, first, let's consider the features of hypothetical cases that are worrisome. Hypothetical cases are under-described: They fail to include crucial features or details that might

be relevant. The selection and the presentation or description of hypothetical cases can be tendentious: a thinker who wants to advocate for a particular conception can choose cases selectively that will favor their own conception, in the sense that they will choose cases that tempt other thinkers into accepting their favored conception. Or hypothetical cases may be packaged in a leading vocabulary by a disputant who has a particular conceptual axe to grind. Moreover, once discussants are no longer limited to the discussion of actual cases, and once stipulation is introduced as an acceptable device, this will permit the invention of and appeal to “hypothetical” cases that are physically or even logically impossible. Every philosophy instructor is familiar with the reaction of her undergraduates to some of the more bizarre hypothetical cases. We can wonder, as our students do, how these cases might play a role in any substantive investigation—even a “merely semantic” investigation.

Second: what are the troubling features of thinkers’ judgments about hypothetical cases? (As I remarked before, some of these can no doubt be understood as consequences of the first type of complaint.) First, our judgments about hypothetical cases are often extremely weak, in the sense that thinkers who contemplate them might not have especially strong intuitions one way or the other about them. Second, the judgments about hypothetical cases might be unreliable, in the following sense: a thinker’s reports about what she *would* say when she considers a case as hypothetical, do not coincide with the judgments that reporters offer when they confront similar actual cases.<sup>2</sup> Or the judgments might be unstable, in the sense that they vary over time, even in the case of a single thinker who is presented with the same case at different times. On the other end of the spectrum, there is the concern that even a sincere thinker’s judgments about a hypothetical case might be subject to a sort of theory bias. If a thinker already has formulated and endorsed a “theory” about the concept in question, that formulation might well make it more likely that she will adjudicate the case in a way that follows that formulation.

Third and finally, then, objectors conclude (usually with some gesture toward the first

---

<sup>2</sup>Much of what is interesting in recent work in “experimental philosophy” is dedicated to conducting psychological experiments that show that non-philosophers’ judgments about hypothetical cases are radically different from philosophers’ judgments.

two sort of concerns to underwrite this concern) that judgments about hypothetical cases are simply too faulty a tool for the role that they play in arguments. Here we would do well to ask what that role *is*. The objector might respond, “Well, does it matter? Assuming that my concerns above hold, then hypothetical cases and judgments about them seem altogether unfit to play much of *any* rational role.” This is a reasonable challenge, and amounts to a sort of shifting of the burden of the argument to those who wish to insist that they do have a role: Given the disturbing features of intuitions adduced above, what sort of role *could* they reasonably play? Nevertheless, we should have a sense of what sort of rational force judgments about hypothetical cases are expected to play before we decide that they are unfit for that purpose.

My strategy in response to this family of objections will be to largely concede many of the first two sorts of complaint about “thought experiments” about what we would say, but then to deny the last. Though hypothetical cases and our judgments about hypothetical cases do suffer from the sort of deficiencies and problems that objectors decry in them, they are nevertheless often sufficient to play—and indeed, they are the only candidates that are available to play—the sort of more modest role that my own theory demands of them. The argument for this claim will have to wait until we have a richer model of the kind of reasoning involved in conceptual disputes, and only appears in §6.2.

But before we get to that stage, we can get a better sense of why the objectors to hypothetical cases are troubled by considering two of the most straightforward ways of characterizing what contests over competing conceptions are supposed to discover—what role they might reasonably be expected to play. The most straightforward way we might understand their purpose would be to characterize them in the way in which those who participate in these disputes characterize them: the dispute is targeted toward getting at the *truth* about what it is for an act to be an apology, for instance. We can understand this objective in a realist sort of way, or we might understand it as simply a question about meaning or content of the corresponding linguistic term.

Understood in the former (realist) way, we might be tempted to understand the use

of judgments at hypothetical cases to serve as evidence of some empirical or metaphysical truth quite independent of our concepts.

Understood in the latter way, contemplation of hypothetical cases is a way of extracting “submerged meanings”. This is something like the hope that Peacocke has for the “offline processing” of hypothetical cases to reveal implicit conceptions: see again §2.2. On this way of understanding judgments about hypothetical cases, they are to be construed as something akin to sonar pings from which we can read off the contours of a subpersonal or implicit conception.

If these are the objectives for which judgments about hypothetical cases are to be recruited, it seems that hypothetical cases might not be of much help for either. And for each objectives, it seems that the deficiencies mentioned in the first two parts of the objection help us to see why. Our judgments about hypothetical cases just appear to be too weak, too unstable, too incoherent, and too easily defeasible to provide much in the way of evidence.

However, there is another type of role that judgments about hypothetical cases might be called upon to play, and, I claim, it is precisely this role that they are called upon to play in disputes over contested concepts. And this role, as we shall see, is one for which they are fit despite these demerits pointed out by their critics.

### **6.1.3 Another role for general principles**

The example disputes I canvassed earlier in the thesis above also demonstrate ways in which general principles play a role in the dispute, in addition to judgments about particular cases. One obvious such role for general principles to play, and which we have already seen them play in the target cases, involves their being offered as candidate conceptions of the concepts—formulations of the concept that attempt to articulate the conceptual norms more fully. In conceptual contests, we have the various parties articulating *general* claims about the correct way of spelling out a concept—e.g., that no act can be cruel in virtue of the humiliation it causes. With those general claims in place, the earlier cases

about which the parties agree can have fairly straightforward relevance to the dispute.

As an example, consider the disputants' appeal in the borbity example to the general principle that borbity admits of degrees (see §4.1 above). The plausibility of that general claim—its fit with the prior history of particular judgments—and the community members endorsement of it, places rational constraints on the conduct of the dispute by putting forward a partial articulation of the norms that constitute the concept. In particular, it places constraints on the conceptions that might be defended: no conception that entailed a flat or digital notion of cruelty could be counted as correct, unless one were willing to give up this other general principle.

So disputes over contested concepts have a place for those general principles that attempt to formulate the concept and spell it out, and for subsidiary general principles that attempt to partially articulate certain structural features of the concept. But there are other types of general principles that may play a role in the dispute, other than those that attempt to articulate or formulate any portion of that concept. There can be other general principles and observations *about* the concept that may be invoked in arguments over which conception of a contested concept is the right one. For instance, these general remarks might express not the content of that concept but rather the role that it plays within the community: its function, significance or point. We have already seen one such example (§5.1.1, p. 89) concerning sayings or bromides that concern the concept APOLOGY. As another brief specimen of this other sort of general principle, consider the example that “cruelty is the worst thing that we can do to one another.” As another example, consider the general pronouncement that, say, the point of the practice of apologizing is to promote reconciliation between aggrieved and offending parties. These general principles would, if accepted, place constraints upon the admissible conceptions of a contested concept. For instance, in the first example, no conception of cruelty that minimized or trivialized cruelty, or made it subordinate to some other sort of wronging of another, would be compatible with the first principle. And if the principle cited in the second example is accepted, then that would count in some measure against any conception of apology that

would admit the off-putting “I am sorry if I offended you” as an apology.

For a more elaborate discussion of the way in which considerations of the theoretical background might play a role, let us return to the counterfactual example discussed above in §5.7.3. In that example, a tweaking of the example from [Wilson \(1982\)](#), we suppose that rather than pressing on across the continent, the zoologist Grant has set up camp and begun to write reports about the population of zebras he has discovered. Only after some such work has been done does Grant discover that the zebras that have been the target of his observations, and which he has named “Grant’s zebras,” are a morphologically distinct subspecies of zebra. Suppose that Grant and a research associate (or rival!) are engaged in a dispute over the question whether the concept GRANT’S ZEBRA picks out the species or the subspecies: whether the broader or narrower conception of GRANT’S ZEBRAS is the correct one. Here is the sort of role that general principles might play in resolving that question. If the body of theory that has grown up around Grant’s zebras has rested heavily on the claim that Grant’s zebras constitute a species, then that would presumably count heavily in favor of the broader conception. However, if that body of theory has heavily emphasized some morphological or behavioral trait that is distinctive only of the subspecies, then that would count in favor of the narrower conception. Other sorts of general background practices and theoretical considerations can also play a role. If zoological tradition has historically favored dubbing subspecies with their discoverer’s name rather than dubbing species, that fact about scientific practice would count in favor of the narrower conception. Inquiry about the correct conception of GRANT’S ZEBRA can proceed by balancing these considerations, including discounting some earlier applications or claims as erroneous.

To see how considerations of fecundity or theoretical fruitfulness might play a role, let us consider the example of the concept NUMBER, which (I hereby tentatively claim) has been contested at various points during its history. In particular, let us consider the case of negative numbers. Some mathematicians (e.g., Carnot) were committed to the general claim that numbers were quantities and thus (since they were also committed to the general idea that there were no negative quantities) the so-called “negative numbers”



could not be numbers at all. We might even suppose that the community's outward and avowed commitment to the connection between number and quantity was greater than its commitment to the algebraic features of numbers. However, the theoretical fruits of adopting the algebraic conception of number were sufficient to dwarf concept-users explicit commitments that ran contrary to that conception. Hence the earlier claims that numbers were quantities were discounted as erroneous, and the corresponding conception of numbers as quantities was ruled out as erroneous.

#### **6.1.4 Interpersonal agreement**

Recall that we observed in our discussion of Gallie (§2.1.3.3) that in order to demonstrate the genuine and rational character of a dispute, it was not sufficient to meet a bare standard that insists only upon the consistency of an individual disputant's judgments. For some disputes that are canonical examples of arational and pointless disputes—e.g., disputes over matters of taste—are subject to that standard. The previous subsections, I believe, have put us on the path to understanding what other sorts of elements might be involved in underwriting the rational and genuine character of those disputes. One further feature of the disputes provides further grounds for this claim. That feature is the interpersonal character of these disputes. Disputants over a contested concept must be willing to take account of and give weight to not only their own judgments, intuitions, and formulations of the concept, but also to those of other concept-users. Thus, for example, the fact that other members of a community strongly favor a judgment about a particular case will exert some rational force on a concept user's own assessment of a dispute. So, for instance, when a disputant finds herself at odds with many of her competent conceptual compatriots in her judgments about controversial or hypothetical cases, or in her own attempts to formulate a conception or to argue for a restriction on the class of admissible conceptions, that would provide strong *prima facie* reasons against her own views.

The divergent judgments of others will not constitute *dispositive* or indefeasible rational force: The fact that the majority of the community thinks that the contested concept

fails to apply to a particular case in a conceptual dispute, for instance, does not, in general, entail that their employment of the concept is correct. A party to the dispute might claim that it does, but in doing so they are either (a) making an erroneous assumption about the way that the content of concepts is fixed in general, or (b) they are making a move in the conceptual dispute by asserting that this concept is one whose content is determined by a decidedly non-standard, democratic sort of way, according to which the meaning of the corresponding general term could be uncovered simply by a descriptive linguistic survey.

Why must a disputant give weight to the judgments and formulations of other parties in a conceptual dispute? Can we give an explanation of the rational force of others' commitments? No doubt it is due in part due to a general demand for intellectual modesty that is not specific to conceptual disputes: In the face of a mass of countervailing judgments, one should never simply presume that one's own view must be correct. But there is complementary explanation of the force of others' judgments which is specific to disputes over contested concepts, and which is due to the Univocality Condition from the original specification of contested concepts.

The Univocality Condition from Chapter 1 is met not because, or not merely because, the patterns of prior interpersonal agreement provide evidence for the claim that a community is using a concept in common. As we have seen, part of the reason why disputes over contested concepts are sustained is due to a conceptual community's commitment to employing the same concept in common. We have also seen that this commitment may be forsaken, and that sustained and intractable disputes might provide reasonable grounds for a break or division within a community, for the various parties to decide that they do not have a concept in common after all. But so long as the commitment to univocality is in place, it entails commitments (a) to giving credence and weight to others' judgments concerning a concept, and (b) to refraining from assigning special weight to one's own prior judgments and convictions simply because they are one's own.

It is certainly acceptable to object to other parties' views in conceptual contests and discount them as erroneous. However, in relying on that maneuver too quickly or re-

flexively, a disputant must be wary of discounting such a wide range of their opponent's judgments that it is no longer plausible to maintain that the concept is being used univocally across the community. Moreover, discounting those views carries with it a price: It entails an intellectual obligation to provide a justification of that maneuver that goes beyond merely rejecting the others' views because one happens to disagree. To say that it incurs such an obligation is not to say that the obligation must be immediately met, lest the charge be counted as false. A disputant may tentatively count others' views as erroneous for reasons she cannot quite express at the moment, in which case the obligation may merely remain on the books as debt or an acknowledged deficiency that must be repaid later. And we have seen some of the other considerations that can be invoked in meeting that obligation: those views' poor fit with the community's pattern of particular and general judgments, or with general principles that posit a plausible view of the point or function of the concept, and the like.

## **6.2 CONCEPTUAL CONTESTS AS EXERCISES IN REFLECTIVE EQUILIBRIUM**

§6.1 was intended to provide a general summary of the sort of rational and principled reasons that may be understood to have force in disputes over which conception of a concept is the correct one. In this section, I will argue that the patterns of reasoning that we have just surveyed are examples of the method of reflective equilibrium. Thus, I will argue, the proper way to construe contested concepts is by understanding them in light of the notion of reflective equilibrium: the aim or goal of disputes over contested concepts, I claim, is best understood as directed toward achieving reflective equilibrium.

What, then, is reflective equilibrium? In order to handle this question sensibly, we must distinguish between reflective equilibrium as a state and the method of reflective equilibrium, which is designed to bring that state about. Let us start with the former.

Reflective equilibrium is a state of an interrelated system of judgments. A system in reflective equilibrium is understood in an *equilibrium* state because such a system has attained a harmony, coherence, and mutual reinforcement among those judgments. A system in reflective equilibrium is understood to be in a *reflective* version of such a state because that equilibrium has been achieved via the considered, deliberative method referred to as the method of reflective equilibrium, to be described below in §6.2.2.

What does it matter whether a system of judgments is in a state of reflective equilibrium? Typically, the point of appealing to reflective equilibrium is to invoke it as a source for conferring justification upon judgments whose justification might otherwise seem mysterious or inexplicable. For example, the first apparent appeal to the notion of reflective equilibrium appears in (Goodman, 1965). Considering the question of how a rule of inference might be justified, Goodman rejects the idea that the justification of deductive inference rules must advert to their self-evidence or to “rules . . . grounded in the very nature of the human mind” (p. 63) Instead, he argues,

[p]rinciples of deductive inference are justified by their conformity with accepted deductive practice. Their validity depends upon accordance with the particular deductive inferences we actually make and sanction. If a rule yields unacceptable inferences, we drop it as invalid. Justification of general rules derives from judgments rejecting or accepting particular deductive inferences.

This looks blatantly circular. I have said that deductive inferences are justified by their conformity to valid general rules, and that general rules are justified by their conformity to valid inferences. But this circle is a virtuous one. The point is that rules and particular inferences alike are justified by being brought into agreement with each other. *A rule is amended if it yields an inference we are unwilling to accept; an inference is rejected if it violates a rule we are unwilling to amend.* The process of justification is the delicate one of making mutual adjustments between rules and accepted inferences; and in the agreement achieved lies the only justification needed for either. (Goodman, 1965, pp. 63–64)

Rawls (1971) explicitly appropriates this notion of justification from Goodman (and coins the phrase “reflective equilibrium”) and applies it to justification of principles of justice and political morality. Indeed, it is perhaps unsurprising that we should discover that method that is employed in disputes over contested concepts should fit the model of reflective equilibrium. After all, the two have a common history in Rawls’s work. Moreover, something

akin to the method of reflective equilibrium seems to be at work in Ronald Dworkin's account of legal interpretation. In the present account, however, I will argue that the method of reflective equilibrium has a role in investigating and deliberating over issues of conceptual content. Indeed, here is an ambitious and immodest way of construing the project of the remainder of the thesis: There is a method that Nelson Goodman invoked for approaching questions involving inferential rules and practices. John Rawls invoked that same method in order to approach questions involving the justification of principles of political morality. What I aim to do is to show that the same method is an appropriate fit for the kind of conceptual contest under discussion in the present work, and thereby show that this sort of enterprise has a role in theories of conceptual content.<sup>3</sup>

### **6.2.1 Structural similarities between reflective equilibrium and contested concepts**

So why should we believe that reflective equilibrium is what lies at the heart of the contested concept phenomenon? Let us begin by getting a more general characterization of the distinctive structural features of the domains in which appeals to reflective equilibrium are invoked, and consider the parallels with what we have discovered about contested concepts.

**6.2.1.1 Levels of generality** First, questions of reflective equilibrium seem to arise only for interconnected systems of judgments that stretch across a spectrum of generality. At the level of greatest generality, we have a (collection of) general principles, rules, commitments, or standards. These might be explicitly formulated, or they might be implicitly held. By contrast, at the level of greatest specificity, we have judgments about individual, concrete cases, about which the members of a community have a history of judgments, and a set of "intuitions" or judgments about hypothetical cases. In between these extremes of generality and particularity, there might be some judgments concerning an intermediate

---

<sup>3</sup>In fact—although I will not have the space to argue for it in this thesis—I believe that each of those previous areas to which reflective equilibrium has been applied can be seen as particular instances of the general phenomenon of contested concepts.

level, say, in the form of subsidiary or supplementary principles, or in the form of judgments that group together individual cases. This system of judgments is interconnected, in the sense that the general standards or rules are properly understood as having specific application to those particular cases, and the particular cases are understood as particular instances or applications of the general rule.

We have noted that the same structural feature holds for disputes over a contested concept F. On the one hand we have general rules or principles—the general concept and the various conceptions that are on offer or under consideration, which may be more or less explicit—that purport to tell us, in general, which things are F (and perhaps why they are F), and the general principles about that concept. On the other hand, we have particular cases, both actual and hypothetical, about which we have specific judgments concerning their F-ness.

**6.2.1.2 Multiple dimensions of interdependence** In the systems of judgments where reflective equilibrium is invoked, the interconnection across those levels of generality is multi-directional. If we consider the question of direction of fit or justificatory primitiveness, neither the general nor the specific judgments will have absolute priority over the other. The general judgments in a reflective equilibrium system are not absolute or foundational; none of them is self-evident or capable of justifying itself non-circularly against other competing or conflicting general principles. Nor are those general judgments “first principles” in the sense of being axioms from which particular judgments must be derived or upon which they must depend for their justification. But neither are the general principles simply summaries or generalizations that tally up and report our judgments about particular cases. The general principles and the judgments about particular cases are understood as “mutually supportive”: to the extent that they coincide or cohere with one another, our commitment to each redounds to the credit of the other.

Much the same sort of pattern of interdependence applies in questions over contested concepts. Concept-users do not understand disputes concerning general principles gov-

erning the application of the term as an issue to be confronted independently of, or prior to, questions about the employment of the concept in particular cases. Rather, as we have seen from examples of conceptual contests, disputants move back and forth between consideration of the two, paying attention to the fit and agreement between them.

**6.2.1.3 Conflicting judgments and commitments** Once we have in place an interconnected system of judgments about general principles and specific cases, this opens up the possibility of discrepancies and conflicts among these commitments. For now we have introduced the possibility of conflict: the general principles might conflict with one another (for instance, two general principles might dictate different results for a particular case), or they might conflict with the particular judgments about earlier or hypothetical cases. Of course there is nothing in the abstract structure of such a system of interconnected judgments that entails that there will be conflict: It might be that the general principles within that system are in harmony with one another and with the judgments about concrete individual cases. In such a case, the system is already in equilibrium and perhaps no further work need be done. However, in most broad and non-trivial systems of judgments, there will be conflicts among the judgments. However, in the case of contested concepts we have a system of judgments and commitments, often involving intrapersonal and interpersonal conflicts of judgment.

**6.2.1.4 Degrees of commitment or “weight”** The systems of judgment that are studied by reflective equilibrium theorists involve different degrees or levels of commitment to various judgments that they trying to bring into equilibrium. Sometimes discussions of this point in the reflective equilibrium literature express it using the metaphor of weight to express degrees of commitment to a judgment: some of the judgments in a system are weightier, in the sense that we will be more strongly committed to some of them than we are others, less willing or able or justified in abandoning some of them than we are others.

This will hold true at each of the different levels of generality. Some of the general

principles will be weightier than others. To take an example from rules of inference: It seems plausible to say that we are more strongly committed to *modus ponens* than we are to the rule of contradiction-elimination—that it would weigh more heavily against a logical system if it were to exclude *modus ponens* than if it were to exclude one of the other two. In a similar vein, we will observe that there differing degrees of commitment to our judgments about particular cases, including hypothetical cases. To use an example from Rawls’s work: We are more deeply committed to the judgment that slavery is unjust than we are to the judgment that income inequity is unjust; our slavery judgment is weightier or counts for more than our judgment about income discrepancy. And this pattern of comparison is supposed to apply *across* levels of generality as well, in a way that lends itself to comparative judgments of weight or strength of commitment between judgments involving general principles and judgments involving particular cases. Thus, the strength of a commitment to a general principle can be assessed against the strength of a commitment to a particular judgment. So too with contested concepts: as we have seen from consideration of the disputes over them, a concept-user or a conceptual community will have different degrees of commitment to the various general and particular conceptual judgments they make.

These, then, I contend are the four distinctive structural features that characterize cases where questions of reflective equilibrium are apt. And all of these features are also present in cases of contested concepts as well. With these structural features in view, we are now in a position to turn to the question of what the *method* of reflective equilibrium is.

### **6.2.2 The methods of argument: weighing, retraction, revision**

The structural features delineated above present us with a system of interconnected, conflicting general principles and particular judgments, where neither the principles nor the particular judgments have priority as such, but where each principle and judgment is understood as having some “weight.” The method of reflective equilibrium is, in part,



a method for inducing or restoring coherence between general principles and particular judgments through a process of “mutual adjustment” of the conflicting principles and particular judgments. Recall the italicized portion of Goodman’s remark about the justification of inference rules, quoted on my page [135](#):

A rule is amended if it yields an inference we are unwilling to accept; an inference is rejected if it violates a rule we are unwilling to amend. (Goodman, 1965, p. 63)

I propose the following variation on Goodman’s remark in order to express the way in which this same method—the method of reflective equilibrium—is involved in conceptual contests:

A conception is amended if it yields a verdict about a particular case that we are unwilling to accept; a judgment about a particular case is rejected if it violates a conception we are unwilling to reject or amend.

Goodman’s remark, and my own variation on it, are a little cursory. We can, and should, amplify upon them a bit.

To simplify things for ease of discussion at the outset, let’s begin by returning to the case of an argument that appeals to a single individual’s own commitments, setting aside the issue of the interplay between different parties to a dispute. Recall the “easy objection” from §6.1.1 (see p. [122](#)). The objection was raised against an individual villager who in the earlier case had judged that a particular act was borb but where the act had caused no physical pain. The objection, then, exploited the inconsistency between the general claim (that any act could only be borb if it caused physical pain) and a specific judgment about a past concrete case (that a painless but humiliating act was borb). Her interlocutor, in pointing out the conflict among her commitments, thereby shows that the system is not in harmony or equilibrium. It appears that her commitment to the general claim commits her to the claim that the earlier case was not borb (which is incompatible with her earlier particular judgment) and her earlier judgment about the particular case commits her to the claim that some borb acts do not involve physical pain (which is incompatible with the general principle she endorses). When a general principle someone endorses conflicts with

one of her judgments about a particular case, a general rational demand for consistency and coherence demands that at least one or the other of these commitments must be renounced and counted as an error.

At least one must be renounced. But which one? Of course, the interlocutor who we imagined posing this easy objection had a specific proposal for a way of returning the system to equilibrium: He proposes that the villager who was target of the objection should continue to endorse her particular judgment that this is an instance of borbity, that she should reject the general principle borbity must involve physical pain, that she should accept, more specifically, that humiliation or emotional pain is a way in which an act might be borb, and, finally, that she should accept that Poppy's branding of his children was borb. For the opponent who mounts the objection, then, this earlier case serves as a counter-example to the villager's general claim.

But as we saw earlier in §6.1.1, the conflict can also be resolved by renouncing the commitment to the earlier particular judgment: Rather than rejecting the conception because she is unwilling to change her judgment about that case, she might instead reject her judgment about the particular case because she is unwilling to reject the conception.

When we broke off our discussion of this point earlier, we wondered how, given that either maneuver would remove the inconsistency, one or the other of these two options might be better than the other, in the sense of being the more rational or the better justified. Proponents of the method of reflective equilibrium have available at least a general strategy for answering this question, based on the structural feature discussed in §6.2.1.4. An individual villager may have different degrees of commitment to the general conception and to the particular case, and one of these commitments may be said to "outweigh" the other. The method of reflective equilibrium, then, when applied to the case of an individual, tells us that an individual ought to retain the commitment to the "weightier" judgment and reject the judgment to which she is less strongly committed.

In the case of the individual villager whose commitments are inconsistent in the way described in the "easy objection" case above, this suggests that an individual should reject

or modify a conception that conflicts with her judgments about a particular case whenever she is more strongly committed to the particular judgment than she is to the conception—the general judgment.

There are some obvious complications that require comment here.

First, note that our discussion thus far has for the sake of simplicity dealt with the question of the equilibrium of an individual's judgments and commitments. Theorists who discuss reflective equilibrium typically pay less attention to this complication, in part because they pay less attention than we are paying to active disagreements among parties. It is apparent given the univocality condition—and given the failures of W.B. Gallie's program we witnessed in §2.1—that only the latter, interpersonal sort of reflective equilibrium holds promise for making sense of contested concepts.

Second, we should make certain that we appreciate the ways in which the simplicity of the “easy objection” has simplified the above illustration of the method of reflective equilibrium. (The simplicity of the slogan from Goodman fed into this oversimplification.) The chief way in which the “easy objection” was easy, recall, is that it involved a fairly localized conflict between a single general judgment and a single judgment about a particular case. Rejecting one or the other was understood as something that could be done in way that was largely insulated from the villager's other past judgments and commitments. Thus the “weighing” involved in the method could be understood as comparing the strength of the villager's commitment to the general principle and the strength of her (perhaps intuitive and unreflective) commitment to the non-borbity of Poppy's act against the strength of her commitment to her classification of that prior pain-free act as borb.

However, in most cases the network of judgments will be broader and its interdependencies more global than we have supposed in this simple case, as we can see by considering even the slightly more complicated example from the borbity case where the purported counterexample involved a commitment to the claim that the borbity of acts come in degrees. There will normally be more than just two or three different judgments that might

be renounced or judgments in order to repair the inconsistency, so there will be several different options that will have to be weighed against one another. Moreover, the rejection of one judgment will more typically require the rejection or adoption of *other* principles and judgments in order to maintain coherence. Thus it seems that the “weight” of a particular judgment within a system will depend not only on its own “weight”—the strength of a thinker’s commitment to it individually—but also upon the cumulative weights of the other judgments that must also be rejected along with it. Thus in general the pattern of “mutual adjustment” will be much broader and more complicated than it might at first have seemed.

On this point, Rawls (2001) helpfully distinguishes between different sort of reflective equilibria: “wide” and “narrow.” The distinction is one of degree, and it reflects the breadth and variety of the judgments that are under consideration in the course of the process of mutual adjustment.

### 6.2.3 Hypothetical cases revisited

In § 6.1.2, I described the concerns about the limitations and defects of judgments about hypothetical cases, and deferred discussion of the question whether these defects might be troubling in the case of disputes over contested concepts until we had a better sense of how those judgments play a role in those disputes. Now that we have on the table a richer sense of how those disputes are conducted, let us turn to that question again.

If the question is, “Should we trust our intuitions?”, the question we want to ask in return is, trust them for what purpose? We have seen that the role that judgments about hypothetical cases are supposed to play in these exercises is much the same as the role that is played by judgments involving the application of the concept to actual, prior cases: The way that they function is to elicit, display, and test a pattern of commitments to use the concept across a range of cases and to test those commitments against general principles in accordance with the method of reflective equilibrium.

As such, we can see the way in which the appeal to hypothetical cases may be justified:

they are justified in much the same way that judgments about actual cases are justified. The role that they play in conceptual disputes is to elicit (often weak, often defeasible) and display commitments about how the concept ought to be employed in these cases, to test them against general principles, and then to be folded into our system of judgments.

### **6.3 CONCEPTUAL CONCEPTS AND REFLECTIVE EQUILIBRIUM**

I am now in a position to provide a succinct statement of my proposed positive account of contested concepts.

Contested concepts are best understood as a system of judgments that are subject to justification by reflective equilibrium. To say that a community has a contested concept *C* is to say that the community members are committed to a variety of conceptual norms governing the employment of the corresponding general term “*C*,” where these conceptual norms are of varying levels of generality, and where these norms conflict with one another. These commitments will generally result from patterns of conflicting but plausible applications to prior cases.

Despite these conflicting norms, the community is also committed to the position that *C* is a single concept, used univocally over time and across different concept users, and they are committed to the position that *C* is a coherent, unified organic whole—that there is a unique best, proper, or correct way of spelling out *C*, resolving the community’s conflicting conceptual norms. They are also committed to the position that this correct way of spelling out those conceptual norms exhibits what “*C*” has meant all along, even if no community member might have been able to explicitly formulate its meaning, or even if this best way conflicts with some of their own convictions involving the concept. Moreover, the members of the community exhibit a willingness to defer to that best conception, in the sense that they understand their own judgments and pronouncements involving the concept are liable to that standard, whatever it might be. And when they engage in ratio-

nal attempts to discover that best conception, community members employ the method of reflective equilibrium.

In light of all of these observations, the most reasonable conclusion to reach is that the best, proper, or correct conception of a contested concept C is that system of conceptual norms that resolves the conceptual community's set of conceptual commitments involving C into reflective equilibrium. Or, in the event that the system of conceptual norms might be resolved into one of several available equilibrium points, then the best, proper, or correct conception will be the system of conceptual norms that resolves the community's commitments into the global maximum equilibrium point.

All of this might be the right account of contested concepts even if the community members do not understand their own conceptual disputes in this way, either because they do not have the theoretical framework to understand it in that way (say, because they are unfamiliar with the notion of reflective equilibrium), or because they explicitly reject it in favor of some other theoretical framework. The case that I have made above rests solely on the claim that the account I have is the most reasonable characterization of the conceptual commitments and behaviors of a community engaged in a dispute over a contested concept. The positive account that I have offered is vulnerable, therefore, to objections that I have failed properly to characterize the conceptual activity of a community with respect to its contested concepts, or that some other account provides a better characterization of that activity.

Of course, my positive account is vulnerable to plenty of other sorts of objections as well, including objections to the notion of reflective equilibrium itself. I turn to those below in §6.4. But first I would like to consider one further significant advantage of my positive account: that it provides us with a way of meeting desideratum (D) by resolving the apparent paradox concerning contested concepts from Chapter 1, on p. 15.

Here's a reminder of that apparent paradox. On the one hand, the Indeterminacy Condition requires that in order for a concept to be contested, there are some cases in which competent concept-users who suffer from no epistemic or linguistic deficiency may

reasonably disagree over whether the concept applies or fails to apply. Moreover, the Conception Condition requires that there are a number of distinct, plausible conceptions available for spelling out the content of that concept. However, on the other hand, the Competition Condition tells us that the candidate conceptions are competing for the status of unique correct conception of that concept, where the sort of correctness involved is understood as revealing what the content of that concept has been all along.

How, then, does the incorporation of reflective equilibrium into the account of contested concepts help to resolve this apparent paradox? It does so precisely by providing a way in which the correct conception can be settled by an issue upon which competent concept-users who suffer from no epistemic or linguistic deficiency can reasonably disagree: the issue of which conception resolves the community's pattern of conceptual commitments into the global maximum reflective equilibrium point.

#### **6.4 OBJECTIONS AND REPLIES**

Tied as it is to the notion of reflective equilibrium, the positive account of contested concepts that I have offered above stands to inherit that notion's features, favors, and faults. The last task of the thesis—the task of this section—will be to consider the potential problems that flow from that inheritance.

There have been a number of trenchant criticisms of the notion of reflective equilibrium over the years. I will have the space here to consider only three of those that are most relevant and potentially fatal to the account of contested concepts that I have developed. Two of these are what we might describe as “internal” to the notion of reflective equilibrium, and the last is what we might describe as “external” to that notion.

### 6.4.1 Internal objections

The first category raises what we might call “internal” worries about the notion of reflective equilibrium—“internal” in the sense that they are ones that might be raised by someone who is generally sympathetic to the notion of reflective equilibrium, but who nevertheless foresees critical difficulties that might well arise in an attempt to execute the method of reflective equilibrium and thwart its practitioners’ aims.

**6.4.1.1 Mistakes in the method of reflective equilibrium?** The first internal objection to the notion of reflective equilibrium focuses on the possibility of errors in the conduct of the method of reflective equilibrium. Even quite adept concept users are finite and fallible, and they are prone to make errors in judgment that will affect the outcome of the method. In fact, thinkers are often quite poor at recognizing *that* a pair of judgments conflict with one another, much less at deciding which member of the pair is “weightier” or which of the pair should be eschewed in order to achieve the maximal equilibrium point if several such equilibria are available. There is no assurance, then, that a conceptual community engaged in the method of reflective equilibrium will reach any equilibrium point at all; and if they do reach equilibrium, there is no assurance that it will be a global maximal equilibrium point. There is nothing to prevent a conceptual community from exercising the method of reflective equilibrium to the best of their abilities but nevertheless getting stuck in a local equilibrium point.

One response to this objection would appeal to the idea that a conceptual community that persistently scrutinized its own commitments for inconsistencies and which strove to attain a wide equilibrium by canvassing a broad range of hypothetical cases and candidate conceptions would eventually discover inconsistencies in its commitments that would throw them out of the local equilibrium point and back on track toward the global maximum. Perhaps this is so, but I will not argue for that claim here. Instead, I will point out that nothing in the positive account I have offered requires that the community actually progress toward achieving equilibrium in their explicit commitments. A community



might never resolve the conceptual norms to which they are explicitly and reflectively committed into any equilibrium point; so long as there is a global maximum reflective equilibrium point that *would* resolve those conceptual norms, that is sufficient on my account to constitute the correct conception of the concept.

The possibility of a community's failure to make progress toward a global maximum equilibrium point does raise one interesting complication for my positive account, but I believe that it is merely interesting and not fatal.

Suppose that a conceptual community does get stuck in a local equilibrium point rather than resolving its conceptual norms into the unique maximum point that (I claim) fixes the best conception of a contested concept. Thus, although the community will have settled upon a stable conception of the contested concept, it will nevertheless be an erroneous conception according to my account, since the conception on which they have settled is not the one at the maximal equilibrium point. However, if the community remains stuck in that local maximum, we may suppose that the community's patterns of judgments and commitments involving that concept—their continued misuse of the concept—will become reoriented around the stable local conception to which they are explicitly committed. The community might even become so strongly committed to their new system of conceptual norms through this misuse that the local maximum will become the global maximum: a shift to the prior maximum equilibrium point might no longer be an improvement, if adopting that conception requires giving up conceptual norms to which the community now has a stronger commitment. Thus, this new equilibrium point fixes a new correct conception, and is treated by the community retroactively as “what it is we have been talking about all along.”

**6.4.1.2 No unique maximum equilibrium point?** The second internal objection to reflective equilibrium raises the possibility that there might be no unique global maximum reflective equilibrium point. This might either be because there is a “tie” between the two greatest equilibrium points, or because two or more equilibrium points are incomparable

or incommensurable. Thus, given the positive account I have provided, there would be no unique best conception to which the community can defer, despite their continued conviction that there is one.

What might bring about such a state of affairs? Some of the circumstances we considered earlier in the thesis might be the cause. For instance, a tie between two equilibrium points might be due to the community's conceptual confusion: the resolution of that confusion might reveal two distinct concepts which the community was running together, each of which fit nicely with one of the equilibrium points but poorly or imperfectly with the other. Or it might be that a tie or two incommensurable equilibria points result from widely divergent, irreconcilable conceptual commitments among the members of the community; perhaps despite community members' convictions that their concept use was univocal across the different members of the community, there is simply no way that this can be sustained upon due consideration because different factions are simply too deeply committed to incompatible patterns of judgment involving that concept. In cases of this sort, we should be pleased with the verdict rendered by my proposed account of contested concepts—that there is no best conception of the concept in question.

There might be other circumstances where there might be no unique global maximum reflective equilibrium point that resolves the community's conceptual commitments. For instance, suppose that a community becomes engaged in a conceptual contest concerning a fairly novel concept: one that the community has not employed very widely at all, and regarding which the community does not yet have a robust or varied history of employing the concept. Nevertheless, the community forges ahead in its employment of the concept, committed to the notion that they are reasoning univocally and quickly founders on a controversial case where opinions diverge in a way that meets the conditions for contested concepts put forth in §1.3.

Now, in such a case it is plausible that the conceptual norms are so scarce and thin that they do not determine a unique global maximum reflective equilibrium point: scarcely any rational constraints are imposed on the method of reflective equilibrium by the scant his-

tory of prior judgments, underdeveloped and weak intuitions about hypothetical cases, and few, tentative general principles that have been introduced. Nevertheless, we may suppose that the concept-users in question remain convinced that their concept has fastened on to some unique noteworthy or salient trait, and they proceed to make judgments, assess arguments, and engage in disputes in light of that shared conviction.

The univocality commitment is what licenses the assumption of a single, best conception, and permits the continued use of contested concepts even when there is not a unique best conception. In those cases where there is unique maximum equilibrium point, a heated conceptual contest might proceed apace as the community continues to develop and enrich the history of that concept's employment. When that history of use develops to the point where there is additional or sufficient material that there is a unique maximal equilibrium point, that conception may retroactively become "the correct conception," "what we were talking about all along," and this conception may be used fruitfully to evaluate earlier judgments and inferences.

#### **6.4.2 An "external" worry**

In addition to these "internal" objections, I want to consider and answer one serious "external" objection to my reliance on the notion of reflective equilibrium in the positive account of contested concepts. By an "external" objection, I mean a worry about whether the method of reflective equilibrium itself is the right sort of technique for addressing questions about the correctness of a conception. In order to mount an external objection to the notion of reflective equilibrium, we can simply stipulate for the sake of argument that the two internal objections above can be met. So let us suppose that the method of reflective equilibrium has been carried out without error, and let us further suppose that it has led a conceptual community that has been engaged in disputes over that concept to explicitly endorse the conception which in fact resolves their earlier conceptual commitments into a genuine global maximum reflective equilibrium point. The objection I have in mind will argue that even in light of the successful execution of the method of

reflective equilibrium, it would be a mistake to identify the system of conceptual norms that provides the global maximum reflective equilibrium with the correct conception of the concept.

To raise this objection, I will have to distinguish between two different ways that the point or value of reflective equilibrium might be understood, or what sort of justification it might be understood to confer.

First, there is what we might call the “constitutive” or “constructive” use of reflective equilibrium. On this model, reflective equilibrium is strictly a self-contained enterprise: it does not involve an attempt to mirror or represent the structure of some facts, features, or properties that are independent of our concepts or practices. This way of understanding the method has intuitive plausibility for concepts like APOLOGY, for example, where the concept seems to derive from a practice that is itself the creation of, and presumably subject to the authority of, the community. Perhaps the concept of apology just amounts to nothing more than whatever the community’s considered, reflective judgments entail it must be. If disputes over contested concepts are to be understood as constitutive or constructive in this sense, then perhaps there is little to worry about here.

However, there is another, less direct way of understanding the point of reflective equilibrium—what we might call the epistemic use. The method of reflective equilibrium might be understood as conferring a sort of indirect warrant for a claim that the concept or beliefs will correctly track some property or some states of affairs that are independent of the method. Let us consider an example where the method might be employed in this way.

Suppose that a team of detectives is investigating the murder of Mr. McGillicuty and has amassed a great many interconnected judgments concerning the crime, virtually all of which point to Mrs. McGillicuty as the perpetrator. As a result, the global maximum reflective equilibrium point drawn from their current judgments leads to a strong commitment or a great deal of weight to the judgment that Mrs. McGillicuty is the murderer.

There are a number of issues that an objector might raise here about whether wide

general reflective equilibrium confers warrant or justification on the detectives' judgment that Mrs. McGillicuty is the murderer. But given the strength of the claim made in the positive account above, the objector can set those aside and simply cut to chase by making the following point:

It is one thing for the judgment that Mrs. McGillicuty is the murderer to be in wide general reflective equilibrium; and it is quite obviously something wholly other for Mrs. McGillicuty to be the murderer. We recognize that we should not identify the two in Mrs. McGillicuty's case. Coherence, even idealized coherence, is simply different from truth. Why, then, should we follow the positive account above in identifying the two when we are confronting questions like whether Goofus's caricature was cruel?

Let us suppose, for example, that our objector avers that she is a Platonist about some notion corresponding to a contested concept in her community: e.g., CRUELTY. Thus, she tells us, she will have none of our nonsense talk about past judgments, intuitions, and reflective equilibrium in order to settle the issue of whether Goofus's caricature was cruel. Goofus's act simply was cruel or it was not, just as Mrs. McGillicuty either is the murderer or she is not, and our judgments about the matter can try to reflect, but not affect or influence, the correctness of that judgment.

This objection has obvious appeal. Nevertheless, I think it is ultimately misguided when it used to address the category of contested concepts. Although it quite clear that the question of whether Mrs. McGillicuty is a murderer is independent of the strongest evidence we could muster in its favor, it is not quite so clear that the question of whether an act is cruel is in a parallel way independent of the strongest argument we could muster in its favor. The case of Mrs. McGillicuty obviously turns upon an empirical fact which the detectives simply do not know. If confronted with the question of how, ideally, with any method of inquiry, they might have investigated the question of Mrs. McGillicuty's guilt, the detectives—indeed, anyone with an interest in the case—can suggest a number of idealized methods of inquiry that do not explicitly avert to anything resembling reflective equilibrium in order to resolve the question, including simply being present to observe the event itself as the crime as committed. But there is no corresponding, idealized method of

inquiry that might be used to resolve the question of whether Goofus's caricature was cruel *other than engaging in the sort of arguments that characterize contested concepts*. What this shows, I think, is that there are good reasons for supposing that the question of whether Mrs. McGillicuty put the poison in her husband's coffee is independent of the question of what sorts of reasons we have for believing that she did, but that those good reasons are not present when we turn to the question of whether Goofus's caricature was cruel.

## 7.0 EPILOGUE: CONCLUSION AND FUTURE RESEARCH

I hope to have demonstrated that there is a philosophically important conceptual phenomenon that is exemplified in the target cases presented in §1.2 and that is specified by the four conditions developed in §1.3: the phenomenon of *contested concepts*. That phenomenon rests on the distinction between an abstract concept and *conceptions* of that concept—different ways of explicating the content of that concept. Contested concepts, then, are those concepts that are liable to what appear to be genuine, principled disputes about which of several *competing* conceptions is the correct one.

I have claimed that in order to give a successful account of contested concepts, we ought to figure out whether there is a way to take seriously the disputants' conviction that their disagreements are genuine and liable to rational standards of argumentation. I also claimed (in Chapter 5) that once we attend to the features of those disputes, we see that the contested concept phenomenon is theoretically novel with respect to theories of content: The disputes that underlie them cannot be resolved or explained by the most likely candidates currently on offer in the philosophy of mind or philosophy of language. Once we understand the kinds of considerations that exert legitimate rational force in these disputes, I claim, we will come to see that conceptual contests are best understood as appeals to reflective equilibrium, and that the best, proper, or correct conception is the one that provides the most satisfactory equilibrium point, bringing the community's conflicting commitments and judgments involving that concept into a coherent, well-justified whole.

As I have noted, a number of philosophers (most notably John Rawls and Ronald

Dworkin) have explicitly appealed to the notion of competing conceptions of a concept in their own work. Moreover, I think that many other philosophers appeal to the notion in their work without explicitly recognizing it. However, after the seminal article by W.B. Gallie (Gallie, 1962), the contested concept phenomenon has seldom been the target of sustained philosophical inquiry in its own right: Its repercussions and consequences for theories of content remain largely unexplored by philosophers of language and philosophers of mind. This is unfortunate, since (as we saw in §2.1), Gallie's article provides only an important and flawed starting point for an inquiry into this rich, important, fertile ground for research, rather than the final word on the theory. Of course, I cannot and do not claim that my remarks in this thesis offer anything remotely approaching the final word on the topic, either. This territory is vast and the terrain is rough, and what I have developed in this thesis is admittedly quite tentative and preliminary. But I hope to have made some progress in the right direction, and perhaps to have offered an advertisement for further discussion of contested concepts among philosophers of language and mind.

I would like to close the thesis with a discussion of my own aspirations for future research on this topic.

Of course, one of the chief tasks that lies immediately before me is the elaboration and defense of the tentative account I have begun to develop in §6.3. Although I have long thought that the notion of reflective equilibrium constituted the key to making sense of contested concepts, these ideas have only recently begun to take shape on the page.

I also plan to develop the technical notion of assessment of a sentence relative to a conception of one of its ingredient concepts—a notion that is sketched in the broadest possible strokes in §3.3.1.

I am also interested in surveying a wider range of concepts that I believe properly count as contested. Of course, if my earlier claims are correct, much of the philosophical work involved in applying the contested concept framework to specific problem areas is already underway; for philosophers and others have been arguing over contested concepts



for years, with varying levels of explicit appeal to that framework. Nevertheless, I think that explicit attention to the way in which the contested concept framework can be used to interpret and evaluate disputes over particular concepts cannot help but sharpen our understanding both of those particular concepts and of the framework itself. One that I am especially eager to consider is the concept of race, which seems to me to involve competing conceptions from a variety of different disciplines and settings.

I would like to turn to a couple of philosophical discussions that seem self-consciously to involve appeals to something much like the account I have begun to offer here. Two particular instances I have in mind to pursue are David Lewis's discussion of pain in "Mad pain and Martian pain" (Lewis, 1980) and Imre Lakatos's discussion of the concept POLYHEDRON in his *Proofs and refutations* (Lakatos, 1976). I do not wish to suggest that either Lewis or Lakatos would endorse my account of contested concepts or that they would accept my characterization of the methods of argument they are considering. Indeed, neither of them explicitly uses the terminology I employ here: neither of them refers to the distinction between concepts and conceptions or to the notion of reflective equilibrium. Nevertheless, the sort of conceptual questions that they confront and the methods of argumentation they employ in their discussions exemplify and illustrate the methods under discussion in the thesis and provide an illustration of the sort of rational force they can exhibit.

I am also convinced that coming to terms with the concept/conception distinction will recast and clarify a panoply of heated debates on philosophical issues that are a little further afield, some of which have been touched upon in the thesis but which require further discussion: essentialism, conceptual analysis, analytic truth, thought experiments, appeals to intuitions about hypothetical cases, and the normativity of linguistic and mental content, among others. All of these are topics I hope to pursue in future work. But I would like to end this thesis by sketching the way in which the contested concept framework might shed light upon the topic that originally launched my dissertation research.

My work on this thesis began with an enquiry into the topic of arguments from con-

ceivability to modality: specifically, inferences from the claim that it is inconceivable that  $p$  to the claim that it is impossible that  $p$ . As I began to work on that topic, I became convinced that conceivability claims were tacit appeals to a sort of conceptual modality. I was led to the topic of contested concepts in part by thinking about the sort of conceptual modality involved and about the setting in which appeals to conceivability are introduced in philosophical arguments over modality.

Here is what I tentatively concluded. The claim that it is conceptually necessary that Fs are Gs, for instance, is to be understood as a claim that the norms that constitute the concept F and the norms that constitute the concept G have the following feature: any (actual or hypothetically entertained) object that the concept F enjoins us to count as an F is such that the concept G enjoins us to count that object as a G as well. Thus, for instance, the conceptual necessity of the proposition that bachelors are unmarried consists in the fact that the norms that constitute the concept BACHELOR are such that every (real or hypothetically stipulated) object that counts as a bachelor also counts as unmarried, given the constraints imposed by the concept UNMARRIED.

The claim that it is conceptually possible that an F could be G, then, is to be understood as a claim that the norms that constitute F and the norms that constitute G are compatible, in the sense that there is some (actual or hypothetically entertained) object such that the norms that constitute the concept F permit us to count that as an F while the norms that constitute the concept G permit us to count that same object as a G. Thus, for instance, the conceptual possibility of the proposition that a woman could be President of the U.S. consists in the fact that the norms that govern “female” and “President of the U.S.” are compatible: we can coherently stipulate a hypothetical situation in which there is an object that the concepts FEMALE and U.S. PRESIDENT each permit us to count as qualifying for application of the corresponding term. And conceptual impossibility, then, amounts to the incompatibility of the concepts involved: The conceptual impossibility of a married bachelor consists in the fact that no real or hypothetically entertained object that qualifies as falling under the concept MARRIED also qualifies as falling under the concept

BACHELOR.

Understood in this way, conceptual modality will turn out to be a much more interesting and robust phenomenon than the simple and uncontroversial examples above might suggest, as we can see when we switch to considering cases involving concepts which are controversial and under dispute. It is in these sorts of cases where we will sometimes find philosophers appealing to conceivability considerations in order to underwrite claims of possibility and impossibility. In fact, it seems to me that most of the interesting cases where philosophers appeal to conceivability considerations or thought experiments turned upon contested concepts of the sort considered here. Such appeals are not, I think, reports of one's own psychological capacities or incapacities, but rather are assertions of positions within the sort of conceptual dispute considered in the present work. For instances, it seems to me that an assertion that it is inconceivable that an F could be G were tacit claims about the admissible or correct conceptions of the concepts F and G: that there is no admissible conception of F and no admissible conception of G that are co-applicable or compatible in the way described in the previous paragraph. In order to make sense of that proposal, of course, I realized that I would have to develop a more complete view of contested concepts. But as that work nears completion, I would like to return to consider its application to the topic that triggered my investigation at the outset.

## 8.0 BIBLIOGRAPHY

- Camp, Joseph L., Jr. 2002. *Confusion: a study in the theory of knowledge*. Cambridge, Mass.: Harvard University Press.
- DeRose, Keith. 1992. Contextualism and knowledge attributions. *Philosophy and Phenomenological Research* LII:913–29.
- Dworkin, Ronald. 1977. *Taking rights seriously*. Cambridge, MA: Harvard University Press.
- . 1986. *Law's empire*. Cambridge, Mass.: Belknap of Harvard University Press.
- Fodor, Jerry. 1964. On knowing what we would say. *The Philosophical Review* LXXIII(2): 198–212.
- Gallie, W.B. 1956. Art as an essentially contested concept. *The Philosophical Quarterly* 6(23):97–114.
- . 1962. Essentially contested concepts. In *The importance of language*, ed. Max Black, 121–146. Englewood Cliffs, NJ: Prentice-Hall.
- Goodman, Nelson. 1965. *Fact, fiction, and forecast*. 4th ed. Cambridge, MA: Harvard University Press.
- Hart, Herbert L.A. 1998. *The concept of law*. 2nd ed. Oxford: Oxford University Press.

- Higginbotham, James. 1998a. Conceptual competence. In Villanueva (1998), 149–162.
- . 1998b. Response to commentators. In Villanueva (1998), 193–196.
- Jackman, Henry. 1999. We live forwards but understand backwards: Linguistic practices and future behavior. *Pacific Philosophical Quarterly* 80:157–177.
- Jacob, Pierre. 1998. Conceptual competence and inadequate conceptions. In Villanueva (1998), 169–174.
- Keefe, Rosanna. 2000. *Theories of vagueness*. Cambridge Studies in Philosophy, Cambridge: Cambridge University Press.
- Lakatos, Imre. 1976. *Proofs and refutations: the logic of mathematical discovery*. Cambridge: Cambridge University Press.
- Lakoff, George. 2006a. *Thinking points: Communicating our American values and vision : a progressive's handbook*. New York: Farrar, Straus and Giroux.
- . 2006b. *Whose freedom?: The battle over America's most important idea*. New York: Farrar, Straus and Giroux.
- Lappin, Shalom, and Herbert J. Leass. 1994. An algorithm for pronominal anaphora resolution. *Computational Linguistics* 20(1):535–561.
- Lewis, David K. 1980. Mad pain and Martian pain. In *Readings in philosophy of psychology*, vol. 1, 216–222. Cambridge, MA: Harvard University Press.
- . 1996. Elusive knowledge. *The Australasian Journal of Philosophy* 74:549–567.
- Peacocke, Christopher. 1992. *A study of concepts*. Representation and mind, Cambridge, Mass.: MIT Press.

- . 1998a. Implicit conceptions, the A Priori, and the identity of concepts. In Villanueva (1998), 121–148.
- . 1998b. Implicit conceptions, understanding and rationality. In Villanueva (1998), 43–89.
- Putnam, Hilary. 1962. It ain't necessarily so. *The journal of philosophy* LIX(22):658–671.
- Rawls, John. 1971. *A theory of justice*. Cambridge, MA: Belknap: Harvard University Press.
- . 2001. *Justice as fairness: A restatement*. Cambridge, MA: Belknap: Harvard University Press.
- Rey, Georges. 1998. What implicit conceptions are unlikely to do. In Villanueva (1998), 93–104.
- Tappenden, Jamie. 1993. Analytic truth—it's worse (or perhaps better) than you thought. *Philosophical Topics* 21(2):233–261.
- Villanueva, Enrique, ed. 1998. *Concepts*. Philosophical Issues 9, Atascadero, CA: Ridgeview.
- Wilson, Mark. 1982. Predicate meets property. *The Philosophical Review* XCI(4):549–589.