

# Imperatives and Logical Consequence

Hannah Clark-Younger

A thesis submitted for the degree of  
Doctor of Philosophy  
at the University of Otago  
Dunedin, New Zealand

# Abstract

The interrelated logical concepts of validity, entailment, and consequence are all standardly defined in terms of truth preservation. However, imperative sentences can stand in these relations, but they are not truth-apt (they do not express propositions). This puzzle can be understood as an inconsistent triad:

**T1** Imperatives can be the relata of the consequence relation.

**T2** Imperatives are not truth-apt.

**T3** The relata of the consequence relation must be truth-apt.

These three claims cannot all be true. So, to solve the problem of imperative consequence we must reject one of the three claims. Solutions can be categorised into three types: those that reject T1, those that reject T2, and those that reject T3.

In this thesis, I first outline and motivate each of these three claims. I then consider, in turn, theories of imperative logic and semantics, all of which fall into one of the three types of solution. I consider and reject two versions of solution type 1, the type that rejects T1. These versions argue that it is impossible for imperatives to stand in logical relations, and attempt to provide alternative explanations for what's happening when it seems like they are doing so. I then consider and reject several versions of solution type 2, the type that rejects T2. These theories claim that, despite appearances based on surface grammar, imperatives are disguised declaratives and thus truth-apt. Each theory proposes a translation schema— it outlines the truth-conditions for imperatives. Next, I outline several versions of solution type 3, the type that rejects T3. These theories aim to develop a formal account of imperative logic. I reject each of these theories in turn, and finally propose a solution that avoids the problems I raise with the other theories.

# Acknowledgements

First and foremost, I would like to thank my series of supervisors. To Josh Parsons, thanks for getting me interested in this project, and for many conversations in pubs in Dunedin and then Oxford which helped my ideas take shape over the first three years. To Zach Weber, thanks for generously agreeing to inherit me and my project, and for helping me to refine my ideas and to sculpt them into this final form over the last year. To Colin Cheyne, thanks for joining me in the final stages and reading near-final drafts. A special thanks must go to Patrick Girard, my honorary supervisor, for working with me on the technical details of my final proposal. Finally, thanks to Charles Pigden, for doing all of these things, for being there with me from the beginning to the end, and for being a constant source of unconditional support and encouragement.

Thanks to the University of Oxford for having me for a year, and particularly to Corpus Christi College for making my visit financially possible.

Thanks to the University of Wisconsin-Madison Philosophy Department for having me for three months, and particularly to Peter Vranas for working with me and giving me immensely helpful feedback on my work. Thanks also to the Claude McCarthy Fellowship and the Fulbright Travel Award for making my trip possible.

This thesis would not exist without the generous financial contribution of the Otago University Postgraduate Scholarship and the Daniel Taylor Scholarship.

Thanks to everyone in the University of Otago Philosophy Department – you have all made it a fun journey. A particular thanks must go to (soon-to-be-Dr) Juan Gomez and (soon-to-be-Dr) Kirsten Walsh, who have been with me throughout this whole journey and without whom I would not have laughed nearly as much.

Finally, thanks to all my family and friends for being unconditionally supportive, even though I know you secretly (or not-so-secretly) wonder why on Earth anyone would do this. Thanks especially to Mum and Paul, Dad and Jo, my siblings Jemimah, Izzy, Sophie, Connor, Sam, Emily, and Harry, and of course the Steins.

# Contents

<b>Introduction</b>	<b>1</b>
<b>I Setting Up The Problem</b>	<b>5</b>
<b>1 Three Plausible Claims</b>	<b>6</b>
1.1 Imperative Arguments Exist . . . . .	6
1.2 Imperatives are Not Truth-Apt . . . . .	7
1.3 Logical Consequence is Truth-Preservation . . . . .	9
1.3.1 Implicature vs Consequence . . . . .	9
1.3.2 The Concept of Logical Consequence . . . . .	11
<b>2 The Problem of Imperative Consequence</b>	<b>17</b>
2.1 Jørgensen’s Dilemma . . . . .	17
2.2 Trilemma . . . . .	20
2.3 Solutions . . . . .	20
2.3.1 Eliminativism about Imperative Consequence . . . . .	20
2.3.2 Imperative Cognitivism . . . . .	21
2.3.3 Formal Imperative Logics . . . . .	22
2.4 The Frege-Geach Problem . . . . .	25
<b>II Eliminativism about Imperative Consequence</b>	<b>27</b>
<b>3 The Argument from Permissive Presuppositions</b>	<b>28</b>
3.1 Permissive Presuppositions and Conversational Implicature . . . . .	31
3.2 Other Valid Argument Forms . . . . .	33
3.3 Alternative Explanations for the Different Permissive Presuppositions . . . . .	38

3.4	Inference in Terms of Contradiction and Negation . . . . .	41
<b>4</b>	<b>Arguments from Rules of Grammar</b>	<b>43</b>
4.1	The Argument From Premise- and Conclusion-Indicating Words . . . . .	44
4.2	The Argument From Motivating Reasons . . . . .	48
4.2.1	Vranas's Response . . . . .	52
4.3	The Argument From The Deduction Theorem . . . . .	53
4.4	The Argument From The Grammar of Subordinate Clauses . . . . .	58
<b>III</b>	<b>Imperative Cognitivism</b>	<b>60</b>
<b>5</b>	<b>Translation Schemas</b>	<b>61</b>
5.1	Schema 1: Reports of Commands Theory . . . . .	62
5.2	Schema 2: Desires Theory . . . . .	63
5.3	Schema 3: Deontic Theory . . . . .	63
5.4	Schema 4: Predictions Theory . . . . .	64
5.5	Schema 5: Elliptic Theory . . . . .	66
5.6	Supposed Advantages of Schema 4 . . . . .	67
5.7	Cognitivist Pluralism . . . . .	68
<b>6</b>	<b>Problems for the Schemas</b>	<b>72</b>
6.1	The Problem of Unwanted Consistencies . . . . .	73
6.2	The Problem of Unwanted Validities . . . . .	76
6.3	The Problem of Soft Imperatives . . . . .	79
6.4	The Problem of Disjunctive Threats . . . . .	83
<b>IV</b>	<b>First-Order Imperative Logics</b>	<b>88</b>
<b>7</b>	<b>Early Attempts</b>	<b>89</b>
7.1	The Neustic-Phrastic Analysis . . . . .	89
7.1.1	Indicative and Imperative Sentences . . . . .	89
7.1.2	Logic . . . . .	91
7.1.3	Problems . . . . .	92
7.2	The Operator Analysis . . . . .	94
7.2.1	Syntax . . . . .	95

7.2.2	Problems . . . . .	96
7.2.3	Imperative Derivability and Consequence . . . . .	97
<b>8</b>	<b>Obedience and Termination Validity</b>	<b>100</b>
8.1	Obedience-Validity . . . . .	100
8.2	Problems . . . . .	103
8.3	Termination-Validity . . . . .	107
8.3.1	Problems . . . . .	110
8.4	Deeper Problems with both Obedience- and Termination-Validity . . . . .	110
8.4.1	Partitioning Language into “Imperative” and “Indicative” . . . . .	110
8.4.2	Validity as Nothing More than Disjunctive . . . . .	113
<b>9</b>	<b>Validity in Terms of Reasons for Obeying</b>	<b>115</b>
9.1	Terminology . . . . .	115
9.2	Types of Argument . . . . .	118
9.3	Definitions of Validity . . . . .	119
9.4	Proof Theory . . . . .	121
9.4.1	Replacement Rules . . . . .	121
9.4.2	(Pure Imperative) Inference Rules: . . . . .	122
9.5	Problems . . . . .	123
9.5.1	Truthmaker Theory . . . . .	123
9.5.2	Reasons for Acting . . . . .	124
9.5.3	Nonmonotonicity . . . . .	125
9.5.4	Validity as Disjunctive . . . . .	126
9.5.5	Counterexamples . . . . .	127
<b>V</b>	<b>Modal Imperative Logics</b>	<b>129</b>
<b>10</b>	<b>Validity in Terms of Imperative Obligation</b>	<b>130</b>
10.1	Terminology . . . . .	130
10.2	Chellas’ Formal System . . . . .	132
10.2.1	Some Results . . . . .	135
10.2.2	Similar to KD45 . . . . .	136
10.3	Chellas’ Theory on the Success Criteria . . . . .	137

<b>11</b>	<b>Functional Preprescription Semantics and KDDc4</b>	<b>140</b>
11.1	Content-Validity . . . . .	141
11.2	Preprescription Semantics . . . . .	142
11.2.1	Simple Preprescriptions . . . . .	144
11.2.2	Complex Preprescriptions . . . . .	145
11.2.3	Conditional Imperatives . . . . .	147
11.2.4	Entailment and Validity . . . . .	152
11.3	Formal System: Accessibility Semantics and KDDc4 . . . . .	154
11.4	Parsons' Theory on the Success Criteria . . . . .	154
11.4.1	Equivalence of Internal and External Negation . . . . .	155
11.4.2	Imperative Exhaustion . . . . .	157
11.4.3	Distribution of ! over Disjunction . . . . .	158
<b>12</b>	<b>Relational Preprescription Semantics and KD45</b>	<b>162</b>
12.1	Relational Preprescription Semantics . . . . .	162
12.1.1	Functional Preprescriptions Again . . . . .	163
12.1.2	Relational Preprescriptions . . . . .	163
12.1.3	RPS Interpretations . . . . .	164
12.2	KD45 . . . . .	164
12.2.1	KD45 Accessibility Semantics . . . . .	165
12.2.2	KD45 Proof Theory . . . . .	165
12.2.3	Soundness and Completeness Proofs . . . . .	166
12.3	Equivalence of RP Semantics and KD45 Semantics . . . . .	167
12.3.1	Equivalence Proof . . . . .	167
12.4	Relationship with Deontic Logic . . . . .	170
12.5	RPS/KD45 on the Success Criteria . . . . .	172
12.5.1	Non-Cognitivism about Imperatives . . . . .	173
12.5.2	Adequacy . . . . .	173
12.6	Imperative Permissions . . . . .	178
12.7	Logical Consequence . . . . .	179
12.8	The Frege-Geach Problem Revisited . . . . .	180

Conclusion	182
Bibliography	188

# Introduction

Suppose you live in King’s Landing, the capital city of the Seven Kingdoms, on the isle of Westeros. You are loyal to House Lannister, and faithful to all born of that name. In particular, you are one of Queen Cersei’s most trusted members of her Queensguard. You love her as your rightful queen, and as such, you consider anything she commands of you to be binding.

One dark and stormy night, you visit Cersei in her chamber, where you find her drinking wine and gazing thoughtfully out the window. You have a message for her: that her husband, King Robert, will be returning to King’s Landing in just a few days. She smiles, smugly. “That is good news, thank you,” she says. “And what of my brother, Jaime?” “No news, your grace,” you reply. “Ah, that’s a shame. Come, friend, join me for a cup of wine, or leave me to drink alone again tonight!” You begin to protest, and she interrupts you, adding “do not make me a lonely queen tonight!” So, you feign reluctance, fill Cersei’s cup, pour one for yourself, pull out a chair, and sit down.

You have inferred the command “join me for a cup of wine!” from the commands “join me for a cup of wine or leave me to drink alone tonight!” and “don’t leave me to drink alone tonight!” This seems like a good inference. “Join me for a cup of wine!” is, intuitively, *entailed by* the first two commands. It is, intuitively, a *logical consequence* of them.

Suppose that some months later, sometime in the final years of the long summer, Cersei travels to visit Eddard Stark and his family in Winterfell with her husband, her brother, and her three children. As always, you accompany the party as a member of the Queensguard. One rainy afternoon, shortly after you arrive in Winterfell, Cersei summons you to her chamber again. “The Stark boy, Bran, has had a terrible fall,” she says. “He is in a long sleep, and will probably not awake from it. However, I need you to be ready in case he does.” She beckons you closer, lowering her voice. “Listen, this is important: if Bran wakes up, kill him immediately!” This puzzles you somewhat, but it is certainly not your place to question the commands given to you by your queen. So, for several weeks, you keep close to

the tower where Bran sleeps, and you keep your ears open for news of Bran's condition. A few days pass, and then one bright and frosty morning you are strolling past Bran's tower, and as you pass the window you overhear one of the guards say to another "thank the Old Gods and the New, little Bran is waking up!" You know what to do. You take out your sword, climb through the window, and start swinging.

This time, you have inferred "kill Bran immediately!" from Cersei's command "if Bran wakes up, kill him immediately!" and the fact that Bran is waking up. Again, this seems like good reasoning. "Kill Bran immediately!" is, again, *entailed by* and *a logical consequence of* "if Bran wakes up, kill him immediately!" and "Bran wakes up."

However, some of these sentences are *commands*. They are in the *imperative mood*, and as such, do not express propositions. This means that they are not standardly thought to be capable of standing in the relations of entailment and consequence, because they are not truth-apt. This is, roughly, the problem of imperative consequence. The central question of this thesis is whether, and how, sentences in the imperative mood can be the relata of the logical relations of entailment and consequence.

Philosophers of logic as far back as Aristotle (1984), through to Tarski (1983), and ultimately still today (for example, Beall and Restall (2006)) have held the view that logical consequence is necessary truth-preservation. The competing modern theories of logical consequence agree that necessary truth-preservation is at least a necessary, if not a sufficient, condition. However, all of these fundamentally "truthy" definitions of logical consequence are apparently vulnerable to counterexamples involving imperatives, because imperatives appear not to be truth-apt. Thus, if these counterexamples are genuine instances of logical consequence (a claim I defend in part II of this thesis), and if imperatives really are not truth-apt (a claim I defend in part III of this thesis), then necessary truth-preservation is *not* a necessary condition for logical consequence. This goes against two and a half thousand years of wisdom on the nature of logical consequence. However, it is not as revisionary as it might appear. In section 12.7, I formulate a theory of logical consequence to replace the truth-preservationist accounts. I claim that logical consequence is the preservation, not of truth, but of *holdingness*. Holdingness is strictly more general than truth – truth is a *special case* of holdingness. I am not advocating throwing out the truth preservationist account of logical consequence and replacing it with something quite different. I am proposing that we generalise the definition, so truth-preservation becomes a special kind of holding-preservation. This is because I argue that truth is a special kind of something more general: a sentence being true is just a special way of *holding*. We should, then, no longer think

that an argument is valid whenever it is impossible for the premises to be true while the conclusion is false, but rather an argument is valid whenever it is impossible for the premises to hold while the conclusion fails to hold. Being true is one way for a sentence to hold, but it is not the only way. Imperatives can also hold, though they are never true (nor are they ever false). All of this relies, however, on taking the imperative counterexamples seriously.

In this thesis, I bring together philosophical work on imperatives from several different approaches: work on imperatives has been done in the areas of practical reasoning and inference; linguistics and philosophy of language; and philosophical and formal logic. Although these approaches appear to be unified only in the sense that each one's subject matter is imperative sentences, and although the theories seem to be answering different questions about imperatives, I show that they can be brought together as different solutions to one problem.

First, I claim that imperatives can stand in the relation of logical consequence, and point out that this provides a counterexample to the conjunction of two standard views: that of logical consequence as truth-preservation and that of imperative sentences as non-truth-apt. I am not the first to point out this problem; it goes back in a developed form as far as Jørgen Jørgensen's "Imperatives and Logic" in 1938. I do, however, formulate a new way of thinking about the problem, and emphasise that this is a problem for our underlying logical concepts, concluding that the problem is best understood as an inconsistent triad. I outline and motivate the three claims involved, and present a brief history of this problem. The problem, then, has three different types of solution which correspond to the rejection of each of the three claims in the triad.

I then consider theories of imperatives from the different areas of literature on imperatives that I co-opt as solutions to this problem and categorise as falling into Type 1, Type 2, and Type 3 solutions. Solutions of types 1 and 2 are less revisionary than those of type 3, so I consider them first.

Type 1 solutions, Eliminativisms about Imperative Consequence, claim, roughly, that imperative logic is impossible, and I draw versions of this solution from the literature on inference and practical reasoning, broadly speaking. I present, reconstruct, and reject several arguments against the possibility of imperative consequence; arguments that we are in some way mistaken that the counterexamples are genuine instances of imperatives standing in logical relations, and that this is in fact impossible.

Type 2 solutions, Imperative Cognitivism, claim that imperatives can be seen as truth-apt, and these theories come from the philosophy of language literature on the semantics of

imperative sentences. Imperative cognitivisms claim that imperatives express propositions, and the different cognitivisms differ on the translation schema that transforms the imperative into a declarative sentence with truth-conditions. This type of solution is only successful if it accurately predicts the logical and semantic behaviour of imperative sentences. I thus reject these theories by providing a series of cases in which they make predictions that, I claim, fail to capture the semantics of imperative sentences accurately.

Because attempts at the first two options have failed, I then turn to type 3 solutions, Formal Imperative Logics. These are those that attempt to formulate definitions of logical consequence and validity that include imperative as well as declarative sentences. These theories are drawn from the literature on formal and philosophical logic. I group these solutions into those that are based on first-order logic and those that are based on modal logic. I present and reject several of these theories, pointing out why they are unsatisfactory. Finally, I propose a type 3 solution that avoids the problems I raise with the other theories, and propose an accompanying definition of logical consequence as the necessary preservation of something more general than truth: holdingness.

# Part I

## Setting Up The Problem

# Chapter 1

## Three Plausible Claims

The problem of imperative consequence is, I claim, an inconsistent triad. That is, it is an inconsistency between three plausible and attractive claims. I will first briefly motivate these three claims in turn, before clarifying precisely what the problem is. I outline the problem as it was first developed, as Jørgensen's Dilemma (Jørgensen (1937)), and then show why it is more helpful to think of it as a *trilemma*. Then, I outline what the three types of solution to the problem are, and then what each type of solution needs to do to be successful.

### 1.1 Imperative Arguments Exist

As Cersei's trusted guard, you made two imperative inferences, and I claimed that they were examples of good reasoning, where the conclusions you drew were genuine logical consequences of the premises you drew them from. We can think of these as the following valid arguments:

- A1** Join me for a cup of wine, or leave me to drink alone again tonight!
- A2** Don't leave me to drink alone tonight!
- $\therefore$  **A3** Join me for a cup of wine!

- B1** If Bran wakes up, kill him immediately!
- B2** Bran wakes up.
- $\therefore$  **B3** Kill Bran immediately!

Both arguments appear to be valid, even though one or both of their premises, and their conclusions, are in the imperative mood. Argument A appears to be an instance of disjunc-

tive syllogism, and argument B appears to be an instance of modus ponens. Arguments A and B pass some intuitive tests for validity: in each case, if someone were to accept the premises, but refuse to accept the conclusion, they would be making a logical error. For example, if you were commanded A1 and A2, you would be correct in saying you had been commanded A3, even if strictly speaking you had not. Cersei has not commanded you to "join me for a cup of wine!", but she has *effectively* commanded you to, because it is *implied by* the premises, that is, by what she *has* commanded. If someone endorsed A1 and A2 by commanding them, yet refused to endorse A3 by denying that they had effectively commanded it, then they would be making a logical error. If someone agreed to obey A1 and A2, they would, again, be making a logical error if they refused to obey A3. Similarly, if someone were to endorse B1 and B2, by commanding or agreeing to obey B1 and believing B2, they would be making a logical error if they refused to endorse B3. A1 and A2 provide the right sort of support for A3, and B1 and B2 provide the right sort of support for B3; they make A3 and B3 *inescapable*, in the right sort of way. In each case, the conclusion intuitively *logically follows* from the premises. It seems to be *logically guaranteed*, or *entailed* by them. So, arguments A and B are valid. A3 is a *logical consequence* of A1 and A2, and B3 is a logical consequence of B1 and B2.

So, imperatives can be parts of valid arguments, imperatives can entail one another, and imperatives can be the relata of the consequence relation. These three logical concepts are interdefinable: An argument is valid if and only if the premises entail the conclusion, or, equivalently, if the conclusion is a logical consequence of the premises. So, I will talk primarily about logical consequence, but what I say applies to each of these three concepts, and I sometimes use them interchangeably.

## 1.2 Imperatives are Not Truth-Apt

By "imperatives sentences," and "imperatives," I mean sentences in which the main verb is in the imperative mood. These are typically used to issue commands, to make requests or even pleas, or to offer pieces of advice. The difference between these will become important for some proposed solutions, but I will deal with these differences when they become relevant. "Shut the door!" is in the imperative mood, and it expresses the command, or request, that you shut the door. It is not the only way of getting that message across, though. For example, "I'm cold!" could in some contexts be taken as a command, or at least a request, to shut the door. So too could "would you mind shutting the door?" or for that matter "were you born

in a tent?” or even a non-verbal action such as pointing at the door. In each of these cases, where the addressee understands that they have been commanded or requested to shut the door, and the speaker intends this, in a relevant sense the command or request “shut the door!” has been issued, but not uttered. So, why not include all instances of commands and requests as the subject matter of imperative consequence? Compare non-literal indicative statements such as “my mouth is on fire,” or “he tries his best at mathematics.” The first of these could be metaphorical; the second could be a nice way of saying “he is not very good at mathematics.” Logic does not deal with these “actual” or intended meanings, but rather deals with the literal meaning of sentences: it is not valid to derive “this food is spicy” from “my mouth is on fire.” Since it is these literal meanings we are interested in when we do logic, I will take as my subject matter sentences in the imperative mood, and not every instance of a command or request. I am interested in everything that is in the imperative mood, from peremptory orders to polite requests or even earnest pleas. “Please Master, don’t beat me!” entails “please Master, don’t beat me with a stick!” just as “Steward, you scum, don’t beat my slave!” entails “Steward, don’t beat my slave with a stick!” Niceties of social status and differences between pleas, requests, instructions and commands do not concern me here, as they are all in the imperative mood.

I am also interested in what it means for the relation of logical consequence to hold between (the contents of) *imperative sentences* themselves, and not between *commands*. Consider the distinction between assertions, sentences, and propositions. Assertions are speech-acts, utterances, or perhaps acts of writing. Asserting is something we *do*, and assertions are *actions*. Propositions are the subjects of these actions. We assert propositions. The exact nature of propositions is a contentious issue, and I will not enter this debate here, but will take them to be sets of possible worlds (this will not be important until the last two chapters of the thesis, and even then I am not committed to any particular view of the ontological status of these entities). Sentences are linguistic entities, made from words and letters. They are the (usual) means by which we assert propositions, which are the *contents* of (declarative) sentences. I take logical consequence to be a relation that holds between propositions, rather than sentences or assertions.

Analogously, then, commands are like assertions: they are actions. To command is to commit a speech-act. Imperatives are sentences, and they are the means by which the command is communicated. What, then, is the analogue of propositions? What are the contents of imperative sentences? This question is an important one, and one that I answer in this thesis. For now, though, we can call them *prescriptions*. Just as logic is usually

concerned primarily with propositions, we will be concerned primarily with prescriptions, but because we won't have any well-formed idea of what these are until the end of the thesis, for convenience we will focus on imperative sentences.

Standardly, imperatives are not thought to be capable of being true or false: they are not *truth-apt*. Imperative sentences do not express propositions. A sentence is true, roughly, when what it *says* is the case *really is* the case. A sentence is false, roughly, when what it says is the case *is not*, in fact, the case. This requires the sentence to *say that something is the case*. Sentences in the declarative mood, like B2, do this. B2 says that it is the case that Bran wakes up. This describes the world. It can succeed in doing so, when Bran really does wake up, or it can fail to do so, when Bran really does not wake up. Imperatives, like B3, however, do not describe the world. They are used to attempt to *change* the world, not to attempt to describe it. So, imperatives do not say that anything is the case. B3 does not *say that it is the case that* kill Bran immediately. B3 *tells you to* kill Bran immediately. Because they do not strive to describe the world, imperatives cannot succeed or fail in doing so, and so they are not truth-apt.

### 1.3 Logical Consequence is Truth-Preservation

So, we have two of our three claims: that imperatives can be the relata of the consequence relation, and that imperatives are not truth-apt. The final claim in the inconsistent triad is that logical consequence is, fundamentally, truth-preservation. I outline a brief history of this view, and show that it is true not only of classical conceptions of logical consequence, but also of (at least) the major non-classical theories of logical consequence. First, I point out what I do *not* mean by “consequence.”

#### 1.3.1 Implicature vs Consequence

Paul Grice, in “Logic and Conversation,” (1989) discussed *implicature* of sentences, which we can distinguish from consequence. The important difference is that consequence occurs when the conclusion follows from the content of the statement, whereas implicature occurs when something follows from the fact that somebody has uttered the statement. There are two kinds of implicature: conventional and conversational (Grice (1989)).

Conventional implicature is the sort that follows from the fact that someone has used one nuanced word over another with the same logical function. A clear example of this is the use of the word “but” instead of “and.” So, for instance, “Sam is intelligent, but is also

good at sports” is logically equivalent to “Sam is intelligent and is also good at sports.” In both cases, two claims are being made about Sam. However, there is a sense in which the “but” in the first sentence conveys extra (or at least different) information. Intuitively, if a person utters this sentence in favour of the “and” version, he is (in some sense) saying that it is unusual for someone to be both intelligent and good at sports, that it is a surprising fact that Sam, being intelligent, is also good at sports. This extra meaning is an example of conventional implicature. It is conventionally implicated by, but not a consequence of “Sam is intelligent, but is also good at sports” that it is unusual for an intelligent person to also be good at sports.

Conversational implicature is the sort that follows from the presupposition that the utterer of a sentence is following conversational maxims. These maxims are based on a general *Cooperative Principle*: “make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged” (Grice (1989): 26). This is broken down into specific maxims:

- Maxim of Quantity:     make your contribution as informative as is required (for the current purposes of the exchange), and do not make your contribution more informative than is required.
- Maxim of Quality:     be truthful; do not say what you believe to be false, and do not say that for which you lack adequate evidence.
- Maxim of Relation:    be relevant.
- Maxim of Manner:     avoid obscurity of expression, avoid ambiguity, be brief, and be orderly.

So, for example, if someone said, “the dairy is on either George Street or Great King Street,” then she would be violating the maxim of Quantity if she knew it was on George Street (even though what she said was true), because she would not be providing the most informative contribution that she was in a position to provide. Thus, when we presume she is conforming to the maxims, it can be assumed that she does not know which of the two streets the dairy is on. If she were to specify further, when she did not know, she would be violating the maxim of Quality, in that she would say something for which she lacked evidence. It is not *entailed by* the content of the disjunctive statement that the speaker does not know which, but it is *con conversationally implicated* that she does not.

These forms of implicature are excluded from the scope of imperative consequence. For the purpose of this thesis, I am interested in logical consequence only: in what can be

extracted from the *content* of the sentences alone.

### 1.3.2 The Concept of Logical Consequence

Standard conceptions of logical consequence hold that it is, fundamentally, *truth*-preservation (appropriately qualified). Our modern conception of logical consequence can be traced back to Alfred Tarski's seminal work "On the concept of logical consequence." However, prior to this 1936 paper, Tarski (along with many logicians in a tradition that stretches back to Aristotle) assumed a purely syntactic, or proof-theoretic, view of the underlying consequence relation. The proof-theoretic view of consequence holds that a sentence  $\phi$  is a logical consequence (in a particular formal system  $L$ ) of a set of sentences  $\Gamma$  if and only if there exists a proof of  $\phi$  that starts with  $\Gamma$  and uses only rules and axioms of  $L$ . A proof, here, is a series of steps from one string of symbols to another string of symbols, where each step must be justified either by one of the rules of inference, or one of the axioms, of  $L$ . I say "string of symbols" and not "sentence" because without a semantics,  $L$  has no interpretation and so really is just a string of uninterpreted, or meaningless, symbols.

In his 1930 paper "On some fundamental concepts of metamathematics," for example, Tarski says that "from the sentences of any set  $X$  certain other sentences can be obtained by means of certain operations called *rules of inference*. These sentences are called the *consequences of the set  $X$* " (Tarski (1983): 30). Similarly, in "Fundamental concepts of the methodology of the deductive sciences," also from 1930, Tarski says "let  $A$  be an arbitrary set of sentences of a particular discipline. With the help of certain operations, the so-called *rules of inference*, new sentences are derived from the set  $A$ , called the *consequences of the set  $A$* " (Tarski (1983): 63). These operations, or rules of inference, that Tarski refers to in these papers are the syntactic steps of a proof.

Tarski himself pointed out that this definition of logical consequence is inadequate, and gave a counterexample: Suppose that we have:

$A_0$ . 0 possesses the given property  $P$ ,

$A_1$ . 1 possesses the given property  $P$ ,

and in general every particular sentence of the form

$A_n$ .  $n$  possesses the given property  $P$ ,

where ' $n$ ' can be any natural number in a given number system. It is clear that

*A. Every natural number possesses the given property P*

is a logical consequence of the fact that we can say of each natural number that it possesses the given property  $P$ . However, this sentence “cannot be proved on the basis of the theory in question by means of the normal rules of inference” (Tarski (1983): 410). That is, no proof can be constructed from the sentences about particular natural numbers (even *all* the particular sentences about the natural numbers) to the universally quantified conclusion, because there is no rule (or series of rules) that allows this in any of the established proof theoretic systems of consequence. Thus, Tarski concluded that “the formalized concept of consequence, as it is generally used by mathematical logicians, by no means coincides with the common concept” (Tarski (1983): 411).

Of course, a formal proof system in which this universally quantified sentence is derivable from the particular sentences can easily be constructed, as Tarski also recognised. For instance, we could add “the so-called rule of infinite induction according to which the sentence  $A$  can be regarded as proved provided all the sentences  $A_0, A_1, \dots, A_n, \dots$  have been proved” (Tarski (1983): 411. This is also known as Hilbert’s  $\omega$ -rule). This rule is very different to other standard rules of inference, but there is nothing to stop a syntactic proof theory including it as a rule. This is all to say that we can include any rules we like in a syntactic proof system. We could, for example, include a rule that says that we can conclude “ $A$ ” from “if  $A$  then  $B$ ” and “ $B$ .” This rule (affirming the consequent) is not just “very different” to other standard rules of inference, it directly contradicts the common concept of logical consequence. So, there must be some constraints on which rules we allow, and which we do not, if we are to claim that our theory accurately tracks logical consequence. We must only allow rules that actually represent instances of consequence. Consequence is something more than derivability in a proof system, it is also a semantic notion.

In 1936, Tarski put forward the most influential specification of the (semantic) concept of logical consequence in the history of philosophical logic. He based his conception on that of Rudolf Carnap. Tarski thought this was the best way of understanding Carnap’s original formulation in Carnap (1937):

The sentence  $X$  follows *logically* from the sentences of the class  $K$  if and only if the class consisting of all the sentences of  $K$  and of the negation of  $X$  is contradictory (Tarski (1983): 414).

Along the same lines, Tarski’s definition of logical consequence is as follows:

The sentence  $X$  follows logically from the sentences of the class  $K$  if and only if every model of the class  $K$  is also a model of the sentence  $X$  (Tarski (1983): 417).

The sentence  $X$  can be thought of as the “conclusion” of an argument, and the class of sentences  $K$  can be thought of as the “premises” of the argument. So, this definition can be rephrased as “a conclusion is a logical consequence of a set of premises if and only if every model of the premises is also a model of the conclusion.”

What does it mean to be a model of a class of sentences, or of a sentence? Tarski specifies that we take any class of sentences,  $L$ , and “replace all extra-logical constants which occur in the sentences belonging to  $L$  by corresponding variables, like constants being replaced by like variables, and unlike by unlike” (Tarski (1983): 416-7). Call this class of resulting sentential functions  $L'$ . Then, any arbitrary sequence of objects which satisfies every sentential function of the class  $L'$  will be called a model of the class of sentences  $L$ . As a special case,  $L$  can consist of a single sentence  $X$ , in which case we can also refer to the model of the class  $L$  as a *model of the sentence  $X$* .

Of course, this hinges on what Tarski means by “satisfaction”: an assignment  $A$  of objects to variables satisfies a sentential function  $x$  if and only if taking each free variable in  $x$  as a name of the object assigned to it by  $A$  makes the function  $x$  into a true sentence (Tarski (1983): 190). Satisfaction, it seems, reduces fundamentally to truth.

For illustration, Tarski considers the simplest case: where there is just one variable. In this case, we can say of each and every object that it satisfies a function  $x$  or that it does not satisfy  $x$ . That is, substituting each object into the variable produces either a true or a false sentence. To this end, consider the following scheme, where  $a$  is an object):

*for all  $a$ ,  $a$  satisfies the sentential function “ $x$ ” if and only if  $p$*  (Tarski (1983): 190).

and then we substitute for “ $p$ ” the sentential function, “ $x$ ,” with “ $a$ ” substituted for the free variable. For example, we can obtain the following formulation:

*for every  $a$ , we have:  $a$  satisfies the sentential function “ $x$  is white” if and only if  $a$  is white* (Tarski (1983): 190).

Then, using this formulation, we can conclude, for example, that snow satisfies the function “ $x$  is white,” because the sentence that results from substituting “snow” for “ $x$ ” (“snow

is white”) is true. Alternatively, grass does not satisfy the function “ $x$  is white,” because the sentence “grass is white” is false.

In the general case, where the sentential function has an arbitrary number of variables, we have the following schema:

*the sequence  $f$  satisfies the sentential function  $x$  if and only if  $f$  is an infinite sequence of classes and  $p$  (Tarski (1983): 192).*

This works in basically the same way as in the simple single-variable case, except now we need an infinite sequence of classes for the sake of uniform expression. Still, though, truth is bound up in Tarski’s definition of satisfaction, so it is also bound up in his definition of logical consequence. He specifies a necessary condition for the sentence  $X$  to be a consequence of the class  $K$ :

If, in the sentences of the class  $K$  and in the sentence  $X$ , the constants—apart from purely logical constants—are replaced by any other constants (like signs being everywhere replaced by like signs), and if we denote the class of sentences thus obtained from  $K$  by “ $K'$ ,” and the sentence obtained from  $X$  by “ $X'$ ,” then the sentence  $X'$  must be true provided only that all sentences of the class  $K'$  are true (Tarski (1983): 415).

Tarski’s formulation of logical consequence is informed by the intuition that whenever a sentence  $X$  follows logically from a class of sentences  $K$ , “it can never happen that both the class  $K$  consists only of true sentences and the sentence  $X$  is false” (Tarski (1983): 414). Furthermore, he adds, “since we are concerned here with the concept of logical, i.e. *formal*, consequence, and thus with a relation which is to be uniquely determined by the form of the sentences between which it holds, this relation cannot be influenced in any way by empirical knowledge of the objects to which the sentence  $X$  or the sentences of the class  $K$  refer” (Tarski (1983): 414-5). Tarski’s formulation is still the most influential definition of logical consequence, and has become immortalised as *Tarski’s thesis*: that an argument is valid when it preserves truth in every model. This standard conception of logical consequence holds that consequence is, essentially, truth-preservation.

Truth-preservation alone is not enough, of course. As John Etchemendy pointed out, “for an argument to be genuinely valid, it does not suffice for it to have a true conclusion or a false premise, for it simply to “preserve truth.” The truth of the premises must somehow *guarantee* the truth of the conclusion” (Etchemendy (1990): 82). That is, it must preserve

truth *necessarily*. It isn't enough that the conclusion is *actually* true or one of the premises is *actually* false, it must be so that in every case where the premises are all true, it is *impossible* for the conclusion to be false. In *every possible case* it must be that either the conclusion is true or one of the premises is false.

Tarski was not the first to put forward a semantic conception of logical consequence. Bernard Bolzano, for example, held a very similar view to that of Tarski nearly a century earlier (Bolzano (1837)). Bolzano's account differs from Tarski's account in that he employs the notion of *substitution* instead of satisfaction. According to Bolzano, sentence is logically true if and only if it is true, and it will remain true even if we substitute any of the variable terms for other terms of the same kind. For Bolzano, a sentence  $\phi$  is a consequence of a set of sentences  $\Gamma$ , then, if and only if the conditional sentence "if  $\Gamma$  then  $\phi$ " is a logical truth (Etchemendy (1990): 28-30).

Going back further in time, Alexander Broadie (1993) claims that, among medieval logicians (who referred to the premises of an inference as the "antecedent" and the conclusion as the "consequent"), "it was commonly stated that an inference is valid if it is impossible for the antecedent to be true without the consequent also being true" (Broadie (1993): 88). Going even further back in time, even Aristotle, who is the quintessential proof theorist, held a version of the thesis that consequence is necessary truth-preservation: a syllogism "is a discourse in which, certain things being stated, something other than what is stated follows of necessity from their being so" (Aristotle (1984): I.i.24<sup>b</sup>19-20). So, Tarski's conception has a long history, but his is the most influential today.

Not everyone today adheres to a Tarskian conception, however. There are also several *non-classical logics*, which specify different conditions on what makes an argument valid. For example, relevant logics specify a further condition: that the antecedents and consequents of implications must be relevantly related. Many-valued logics allow more than just two truth-values to be specified, and modify their respective definitions of logical consequence accordingly. Intuitionists claim to reject truth as the thing to be preserved in their definition of consequence, replacing it with *constructive provability* (Dummett (1977): 9). However, as a constructive proof of  $p$  is a (particular kind of) proof *that  $p$  is true*, it is not a genuinely "non-truth-centric" conception of logical consequence. Michael Dummett even specifies that "a formula is constructively valid if *it comes out true under every internal interpretation* given in terms of constructive functions and completely defined species" (Dummett (1977): 259, my emphasis).

Preservationists, also, seem to reject *truth*-preservation in favour of the preservation of

*levels of consistency*, where a “level” of consistency is defined in terms of “how finely a set must be divided before all parts are consistent” (Payette and Schotch (2009): 93-4). However, the definition of the consequence relation, at least as given by Payette and Schotch (2009), is not really a definition as much as a definition *schema*, where the relations of  $\vdash$  and  $\Vdash$  are indexed to an underlying logic (Payette and Schotch (2009): 88). If the underlying logic has a consequence relation that is essentially truth-preservation, the corresponding *derived* preservationist consequence relation will inherit this feature (albeit with some further specifications with respect to levels of consistency).

JC Beall and Greg Restall generalise Tarski’s thesis to cover this variation in definitions. They define the *Generalised Tarski Thesis* (where the subscript “ $x$ ” indexes both validity and allowable cases to a specific logic):

An argument is valid <sub>$x$</sub>  if and only if in every case <sub>$x$</sub>  in which the premises are true, so is the conclusion (Beall and Restall (2006): 29).

That is, an argument is valid (in logic  $x$ ) if and only if in every case (of type specified by logic  $x$ ) in which the premises are all true, the conclusion is also true. Beall and Restall hold that logical consequence is a matter of the preservation of truth in all cases, where the word “cases” is neutral— it stands for any specific logic’s account of models, or worlds, or situations, or whatever plays this role. An argument is valid (in logic  $x$ ) when there is no counterexample to it: that is, there is no case (of an  $x$ -kind) in which the premises are true and the conclusion is not true (Beall and Restall (2006): 23).

So, definitions of logical consequence all fundamentally reduce to truth-preservation (in all cases). Because it is *truth* that the relation of logical consequence preserves, the relata of the consequence relation must be truth-apt.

We have now seen some examples of apparent cases of imperative consequence that provide strong *prima facie* evidence that imperatives can be the relata of the consequence relation. We have also seen that, standardly and plausibly, imperatives are not truth-apt. Finally, we have seen that standard conceptions of logical consequence all hold that it is, essentially, (necessary) *truth-preservation*. As we will see in the next chapter, these three claims constitute the problem of imperative consequence.

## Chapter 2

# The Problem of Imperative Consequence

### 2.1 Jørgensen's Dilemma

Jørgen Jørgensen put forward the most influential set-up of the problem of imperative consequence in his 1938 paper “Imperatives and Logic.” In this paper, Jørgensen first outlines the usually accepted definition of “inference” as a particular process of thought, which leads from one or more judgements to a further judgement that can be guaranteed to be true, given the truth of the former judgements (Jørgensen (1937)). Everything he says also applies to concepts of logical consequence and validity, which are more fundamental and more important as they are the underlying relations of inference (at least, of good, or licensed, inference). The condition for the validity of an argument from premises  $p_1, \dots, p_n$  to a conclusion  $q$ , is that it is impossible for  $p_1, \dots, p_n$  to all be true and  $q$  false. That is, there is no *counterexample* to the argument, where a counterexample is a possible state of affairs in which  $p_1, \dots, p_n$  are all true and  $q$  is false. Conversely, an argument is *invalid* when it *does* have a counterexample. According to this condition, arguments can only be valid (and they can only be invalid) when the premises and the conclusion are all sentences that are capable of being true, because if they cannot be true, then they cannot fulfill either of these conditions.

Jørgensen agrees that imperatives cannot be true or false – to ask of “be quiet” or “do your duty” whether it is true or false would be an ill-formed question. As Jørgensen says, “these two commands may be obeyed or not obeyed, accepted or not accepted, and considered justified or not justified, but to ask whether they are true or false seems without any sense as well as it seems impossible to indicate a method by which to test their truth or falsehood” (Jørgensen (1937): 289). It makes no sense to *ask* of an imperative whether it is true or false,

and also there is no way of *checking* whether it is true or false. So, imperatives are incapable of being true or false; they are not truth-apt. Thus, according to the generally accepted definition of validity, imperatives can be neither conclusions nor premises of arguments.

On the other hand, examples can be given of instances that look like clear examples of valid arguments. Arguments A and B from chapter 1 are two examples. Jörgensen's examples are (see Jörgensen (1937): 290):

**C1** Keep your promises!

**C2** This is a promise of yours.

∴ **C3** Keep this promise!

and:

**D1** Love your neighbour as yourself!

**D2** Love yourself!

∴ **D3** Love your neighbour!

In Argument C, the first premise and the conclusion are in the imperative mood, and in Argument D, both the premises and the conclusion are in the imperative mood. Yet, the conclusion is just as *inescapable* as the conclusion in any valid argument containing only sentences in the indicative mood (Jörgensen (1937): 290). That is, it is intuitive that if the premises were accepted, the conclusion could not be denied. So, they are intuitively valid, but they cannot be valid according to standard definitions of validity, because “keep your promises,” “love your neighbour” and so forth are not capable of being true or false; they are not truth-apt.

So, Jörgensen arrives at his dilemma. Although he uses the term “inference,” what he says is more apt if we take him to mean “valid argument”:

According to a generally accepted definition of logical inference only sentences which are capable of being true or false can function as premisses or conclusions in an inference; nevertheless it seems evident that a conclusion in the imperative mood may be drawn from two premisses one of which or both of which are in the imperative mood (Jörgensen (1937): 290).

Alf Ross (1944) gives an often cited formulation (for example, in Rescher (1966): 75):

According to the usually accepted definition of a logical inference, an imperative is precluded from being a constituent part of such inference. Nevertheless instances

may be given of inferences the logical evidence of which seems obvious in spite of the fact that imperatives form part of them. (Ross (1944): 32).

So, the dilemma here is between the following two claims:

**JD1** Validity requires the premises and conclusions of arguments to be truth-apt, so there are no valid arguments containing imperatives.

**JD2** There are valid arguments containing imperatives.

Most solutions to the problem of imperative consequence have been to reject either (the first part of) JD1, and formulate a formal system of logic for imperatives that does not require the premises and conclusions of arguments to be truth-apt, or to reject JD2 and explain why and how there cannot possibly be imperative arguments. So, the debate has generally been about whether there could be valid imperative arguments (for example, the argument given by Bernard Williams (1963), which I discuss in chapter 3) and the arguments given by Gary Wedeking (1970), Jonathan Harrison (1991), and Jorg Hansen (2008), which I discuss in chapter 4), and (among those who thought there could), how exactly imperative validity and entailment should be formalised (for example, see the imperative logics of R. M. Hare (1952) and Hofstadter and McKinsey (1939), which I discuss in chapter 7, the imperative logics of Jack Smart and Charles Pigden (unpublished) and Nicholas Rescher (1966), which I discuss in chapter 8, the imperative logic of Peter Vranas (2008, 2010, 2012, 2013), which I discuss in chapter 9, the imperative logic of Brian Chellas (1971), which I discuss in chapter 10, and the imperative logic of Josh Parsons (2012, 2014), which I discuss in chapter 11).

Jørgensen's two claims, that validity requires the premises and conclusions of arguments to be truth-apt, and that there are valid arguments containing imperatives, are only inconsistent if (that is, there is no dilemma unless) we assume a third claim, that imperatives are not truth-apt. We can keep them both if we do not mind claiming that imperatives are capable of being true or false, as several theories of imperatives have proposed (see chapter 5).

It must be stressed that this claim is just as important for producing the inconsistency. Without it, the first part of JD1 (validity requires the premises and conclusions of arguments to be truth-apt) does not get us to the second part (there are no valid arguments containing imperatives). Without it, there is no contradiction with JD2 (there are valid arguments containing imperatives), because the first part of JD1 does not contradict JD2.

## 2.2 Trilemma

I propose, therefore, that we treat the problem not as a dilemma, but as an inconsistent *trilemma* of three plausible claims, each of which I have already motivated in sections 1.1, 1.2, and 1.3, respectively:

**T1** Imperatives can be the relata of the consequence relation.

**T2** Imperatives are not truth-apt.

**T3** The relata of the consequence relation must be truth-apt.

These three claims cannot all be true, they are jointly inconsistent. Any two of the three are consistent with each other and entail the negation of the third. This, I claim, is the problem of imperative consequence. There are three claims, all of which seem plausible and attractive, but which together are inconsistent.

## 2.3 Solutions

To solve the problem, then, we must deny one of the claims of the trilemma. We can group solutions into three kinds, each corresponding to the rejection of one of the claims. We could deny T1, that imperatives can be the relata of the consequence relation, and instead claim that imperatives *cannot* be constituent parts of valid arguments. I call this view “Eliminativism about Imperative Consequence.” We could, instead, deny T2, that imperatives are not truth-apt, and instead claim that they are capable of being true or false, that they have truth-conditions. This view can be called “Imperative Cognitivism.” Finally, we could deny T3, that the relata of the consequence relation must be truth-apt, and instead reformulate our definition of logical consequence. This view requires developing an alternative formal system of imperative logic.

### 2.3.1 Eliminativism about Imperative Consequence

The first type of solution rejects T1 of the inconsistent triad, and instead holds that imperative consequence is impossible. Type 1 solutions must offer a convincing argument that apparent cases of imperative consequence, such as arguments A and B, are not genuine cases of logical consequence.

One way of arguing for eliminativism about imperative consequence is to appeal to the trilemma of solutions, and claim that rejecting T2 and T3 are both far more revisionary options than rejecting T1. T2 and T3 are both well-established claims, whereas T1 is not, so we should reject T1 as it will cause the smallest stir. Although it is perhaps true that rejecting these instances of imperative consequence as not genuine cases of consequence is the least revisionary solution, the examples provide strong *prima facie* evidence that imperatives can be the relata of the consequence relation. Thus, more argument is needed as to why the examples should not be considered as genuine cases of imperative consequence.

There have been several such arguments offered, and I will outline and reconstruct these arguments in chapters 3 and 4. In chapter 3, I discuss an argument that claims that imperatives are “essentially discontinuous,” because they always have conflicting permissive presuppositions. For this reason, they cannot be accumulated into a set of premises, and thus they cannot be taken together to imply anything. In chapter 4, I discuss a series of arguments that claim that imperatives cannot be parts of valid arguments because of facts about grammar. That is, imperatives cannot, for example, follow premise-indicating words (words that indicate that what follows is a premise), such as “because.” I outline and reconstruct these arguments, and conclude that they are unsuccessful.

### 2.3.2 Imperative Cognitivism

The second type of solution rejects T2 of the inconsistent triad, and instead holds that imperatives *do* have truth-values. Adherents of this view must explain how this is, given they do not appear to describe the world. These theories claim that imperatives, despite appearances based on surface grammar, do in fact have *propositions* as their contents. Type 2 solutions must give a plausible and principled account of the truth-conditions of imperative sentences. That is, the theory in question must provide some account of which proposition any given imperative expresses. This account must be principled: it must give us a translation *schema* of how to find out which proposition is expressed by an imperative sentence, rather than just specify them on a case-by-case basis. It must also be plausible. In particular, the translation schema must *get the logic right*. That is, it can’t offer a translation that yields results that grossly contradict our considered intuitions about what imperatives are like and how they behave.

There have been several such translation schemas proposed, and I will outline them in chapter 5 and discuss them in chapter 6. The translation schemas are: the Reports of Commands theory, which claims that imperatives such as “do x!” express the proposition “I

command that you do x,” the Desires theory (“I want you to do x”), the Deontic theory (“you ought to do x”), the Predictions theory (“you will do x”), and the Elliptic theory (“you will do x or else...”). I respond, in chapter 6, that these theories all have several problems: they each suffer from some combination of: the problem of unwanted consistencies, the problem of unwanted validities, the problem of soft imperatives (such as requests), and the problem of disjunctive threats.

### 2.3.3 Formal Imperative Logics

The third type of solution rejects T3 of the inconsistent triad, and so holds that the relation of the consequence relation need not be truth-apt. As we have seen, T3 just follows from standard definitions of logical consequence and validity. So, this solution requires redefining these logical concepts to include imperatives as well as declaratives, or to include *prescriptive* content as well as *propositional*.

I assess each theory on the following criteria, based on those specified by Parsons (2013). Any criterion of logical consequence must have these four virtues, to be deemed successful:

<b>Conservative</b>	The declarative-only fragment of the logic is classical: we get all the usual results when imperatives are not involved.
<b>General</b>	The theory applies the same criterion of logical consequence to arguments that consist entirely of imperatives, those that consist entirely of declaratives, and those that consist of a combination of imperatives and declaratives.
<b>Adequate</b>	The theory gets the logic right: it does not say there is consequence when there is none, or that there is none when there is.
<b>Non-cognitivist about imperatives</b>	The theory must have an interpretation (semantics) that makes declaratives, but not imperatives, truth-apt.

The logic must be conservative, not because I have any particular commitment to classical logic, but for simplicity; in fact, it would be a bonus if a successful theory of imperative consequence could be generalised to an imperative version of any non-classical logic. However, non-classical logics deviate from classical logic orthogonally to the way that imperative logic deviates. That is, they deviate in various ways, but are all truth-centric. Imperative logic, on the other hand, generalises from the truth-preservation case, but can retain the other elements of the definition of logical consequence, including the variations in these elements.

For simplicity, then, it is desirable to deal with one deviation at a time so as to avoid confusion. I see no reason to suppose that my final solution, for example, cannot be modified in any number of ways that are analogous to the modifications to standard (that is, non-imperative) classical logic that produce the various non-classical logics. This is because the aspect of the standard classical view that I propose to change is distinct from any of the aspects that non-classical logics change. My theory, then, I believe, is more general than just classical or any non-classical logic: it runs parallel to all of them. So, the reason for the conservativeness criterion is not a commitment to classical logic, but rather a way of ensuring that we do not, for example, lose modus ponens or disjunction introduction *just because* we have modified an unrelated part of the logic (made the consequence relation non-truth-centric). The aim is to *extend* our existing logic, not revise it unnecessarily.

The logic must be general, because if it does not apply the same criterion of logical consequence to imperative, declarative, and mixed arguments, then it must have (at least) two distinct criteria of logical consequence. In this case, it is difficult to see how they would both (or all) be defining the (same) concept of *logical consequence*. If there are two or more distinct definitions of logical consequence, then we have two *different* things. So, to say that they are both kinds of logical consequence is to make logical consequence into a disjunctive relation— it holds *either* when set of conditions *A* holds, *or* when set of conditions *B* holds. This is undesirable because it makes it unclear how it is one unified concept at all, and not just two different relations entirely. In the original examples, arguments A and B, the conclusions seem to *logically follow* from the premises in the *same way* as conclusions in valid declarative arguments follow from their premises. So, a successful account of logical consequence that includes imperative sentences must be general: it must apply the same criterion of logical consequence to pure imperative arguments, pure declarative arguments, and arguments that contain both imperatives and declaratives. At the very least, having a unified consequence relation is *preferable* to having a disjunctive consequence relation. However, it should be noted that some theorists do not accept this constraint. Rather than talking about *a* consequence relation, they are happy to talk about consequence *relations* linked by a common relation of family resemblance. After all, if logical pluralism looks like a plausible view, why not add another dimension of plurality by considering a concept or concepts of imperative consequence, like (but not identical to) declarative consequence? Perhaps a unified conception of consequence is not essential. However, the fact remains that a unified consequence relation would be *preferable*, if possible. A solution that meets this constraint would be better than one that does not, even if a pluralistic account would be

better than none at all. Since I think it *is* possible to develop such a unified account, I include it as a criterion of success.

The logic must also be adequate: it must not say there is consequence when there is none, or that there is no consequence when there is. This seems relatively straightforward as a criterion for a successful theory of logical consequence: it must be correct. However, the problem with this lies in determining whether or not a theory meets the criterion. It seems that we need a criterion of logical consequence in order to assess a proposed criterion of logical consequence, but if we had one already, the whole project would be unnecessary. So instead, we need an independent (and informal) test for logical consequence that we can use to guide us in the quest for the formal account. I propose two such informal tests:

**Test 1)** If someone were to endorse the premises (perhaps by asserting or commanding them, or by nodding assent to them), we would be licensed to take it that they had already effectively endorsed the conclusion.

**Test 2)** If someone were to accept the premises (perhaps by believing them, assenting to them, or agreeing to obey them), yet refused to accept the conclusion, and they understood the meanings of all the premises and the conclusion, they would be making a logical mistake.

These are intended to coincide, to be two slightly different ways to think about what we expect to happen when the relation of logical consequence holds.

Finally, the logic must be non-cognitivist about imperatives. If it fails this criterion, and makes imperatives truth-apt, then it would collapse into a type 2 solution. This means that it will, firstly, potentially fall foul of the objections I raise in chapter 4 on imperative cognitivisms. It also means that the theory is either not a type 3 solution at all (if it still claims that the relata of the consequence relation must be truth-apt), or else it is more revisionary than required. We do not want to reject *both* T2 *and* T3, as this is unnecessary. So, if a theory makes imperatives truth-apt, we would be better to keep our existing definition of logical consequence and apply it to imperatives than to *also* substantially revise this definition. The motivation for revising the definition of logical consequence is that we want to include imperatives but they are not truth-apt, so we can't (as the definition stands). If our theory makes them truth-apt, that motivation disappears. So, in order to count as a successful type 3 solution, it must make imperatives non-truth-apt, while still making declarative sentences truth-apt.

## 2.4 The Frege-Geach Problem

The problem of imperative consequence is closely related to the Frege-Geach problem in metaethics. The Frege-Geach problem, or, the problem of unasserted contexts, arises for non-cognitivists about morality. Non-cognitivists about morality hold that moral statements are not truth-apt; that they do not express propositions. Different forms of non-cognitivism differ on what moral statements do express. Emotivists hold that they express approval and disapproval. That is, for example, “murder is wrong” expresses something like “boo to murder!” Prescriptivists hold, instead, that moral statements express prescriptions, which is also what imperative sentences express. Thus, moral statements like “murder is wrong” are, for the prescriptivist, disguised imperatives. The Frege-Geach problem arises for each kind of non-cognitivist, when we consider patterns of reasoning, such as:

**E1** Stealing is wrong.

**E2** If stealing is wrong, then so is committing copyright infringement.

∴ **E3** Committing copyright infringement is wrong.

The problem for the non-cognitivist is that it is unclear what role “stealing is wrong” is playing in E2, that is, what role a moral statement can play as the antecedent of a conditional. In general, antecedents of conditionals are not being *asserted*, they are being *entertained*. So, it cannot be correct to think of the antecedent of E2 as “boo to stealing!” or as “do not steal!” Stealing is not being disapproved of, and refraining from stealing is not being commanded. However, the occurrence of “stealing is wrong” in E1 is not merely being entertained. It *is* being used. Stealing *is* being disapproved of, or refraining from stealing *is* being commanded. So, the two occurrences of “stealing is wrong” have different semantic functions. They are not playing the same semantic role in each instance. Argument E, then, should be invalid, as it contains an *equivocation* (see Geach (1960) and Geach (1965) for more on the Frege-Geach problem).

My solution to the problem of imperative consequence, it will turn out, also provides a solution to the Frege-Geach problem for one form of prescriptivism. The Frege-Geach problem is a two part problem: first, how can the non-cognitivist plausibly translate moral conditionals? Second, how can these translations exhibit the expected logical behaviour? My solution, then provides a solution to the second problem, but only for one answer to the first problem (and because it is not a theory of non-cognitivism about morality, it has nothing to say about the plausibility of this answer). See chapter 12 for the details of my solution, and in particular section 12.8 for how it solves the Frege-Geach problem.

Now that we have a clear idea of what each type of solution needs to do in order to be considered successful, we can begin to look at the various solutions and see how they fare. In Part II, I will consider type-1 solutions, arguments for Eliminativism about Imperative Consequence. Later, in Part III, I will turn to type-2 solutions, Imperative Cognitivism. Then, in Parts IV and V, I will discuss type-3 solutions, Formal Imperative Logics. Part IV will cover First-Order Imperative Logics, and Part V will cover Modal Imperative Logics. Finally, in chapter 12, I present my own solution, and show how it solves the problems of the other theories. For now, let us turn to the first type of solution: Eliminativism about Imperative Consequence.

## Part II

# Eliminativism about Imperative Consequence

## Chapter 3

# The Argument from Permissive Presuppositions

In this chapter, I will discuss an argument, offered by Bernard Williams (1963), against the possibility of imperative arguments. If successful, it would lead us to favour the rejection of T1. I claim that not only does Williams' argument fail, but also that imperative consequence is possible by Williams' own lights. Williams claims that there is only one logical relation between imperatives: that of contradiction. Two commands contradict when it is logically impossible for both commands to be obeyed. However, he claims, this is the only relation that can hold between commands, and there is nothing that can be generally thought of as imperative consequence.

Williams claims that strings of imperatives cannot form anything analogously like an argument because they are essentially "discontinuous". That is, a second imperative uttered after a first imperative represents a change of mind on the part of the commander, because they necessarily have conflicting permissive presuppositions. As arguments require all the premises to be accumulated and accepted at once (Williams claims), strings of imperatives cannot form the premises of an argument.

To illustrate this accumulation point, suppose a person, Alice, said "I want an icecream" then changed her mind and said "actually, I want some orange juice." The two statements "Alice wants icecream" and "Alice wants orange juice" cannot be used together as premises of an argument, for example to the conclusion "Alice wants icecream and orange juice" because she changed her mind between uttering them, and thus they are not meant to both be considered true at the same time. So, the conjunction "Alice wants icecream and orange juice" is not true at any time in this scenario. Williams claims that imperatives uttered in

succession always exhibit the same kind of discontinuity.

He asks us to consider the following example:

**A1** Do x or do y

**A2** Do not do x

∴ **A3** Do y

Compare (where p and q are declarative sentences):

**B1** p or q

**B2** not-p

∴ **B3** q

Argument A seems to be a parallel of Argument B, which is an instance of disjunctive syllogism. It is intuitively valid; that is, intuitively, A3 genuinely follows from A1 and A2 in much the same way that B3 follows from B1 and B2. But, Williams argues, we have to imagine it in a context – as commands are not commands unless they are commanded by somebody to somebody. He claims that neither the commander, nor the agent, makes a genuine imperative inference. In this chapter I will not discuss his claim that commands cannot be abstracted from their context. This is because, *even if* we take it that commands cannot be abstracted from their contexts, Williams is mistaken in his belief that commanders and agents can never make an imperative inference.

Although Williams appears to be denying a different claim to the one I want to defend, namely that imperatives can entail one another, his argument is still relevant. If he were correct that imperative inferences can never happen, then imperative logic could still proceed but it would, perhaps, be somewhat less interesting as it would not apply to anything we actually do. However, because Williams' arguments fail, and in fact imperative logic is possible even by his own lights, he gives us no reason to think that we can't make imperative inferences. We are, of course, ultimately interested not in which imperative inferences are *actually* made by people, but which imperative inferences are *rationally licensed* to be made by people. That is, which imperative entailments hold.

So, what does Williams claim is going on in argument A? Suppose a commander utters A1, then some time later utters A2. This commander, according to Williams, must change his mind between the two utterances, as he gives conflicting permissions in commanding each; namely he gives permission to do x and then denies permission to do x.

It is reasonably straightforward that command A2 denies permission to do x. "Do not

do x” and “you are not permitted to do x” seem to at least confer identical information to the hearer, if they are not synonymous (whether a command can be synonymous with any declarative is discussed in chapters 5 and 6). The more interesting claim here is that the command “do x or do y” implicitly gives permission to do x, and permission to do y. Williams demonstrates this by insisting that when a commander commands “do x or do y,” he permits the agent not to do x, if he does y, and he permits him not to do y, if he does x. This, in turn, permits the agent to do x, and permits the agent to do y, because ““I permit you not to do x, if you do y, but I do not permit you to do x” does not confer any permission at all” (Williams (1963): 31). When the commander commands A1, what he effectively does is give the agent a choice. That is, he says “you can choose between x and y, but make sure you do one of them.” So, Williams argues, this shows that the permissive presuppositions of the disjunctive command “do x or do y” include permission to do x and y themselves.

However, to command A2, “do not do x” is to explicitly express that x is not permitted. Thus, the same commander utters A1, which gives permission to do x, then utters A2, which disallows x. So he in fact reneges on the permission given when he uttered A1.

In order to give permission to do x and then take it away, Williams claims that the commander must have changed his mind about what he wants the agent to do. So it follows that the speaker is not making anything that can be called an imperative inference, as he does not endorse the premises all together at any time.

To demonstrate this point, Williams asks us to suppose that the commander utters the second command a few moments too late – the agent has already trotted off to obey the first command, and fails to hear the second command. Then the agent comes back, having done x but not y. He has obeyed the command which he heard, but the commander would be disappointed (or even angry). This indicates a change of mind or desire.

To further illustrate Williams’ claim that moving from “do x or do y” to “do not do x” represents a change of mind, suppose that A1 and A2 were commanded in the reverse order; that is, the commander first says “do not do x” then after some time says “do x or do y.” For example, suppose the commander first says “don’t go into the lounge” then a few moments later says “go into the lounge or the kitchen.” This seems to more clearly demonstrate a change of mind. It would even seem rather deceitful of someone to do this, and to expect their first command to remain binding. This demonstrates Williams’ description of the permissive presuppositions of the respective commands. That is, if “do not do x” is commanded, and then “do x or do y” is subsequently commanded, and we take this to nullify the first command, then the permissive presuppositions Williams outlined must be correct;

namely that “do x or do y” gives permission to do x (and to do y), otherwise there would be no conflict.

I offer four objections to Williams, taken up in turn over the next four subsections:

- 1 Permissive presuppositions are analogous to conversational implicature, and so should be ignored when formulating imperative arguments.
- 2 There are other valid argument forms, which Williams does not consider, that do not suffer from different permissive presuppositions.
- 3 There are explanations for the change in permissive presuppositions in Williams’ example other than a change of mind, so the different permissive presuppositions do not have to prevent accumulation.
- 4 Williams accepts enough logical relations between imperatives (contradiction and a form of negation) for a definition of a valid imperative inference to follow naturally.

### 3.1 Permissive Presuppositions and Conversational Implicature

Consider again the parallel, declarative case (Argument B) of disjunctive syllogism:

- B1** p or q
- B2** not-p
- ∴ **B3** q

In this case, the move from B1 to B2 appears to have an analogous change of mental state. Ordinarily, for the utterer (or hearer) of these two premises to sincerely leave open “p or q,” he must not know that not-p. But the second utterance (of not-p) obviously demonstrates that he does know that not-p. So is this not relevantly like the imperative case?

Williams claims that it is not, because it represents a movement from ignorance to knowledge. That is, the first claim displays less definite knowledge than the second. The disjunction “p or q” can only sincerely be asserted if the speaker genuinely does not know which. Stating “p” some time later simply means that the speaker has gained more knowledge. This does not upset the consistency of a set of premises. This, he claims, is due to the fact that the truth of a statement of fact depends on whether it is in accordance with the way the

external world is, not on whether or not the utterer knows it. That is, statements of fact (assertions) attain their fulfilment value when they correctly represent the world. In contrast, commands are obeyed, or complied with, when the world is in accordance with *them*. That is, commands attain their fulfilment value by the world being modelled on them. What is commanded depends entirely on the mental state of the commander. Which commands are appropriate depends on what the commander wills; it is up to him, at each stage in time, to decide what changes he wants initiated by means of his commands. If he is content with either x or y, he can command “do x or do y,” but if he wants “y and not x” then he is free to command this. They represent different desires or wills, and so if he moves from the first to the second, then he has to have changed his mind about what he wants. So, although the declarative case may represent an analogous change of mental state to the imperative case, it makes a difference only to the imperative case. More specific knowledge can be gained over time without the consistency of the premises being compromised, but a more specific command cannot be uttered without a change of mind and thus disjointed premises.

Consider the preparatory statement “I am going to command you to do x or to do y.” This prepares the agent for a later command, but is itself not a command. The declarative disjunction, “p or q,” is in one relevant way more like the preparatory statement than the disjunctive command: it allows the speaker to add “and I may be able to tell you later which of the two.” This is not true of the disjunctive command, “do x or do y,” as this in effect tells the agent to choose; to decide for himself whether to do x or to do y (I come back to a similar point in section 11.4.3).

Williams’ claim that in the declarative case the move from B1 to B2 represents a move from ignorance to knowledge relies on B1 coming first and B2 coming second. So we apply the same reversal technique: now we have someone uttering “not-p” and then “p or q.” This cannot be thought of as a move from ignorance to more specific knowledge, as by Williams’ own admission “p or q” cannot seriously be left open if it is known that not-p. Consider, as an illustration, someone uttering “Mark is not a doctor” and then a few moments later uttering “Mark is either a doctor or a nurse.” In this sort of case, it is clear (at least in some contexts) that we would take the second utterance to cancel the first utterance.

In the declarative case, at least, we surely agree that the order of utterance makes no difference to their being premises of an argument. My reversal example and the Alice example above, however, show that the order makes a difference to the overall intended (and understood) meaning. If logic does not care about the order of two or more utterances, then the order must be making a difference to the *conversationally implicated*, rather than the

*logically implied*, meaning.

The order seems to make a difference because of Gricean conversational implicature. Uttering “p or q” conversationally implicates that the speaker does not know which of p and q is true, because if they did know which one was true then they would be violating the maxim of Quantity by not being as specific as possible. It is because of this conversational implicature that reversing the order of “p or q” and “not-p” makes a difference to what we “conclude” from these two utterances. When “p or q” is uttered first, followed by “not-p,” it is plausible that the speaker could have gained some more specific knowledge. When “not-p” is uttered first, followed by “p or q,” it seems that the speaker must have made some sort of mistake.

We ignore this conversational implicature when we accumulate utterances into a set of premises. If we strip declarative statements of their conversational implicature in order to make them into premises of an argument, then ought we not do the same to imperative sentences?

Logic, in principle, deals with what can be concluded from (the contents of) sentences themselves, not with what can be concluded from the fact that a certain person uttered a sentence at a certain time. If imperative logic is to be a deviation from this practice, it must be shown that imperative sentences *cannot* be stripped of their conversational implicature in the way that declarative sentences can be.

## 3.2 Other Valid Argument Forms

There is another serious problem with Williams’ argument against the possibility of imperative inference. Even if Williams’ analysis of the apparent imperative disjunctive syllogism is correct, he does not consider other cases of apparently valid imperative arguments. His analysis of it as essentially discontinuous due to a change of mind on the part of the commander seems as though it could be specific to this form of argument, and will not hold in cases where there is no change of mind. The simplest example of this is the inference to a command from itself. For example:

**C1** Do x!  
∴ **C2** Do x!

In this case, the permissive presuppositions are the same in both C1 and C2, so there is no change of mind required to command C2 after commanding C1. Williams makes no

mention of this kind of case, so perhaps on his behalf we can point out that this is a very strange inference, even for its declarative counterpart. There are not really two commands here, but one, given twice. No inference needs to occur to get from C1 to C2, because they are one and the same command. Perhaps “inference” is a process that must occur between at least two different statements. That is, “X entails Y” requires that X is distinct from Y.

So, Williams could say that there is no inference going on in Argument C because there is no movement from one command to another. There is just the same command twice. He could say the same thing, presumably, about the declarative case; there is no inference from “p” to “p.” While this may be an unpopular claim, I intend only to reconstruct a consistent theory on behalf of Williams so that his argument can get off the ground.

There are other forms of inference, however, that have two or more different commands and are not considered by Williams. For example, consider the following cases:

**D1** Do x and do y!  
∴ **D2** Do x!

And, similarly:

**E1** Do x!  
**E2** Do y!  
∴ **E3** Do x and y!

These represent cases of conjunction elimination and conjunction introduction, respectively. There does not seem to be any way of construing these cases as changes of mind. In Argument D, furthermore, there is no accumulation of premises required as it only has one premise. In Argument E, there are two premises, but it does not seem impossible that both could be endorsed by the commander at the same time and thus form a set of two premises. Suppose that somebody said to you “open the window!” and then said “open the door!” It seems as though, in some cases, he may have changed his mind about what he intended for you to do, but not in every case. It seems perfectly reasonable that he simply intends for you to do both.

So, a commander who uttered E1 and then E2 some time later *may* have changed his mind about what he wanted the agent to do, but he *need* not. For Williams’ account to be correct it must *always* be the case that the commander changes his mind between commanding E1 and E2.

The only way I can see for Williams to plausibly deny that these are arguments is to

deny that the conjunction of two commands is itself a command. He could instead claim that a conjunction of commands is still two commands. So, for example, “do x and do y” is the same as “do x” and “do y.” If this is the case, then the inference from “do x and do y” to “do x” is really an inference from “do x” to “do x,” which is the same as Argument C and thus, according to my proposed analysis, not a real inference. Similarly with Argument E; the conclusion is just restating the premises: it moves from “do x” and “do y” to “do x” and “do y,” which is no movement at all. So again, there is no real inference.

Note that if this line was taken here, Williams would not be committed to denying conjunction elimination and conjunction introduction in ordinary declarative inferences. The claim here is that the conjunction of two commands is not itself a (different) command. This does not require that the conjunction of two declaratives is not a separate statement. Williams could still maintain standard declarative logical inferences and deny that conjunctive commands exist. As for mixed conjunctions, like “p and do x!,” Williams will have to group them in with conjunctions of two commands. That is, “p and do x!” is not a distinct sentence from “p” and “do x!” Otherwise, he would not be able to explain why it was not an inference to go from “p and do x!” to “do x!” Consequently, though, he will have to also accept that there is no inference from “p and do x!” to “p.”

What about equivalent commands, that nonetheless seem like they are different? For example, consider two De Morgan-like imperative equivalences:

**F1** Do neither x nor y!

∴ **F2** Don’t do x and don’t do y!

and:

**G1** Don’t do both x and y!

∴ **G2** Don’t do x, or don’t do y!

Williams will have to say that F1 and F2 are not really different commands, but the same command issued using different words. They have the same permissive presuppositions as each other, namely “you are not permitted to do x and you are not permitted to do y.” Similarly, he will have to say that G1 and G2 are the same command. Again, they have the same permissive presuppositions, namely “you are permitted to do x, as long as you don’t do y, and you are permitted to do y, as long as you don’t do x.”

Perhaps having the same permissive presuppositions is enough for F1 and F2 (and G1 and G2) to be one thought of as one and the same command. The plausibility of this

becomes somewhat strained when we modify F2 to only include one of the conjuncts (for example, F2\* “don’t do x!”). Intuitively, F1 and F2\* are different enough to be considered distinct commands. However, when combined with the claim that a conjunction is not a different command to its two conjuncts separately, this response does not seem implausible (if somewhat ad hoc).

Next, consider arguments with conditionals:

**H1** If do x then do y!

**H2** Do x!

∴ **H3** Do y!

H1 is not grammatically acceptable; examples simply do not occur in English. It would be something along the lines of “if I command you to go to bed, then I command you to brush your teeth first,” where “I command” must be taken as performative, not descriptive. However, because subordinating conjunctions (such as “if”) cannot precede imperative sentences (that is, imperatives cannot occur as subordinate clauses), this is impossible to express (at least in English) (for an argument from this fact to the impossibility of imperative inferences, see Wedeking (1970). I discuss this argument in chapter 4).

So we can turn our attention to conditional commands such as “if there is a stop sign, then stop!” or “if the weather is fine, then attack at dawn!” or even “attack at dawn only if the weather is fine!” (these examples are due originally to Charles Pigden, and first appear in print in Josh Parsons (2012)). There are two main forms of inferences involving conditional commands:

**I1** Attack at dawn only if the weather is fine!

**I2** Attack at dawn!

∴ **I3** The weather is fine.

This can be generalised to:

**J1** Do x only if p!

**J2** Do x!

∴ **J3** p

Secondly:

**K1** Attack at dawn if the weather is fine.

**K2** The weather is fine.

∴ **K3** Attack at dawn.

which can be generalised to:

**L1** Do x if p!

**L2** p

∴ **L3** Do x!

The first case, Argument I, can be analysed in Williams' terms. There does seem to be a change of mind – that is, after commanding “attack at dawn only if the weather is fine,” in commanding I2 the commander appears to be saying something like “scrap that, attack at dawn regardless of the weather.” This can be demonstrated by means of permissive presuppositions: in the case of I1, the commander is saying “you are not permitted to attack at dawn except if the weather is fine,” or “you are not permitted to attack at dawn if the weather is anything but fine.” On the other hand, in the case of I2, there is no condition placed upon the command. It simply says “attack at dawn (regardless of the weather – or any other considerations).” In permissive presupposition terms, “you are not permitted to not attack at dawn” – there is no escaping the duty.

The reversal technique also demonstrates that these permissive presuppositions are at work. Suppose we reverse the order of these premises in the same way as we reversed A1 and A2 previously. Suppose someone were to command I2 and then I1: that is, “attack at dawn!” and some time later, “attack at dawn only if the weather is fine!” It is clear, here, that we are not expected to infer “the weather is fine” from this (nor does it seem that the commander ought to, or could, do so). Rather, we (under normal circumstances) have been given a new, overriding command.

The second case, Argument K, is however not so straightforward for Williams' theory. There is certainly no change of mind required to explain the move from “attack at dawn if the weather is fine” and “the weather is fine.”

Actually, K1 “Attack at dawn if the weather is fine” is potentially ambiguous between:

- a) “At dawn: if the weather is fine, attack!” (or “if the weather is fine at dawn then attack!”) and
- b) “If the weather is fine (now) then attack at dawn.”

This is important because if it is the first (as it seems reasonable to suppose is intended in this case), then it is an instruction to consider a certain command; “attack at dawn” binding at a future time if the weather is fine at that time, and dismiss the command, that is, consider no command to be binding, if the weather is not fine.

There are other instances of the same form which seem to clearly be cases of the second. For example: “cover your eyes if you are afraid” or “don’t eat it if you don’t like it.” These seem to be of the second kind, that is, they are not instructions for a future time. The antecedent is about the mental state of the agent; “if you are afraid (now), then cover your eyes.”

In response to these examples, generally, perhaps Williams could respond that the commander never holds both L1 and L2 (to stand for any of the examples given) in his mind at the same time. If he knew  $p$  to be true, and he wanted the agent to do  $x$  if  $p$  were true, then he should just command “do  $x$ .” On the other hand, if he does not know (or even believe) that  $p$  is true, he would not assert it. In the examples above, the commander does not know that the antecedent is true either because it is a statement about some future time, or because it is about the mental state of the agent. In either of these scenarios, the commander is in no position to make the inference as he does not know  $p$ .

This is a plausible account of what might be going on, at least some of the time. Suppose, however, that the commander says “if you are afraid, come here” and the agent says “I am afraid.” Now, surely, the commander knows the antecedent to be true, so if he was to command “well then, come here!” then it seems that he has made a genuine inference.

### 3.3 Alternative Explanations for the Different Permissive Presuppositions

Another objection to Williams’ analysis of (any instance of) Argument A is that he only considers one explanation for the different permissive presuppositions of A1 and A2. He assumes that the commander must have to change his mind in order to command A1 and then A2. This is important because the change of mind is crucial to his position that the commander does not endorse the two at the same time, so they cannot be accumulated into a set of premises. If there is another adequate explanation for the move from A1 to A2, and if that explanation does not involve a change of desire or will on the part of the commander, then Williams’ argument that the commander never makes an imperative inference fails. So, do any such explanations exist? I shall offer and discuss four candidates.

(1) Perhaps the commander is logically incompetent, and does not realise that there is anything wrong with his commands because he cannot see the inconsistency. This would not require a change of mind. It also, however, is no help with attributing an imperative inference to him, because it is problematic to ascribe any kind of inference to a logically inept person.

(2) Perhaps the commander wants to do y (A3), and he uttered the premises A1 and A2 for rhetorical flourish, emphasis, or to otherwise avoid directly commanding the conclusion. For example, consider the following instance of Argument A:

**M1** Release the prisoners or execute them!

**M2** Do not release the prisoners!

∴ **M3** Execute the prisoners!

Suppose that a drug lord says to one of his inferiors, Mike, “head to the warehouse in Albuquerque tomorrow afternoon, and when you get there, either release the prisoners or execute them!” The next day, the drug lord learns that the prisoners are more dangerous than he originally thought, so he phones Mike on his cellphone. He is worried about being overheard, so he chooses his words carefully: “remember what I said yesterday? That you were to do one of two things?” “Yes, Boss” “well, don’t release them!” The drug lord has made it explicitly clear that the original disjunctive command is still in force, by referring back to it. He does *not* mean “scrap that, here’s a new command: whatever you do, don’t release the prisoners!” It is quite clear that he means for an inference to be drawn from the disjunctive command “release the prisoners or execute them!” in conjunction with the new command “do not release them!” Mike is expected to infer “execute the prisoners!”

The commander probably cannot be said to be making a genuine inference, but he is *exploiting a logical connection* in order for this unusual way of commanding to work. He is asking the agent to make an inference, but he has not arrived at the conclusion through any inference of his own. Rather, he desires the conclusion, and commands it in an unusual way, a way that requires that there is a logical connection between the imperatives in question.

(3) Perhaps, instead of being logically incompetent or not intending the disjunctive command to be taken seriously, the commander simply does not realise that his two commands carry the permissive presuppositions that it is claimed they carry. So the commander commands A1 and A2 of Argument A, but does not realise that “do x or do y” presupposes that x and y are both permitted by this command. In this case, he would not need to change his mind to subsequently command A2, as he would not realise that it would renege the

permission he had just given.

This is not particularly helpful, because if Williams is correct about the presuppositions inherent in a disjunctive command, then what this means is that a disjunctive command by definition includes its permissive presuppositions, and thus the commander does not realise he is commanding a disjunctive command. This means either that he is not really commanding a disjunctive command (if intention determines or is a necessary factor in determining what is uttered), or (if intention is not necessary in determining what is uttered) that he has such a lack of understanding of what he is commanding that he cannot be said to be making an inference.

(4) Perhaps the commander actually desires two conflicting things. That is, he fully understands that his commands carry the permissive presuppositions that they do, he intends them to do so, but he really does desire both things. He wants the agent to choose between  $x$  and  $y$ , but he also wants the agent not to do  $x$ . So he both gives permission and denies permission to do  $x$ . It is certainly possible that someone could desire conflicting things, even when he understood that they conflicted.

The problem with this scenario, however, is that the conclusion, “do  $y$ ” does not satisfactorily follow from his desires. If the commander really does want the agent to choose between  $x$  and  $y$ , then the commander does not want to end up effectively commanding that he must do  $y$ . Perhaps it could be said that the command “do  $y$ ” is an unpleasant consequence of his two somewhat conflicting desires, but a consequence nonetheless. For example, consider the following argument:

**N1** Bow to me, or leave my court!

**N2** Stay in my court!

$\therefore$  **N3** Bow to me!

Suppose that a king knows that it is a very strict rule that his subjects must bow to him or be cast out of his court, and furthermore he must command this, but he also might rather like this particular subject and wish him to remain. So he says “Bow to me or leave my court!” in his kingly, authoritative way, and then softly, perhaps pleadingly, “stay in my court, your grace!” This instance does not seem to demonstrate a change of mind, nor does it seem to be given for emphasis or flourish.

In fact, what appears to be going on here is that the second imperative, “stay in my court, your grace!”, is not a command, but a request or even a plea. However, as they are both in the imperative mood, this is still an instance of imperative inference. In particular,

it is an instance of Argument A in which the commander utters the premises (with their contradictory permissive presuppositions) but does not change his mind between uttering the first and uttering the second. Because there is no change of mind involved, Williams' argument that the premises cannot be accumulated into a set of premises fails, even for his own paradigm case.

A similar example of this is given by Geach (1963) in his reply to Williams: suppose you are seeking counsel from somebody, and they give you the advice "do x or do y." You then give him a list of reasons why you should not do x, and he replies "very well, do not do x, so do y." It is not required that the commander in this case change his mind. He may still endorse the disjunctive command; that is, his advice to choose between doing either x or y may still stand. However, in light of some new information, he also (perhaps for quite a different reason) endorses the advice not to do x.

### 3.4 Inference in Terms of Contradiction and Negation

A more fundamental objection to Williams' argument is that he contradicts himself: he claims that there is no such thing as imperative consequence, but he accepts enough about the logical relations between imperatives for an account of imperative consequence to follow naturally.

Williams admits that "there are certain logical relations between imperatives." They can contradict each other: "two imperatives may be said to be inconsistent, if and only if it is logically impossible that they should both be obeyed" (Williams, 1963: 30). He also accepts a negation-like operator. In his example (Argument A), the second premise was "do not do x," which he takes as denying permission to do x. This is a kind of negation of "do x," which he takes as denying permission not to do x. It may be objected that negation is defined by its truth-table, and so, by definition, is something that operates only on truth-apt sentences. In that case, think of this as a different operator, and call it "imperative-negation" or "i-negation" instead.

An account of imperative consequence can be formulated simply in terms of contradiction and i-negation: "an imperative conclusion follows from a set of imperative premises if and only if the set of sentences consisting of the premises and the i-negation of the conclusion contains a contradiction." In other words (using Williams' definition of imperative contradiction): "an imperative conclusion follows from a set of imperative premises if and only if it is logically impossible for the premises and the i-negation of the conclusion to all be obeyed."

This is a definition of imperative consequence, not imperative inference, but once we further specify that an imperative inference is licensed whenever there is imperative consequence, and not licensed if there is no imperative consequence, we have an account of permissible imperative inference.

I do not wish to claim that this is the best possible account of imperative consequence, nor even that it is adequate (for instance, it will not account for inferences with some imperative and some declarative sentences, like instances of Arguments J and L). I simply claim that it is an account of imperative inference that uses only the tools that Williams accepts. Without demonstrating that this proposed account is impossible, and because he accepts that imperatives can be contradictory and that a form of (imperative) negation is possible, Williams cannot consistently hold that there is no such thing as imperative inference.

The argument from permissive presuppositions fails for four reasons: because permissive presuppositions are analogous to conversational implicature, and so should be ignored when formulating imperative arguments; because there are other valid argument forms, which Williams does not consider, that do not suffer from different permissive presuppositions; because there are explanations for the change in permissive presuppositions in Williams' example other than a change of mind, so the different permissive presuppositions do not have to prevent accumulation; and because Williams accepts enough logical relations between imperatives (contradiction and a form of negation) for a definition of a valid imperative inference to follow naturally. So, we will now move on to another family of arguments for eliminativism about imperative consequence: arguments from the rules of grammar.

## Chapter 4

# Arguments from Rules of Grammar

In this chapter I focus on a group of arguments, offered by Gary Wedeking (1970), Jonathan Harrison (1991) and Jörg Hansen (2008), that go from premises about the rules of grammar to the conclusion that imperative arguments are impossible. They claim, broadly, that sentences expressing premises cannot be in the imperative mood, and so imperatives cannot be premises of arguments. Wedeking, Harrison, and Hansen, like Williams, want to deny that imperative inferences and imperative arguments are possible, and if they are successful, this will prevent the project of imperative logic from getting off the ground.

I will look at three instances of this type of argument. The first two appeal to the use of premise- and conclusion-indicating words such as “so,” “thus” and “because.” The first claims that they cannot (without violating rules of grammar) be used to indicate that an imperative is a premise, and together with Hare’s rule, this means that imperatives cannot be conclusions either. Hare’s rule states that “no imperative conclusion can be validly drawn from a set of premises which does not contain at least one imperative” (Hare (1952): 28). I will discuss it in more detail in chapter 7, but one of Wedeking’s arguments relies on it, so I will accept it for the purpose of this chapter.

The second argument claims that although conclusion-indicating words can precede imperative sentences, they are never used to indicate an inferential reason (that is, a premise), only as reasons *for carrying out* the order or request.

The third argument appeals to the deduction theorem. It claims that whenever we have a valid argument, we can make a conditional that is analytically true that has as its antecedent a conjunction of the premises of the argument, and as its consequent the conclusion of the argument. Imperatives cannot be antecedents of conditionals, so imperatives cannot be premises of arguments.

How is this relevant to the question of whether it is *possible* to formulate a logic of imperatives?

If Wedeking, Harrison and Hansen are right that we do not make imperative inferences in real-life discourse due to grammatical impossibilities, then at the least, it will show that we cannot base our logic of imperatives on ordinary language – that is, we cannot simply read off the logical relations from ordinary language discourse. They, notably Hansen, also claim that the lack of imperative inferences in ordinary language means that formulating a logic of imperatives is pointless. Be that as it may, the arguments cannot show that a logic of imperatives is impossible.

At best, this would be *prima facie* evidence that a logic of imperatives is impossible, if you think that ordinary language has the ability to express almost everything that it is possible to think. If this is compelling, and if imperative logic has a point and imperative inferences are possible, it would be strange if English had trouble expressing such inferences. At worst, it would be evidence that the logic of imperatives is uninteresting and useless, as it would not map onto anything we actually do (or even *could* do).

## 4.1 The Argument From Premise- and Conclusion-Indicating Words

*Premise-indicating words* (Wedeking (1970): 162) are those that indicate that what follows is a premise, such as “since,” “because” and “for.” For example:

- (a) Since it is a form of dishonesty, lying is wrong.
- (b) He must either live or work down my end of town, because I see him around there so often.

*Conclusion-indicating words* (Wedeking (1970):162) are those that indicate that what follows is a conclusion, such as “so,” “thus,” “therefore” and “hence.” For example:

- (c) I can feel a breeze, so there must be a window open somewhere.
- (d) She is a librarian, therefore she likes books.

So now let us consider some examples of these words in use in imperative sentences:

- (e) Since he is too busy right now, sit down and wait!
- (f) Open the window, for it is too drafty and the draft will worsen his condition!
- (g) That's the best thing for you to do, so do it, and don't hesitate!

Notice that in all of these cases, the imperatives are conclusions, never premises. Sentences where the premise-indicating word precedes an imperative, such as "since open the window..." or "...for sit down and wait" are ungrammatical.

Here are some more examples:

- (h) Since go outside and play, go outside!
- (i) The window will be open, for open the window!
- (j) Stop doing that and come here, so I'm angry!

These are all grammatically unacceptable. These words cannot be used to indicate that an imperative can be a premise. So, the argument goes, the use of premise- and conclusion-indicating words shows that imperatives can be conclusions, but not premises. However, add Hare's rule that "no imperative conclusion can be validly drawn from a set of premises which does not contain at least one imperative," (Hare (1952): 28) it follows that there are no premises from which imperatives can be validly derived.

Here is a reconstruction of this argument:

- 1.1 Premise-indicating words cannot precede imperative sentences.
- 1.2 If premise-indicating words cannot precede imperative sentences then imperative sentences cannot be premises of arguments.
- ∴ 1.3 Imperative sentences cannot be premises of arguments.
- 1.4 No imperative conclusion can be validly drawn from a set of premises that does not contain at least one imperative.
- ∴ 1.5 Imperative sentences cannot be conclusions of valid arguments.
- ∴ 1.6 Imperative sentences can be neither premises nor conclusions of valid arguments.

1.1 is the claim about the use of premise-indicating words, and 1.4 is Hare's rule. These are the substantive claims involved. 1.3 and 1.5 must be added as sub-conclusions. However, there is still a gap remaining: from 1.1 (premise-indicating words cannot precede imperative sentences) to 1.3 (imperative sentences cannot be premises of arguments). Something like

1.2 (if premise-indicating words cannot precede imperative sentences then imperative sentences cannot be premises of arguments) is required to bridge this gap. Perhaps it comes from a principle, such as “in order for a sentence to be a premise of an argument, it must be grammatical to precede that sentence with a premise-indicating word,” but this is more than is strictly required to get from 1.1 to 1.3.

The first possible response is to deny 1.1. This requires accepting that there exists a grammatically correct sentence with a premise-indicating word preceding an imperative.

If we are allowed to interpret this as “there exists *in a possible language* a grammatically correct sentence with a premise-indicating word preceding an imperative,” then this is obviously true. Let L be a language in which the word “because” functions as a premise-indicating word that can precede an imperative, and which is otherwise identical to English. Then, here is such a sentence:

**(k)** Because open the door and close the window, open the door!

The sentence (k) is not an English sentence, but a sentence in L. In this language, all the words in the sentence have the same meaning as their counterparts in English, except for “because,” which (as specified) functions as a premise-indicating word that can precede an imperative. So, we have our counterexample, and 1.1 is false.

This cannot be what is meant. 1.1 must mean “premise-indicating words cannot precede imperative sentences *in English*”. For now, just note that this step is required.

The second possible response is to deny 1.4. Then, we need an example of a valid argument that has an imperative conclusion, but no imperative premises. Here are some suggested counterexamples:

**A1** It is raining and it is not raining.

∴ **A2** Open the door!

Or perhaps:

∴ **B1** Close the window or don't close the window!

It could be objected that these are not legitimate counterexamples, because there is no connection between the premises and the conclusion (the single premise in argument A and the lack of premises in argument B). There is a lively debate still going on over these types of argument, even in the declarative case. I will not go into this debate, because (as I outlined

in section 2.3.3) I am assuming classical logic for the purposes of this thesis, and as such these types of arguments can be considered valid, at least in the declarative case.

However, even granting that these are legitimate counterexamples, it would not be particularly satisfactory if these were the *only* types of counterexamples, because it is desirable to say that our original examples (arguments A and B of chapter 1) are valid. In any case, 1.4 is really just a special case of the conservativeness principle in logic: something along the lines of “you can’t get out what you haven’t put in” (see Pigden (1989)). I’m happy to grant it for the purposes of this discussion.

Finally, we could deny 1.2. Recall that 1.2 was:

- 1.2** If premise-indicating words cannot precede imperative sentences then imperative sentences cannot be premises of arguments.

Recall, also, that we had to add “in English” to the end of 1.1. So, modify 1.2 accordingly;

- 1.2\*** If premise-indicating words cannot precede imperative sentences in English then imperative sentences cannot be premises of arguments.

The obvious response to this (modified) 1.2\* is that premises can be indicated in ways other than by the use of premise-indicating words. Premises can be indicated by appearing in a list that is known, by the context, to be an argument. For example:

“Here is an argument:

**C1** Open the door or close the window!

**C2** Don’t close the window!

∴ **C3** Open the door!”

Or, indeed, by preceding a conclusion-indicating word, for example:

- (1)** Open the door and close the window, so open the door!

Compare (1) with:

- (1<sup>1</sup>)** Since open the door and close the window, open the door!

As Castañeda points out, the fact that (1<sup>1</sup>) is ungrammatical does not show that “open the door and close the window” is not a premise in (1). It may be that it is not a premise in (1<sup>1</sup>), but (1<sup>1</sup>) and (1) are not equivalent because (1) is grammatically correct and meaningful

whereas (1<sup>1</sup>) is grammatically incorrect and meaningless. Because they are not equivalent, they do not express the same sentence, and so (1<sup>1</sup>) cannot show that (1) is not an argument (Castañeda (1971): 14).

I will come back to Argument 1, but first let us turn to the second argument.

## 4.2 The Argument From Motivating Reasons

Wedeking (1970), Harrison (1991), and Hansen (2008) offer versions of this argument, but Hansen's is the most detailed investigation of the use of conclusion-indicating words. Often we use these words before sentences, both imperative and declarative. Even in cases where they precede indicatives, it is sometimes difficult to tell whether they are indicating that what comes before is a logical (inferential) reason, that is, what comes after is a conclusion of a logical argument or inference, and when they have some other purpose such as to provide extra information or reasons why you should believe the claim that comes after the conclusion-indicating word. For example,

(m) I've just been to the gym, so I'm just going to have a shower.

Hansen proposes the following test: replace “so...” with “it follows logically from this that...” (Hansen (2008): 41). Now, we can more easily determine which case a given example falls into. For example,

(n) I have read all of Vladimir Nabokov's novels, so I have read Pnin.

can be changed to:

(n<sup>1</sup>) I have read all of Vladimir Nabokov's novels. It follows logically from this that yes, I have read Pnin.

without noticeably changing the sense of what the speaker is saying. On the other hand,

(o) There were holes in the roof, so birds had come in and were roosting in the rafters.

becomes:

(o<sup>1</sup>) There were holes in the roof. It follows logically from this that birds had come in and were roosting in the rafters.

This is obviously silly, and not what the utterer of (o) would most likely mean. They mean that the fact that there were holes in the roof *made it likely that* birds would come in and roost in the rafters. Next, we can apply this test to sentences where a conclusion-indicating word precedes an imperative, using the following example of an imperative sentence following “so”:

(p) Read all of Nabokov’s novels, so read Pnin!

We can try out various formulations:

(p<sup>1</sup>) Read all of Nabokov’s novels. It follows logically from this that read Pnin!

This is not grammatical, so instead replace the “that” with a colon, corresponding to a pause in spoken language:

(p<sup>2</sup>) Read all of Nabokov’s novels. It follows logically from this: read Pnin!

Hansen argues that this seems “strangely detached” (Hansen (2008): 43). It is unclear whether the “read Pnin!” is being commanded or just mentioned; for example, for the purpose of giving an example of imperative inference. It seems even stranger if the subject is added:

(p<sup>3</sup>) John, read all of Nabokov’s novels. It follows logically from this: John, read Pnin!

This seems even less like a real command and more like a logician giving an example of imperative inference. So Hansen tries out another phrase:

(p<sup>4</sup>) John, read all of Nabokov’s novels. We can conclude from this: John, read Pnin!

Now it is unclear who is giving the command “read Pnin!” – the speaker, the hearer/s, or the speaker *and* hearers. So perhaps change the pronoun:

(p<sup>5</sup>) John, read all of Nabokov’s novels. I conclude from this: John, read Pnin!

Hansen claims that this is the worst alternative so far. It is unclear whether the speaker is concluding or commanding “John, read Pnin!” He thinks these two things cannot both be happening at once: “the performative acts of concluding and commanding seem to collide, whereas the acts of stating and concluding seemed to go hand in hand” (Hansen (2008): 44).

Hansen thinks the next formulation is the best so far:

(p<sup>6</sup>) Read all of Nabokov’s novels. So you can conclude for yourself: read Pnin!

However, “it seems necessary that the “you” is the person to whom both commands are addressed” (Hansen (2008): 44). So we make the addressee explicit:

(p<sup>7</sup>) John, read all of Nabokov’s novels. So John, you can conclude for yourself: read Pnin!

(p<sup>8</sup>) John, read all of Nabokov’s novels. So Mary, you can conclude for yourself: read Pnin!

(p<sup>9</sup>) John, read all of Nabokov’s novels. So Mary, you can conclude for yourself: John, read Pnin!

Of these three formulations, only (p<sup>7</sup>) seems acceptable. In (p<sup>8</sup>) it seems that Mary is being commanded to read Pnin, which does not follow from the command issued to John. In (p<sup>9</sup>) it seems that Mary is being asked to give a command to John (and not just draw a conclusion). However, even (p<sup>7</sup>) becomes less acceptable when the addressee is explicit:

(p<sup>10</sup>) John, read all of Nabokov’s novels. So you can conclude for yourself: John, read Pnin!

It seems that in (p<sup>6</sup>) and (p<sup>7</sup>) the speaker has not only asked the addressee (John) to draw a conclusion but also to “give himself” the second command. This explains why making the addressee explicit makes it seem strange: John is being asked to use his own name when giving himself a command. If this interpretation of (p<sup>6</sup>) and (p<sup>7</sup>) is correct, then there is a further problem that if the commander of the first command “read all of Nabokov’s novels” and the commander of the second command “read Pnin,” are different people, then it is difficult to see who is doing the inferring and how this is possible.

By contrast, all of these examples can be converted into corresponding deontic sentences (statements about what ‘ought’ to be done):

(q) You ought to read all of Nabokov’s novels, therefore you ought to read Pnin.

(q<sup>1</sup>) John ought to read all of Nabokov’s novels, therefore John ought to read Pnin.

(q<sup>2</sup>) John ought to read all of Nabokov’s novels. It follows logically from this that John ought to read Pnin.

(q<sup>3</sup>) John ought to read all of Nabokov’s novels. We can conclude from this that John ought to read Pnin.

(q<sup>4</sup>) John ought to read all of Nabokov’s novels. I conclude from this that John ought to read Pnin.

(q<sup>5</sup>) John ought to read all of Nabokov’s novels. You can conclude for yourself that John ought to read Pnin.

(q<sup>6</sup>) John ought to read all of Nabokov’s novels. Mary, you can conclude for yourself that John ought to read Pnin.

All of these sentences are grammatical, meaningful and not ambiguous or confusing.

Given that none of the imperatives (p)-(p<sup>10</sup>) are straightforward and the deontic counterparts are all fine, Hansen proposes that it is because “these adverbs really are not used to indicate a claimed analyticity when linking imperatives” (Hansen (2008): 45).

This method of replacing “so” (or other conclusion-indicating words) with a clause making it explicit that the “so” is indicating a logical conclusion fails to produce meaningful, clear sentences in imperative cases, whereas it produces perfectly ordinary sounding sentences in declarative cases.

This is not surprising if we accept that conclusion-indicating words “provide only reasons, explanations or motives in the case of imperatives, and do not indicate claims of analyticity” (Hansen (2008): 46).

This also suffices to explain why (p) is meaningful: “the teacher, perhaps asked by John whether he also has to read Pnin, motivates the more specific imperative to read this book by prefixing to it the general requirement to read all of Nabokov’s novels, thus making it clear that Pnin is in fact one of the books that John has to read” (Hansen (2008): 46).

So, Hansen claims, there are no imperative inferences in ordinary language, and we cannot make imperative inferences because there are no such things as imperative arguments.

Here is a reconstruction of this argument:

- 2.1** In every instance where a conclusion-indicating word precedes an imperative sentence, it provides a motivating reason rather than a logical reason.
- 2.2** Motivating reasons are non-inferential.
- ∴ **2.3** Conclusion-indicating words never indicate an imperative conclusion of an argument.
- 2.4** If conclusion-indicating words never indicate an imperative conclusion of an argument then imperative sentences cannot be conclusions of arguments.
- ∴ **2.5** Imperative sentences cannot be conclusions of arguments.

Consider  $(p^2)$ - $(p^{10})$  again. Hansen argues that these are each unacceptable translations of  $(p)$ , for various reasons. However, most (all) of these reasons are to do with confusion over whether the conclusion (and perhaps the premise) is being commanded (that is, used) or merely mentioned. Hansen claims that this is evidence that the statements do not express arguments. This assumes, though, not only that arguments must be expressible in English, but also that they must be expressible by *using* the premises and conclusion (actually issuing commands with them) rather than mentioning them.

Does this not express an argument?: “if someone were to tell you “the door is open or the window is closed,” and you also knew that the door was not open, then it would be rational for you to conclude that “the window is closed” is true.” Here, for example, it appears that the conditional is true; it would be rational for someone to conclude that, and it would be rational because (perhaps among other reasons), the “conclusion” there follows validly from the “premises.” But there is a mixture of using and mentioning the premises in this case. Similarly,  $(p)$ ,  $(p^2) - (p^9)$  can be thought of as arguments with the premises and/or conclusion just being mentioned.

### 4.2.1 Vranas’s Response

Peter Vranas (2010) offers another response to Hansen’s argument. First, Vranas produces his own examples, which he claims bring out the problem and the solution more clearly:

(**r**) John, watch TV if and only if you finish your homework. So if you don’t finish your homework, don’t watch TV.

(**r**<sup>1</sup>) John, watch TV if and only if you finish your homework. It follows logically from this: John, if you don’t finish your homework, don’t watch TV.

(**r**<sup>2</sup>) John, watch TV if and only if you finish your homework. I conclude from this: John, if you don’t finish your homework, don’t watch TV.

Vranas says that he understands Hansen “as inferring that the speaker is not commanding (and thus the speaker is not both concluding and commanding) from the premise that, if the speaker were commanding, then she would not be saying that she is concluding” (Vranas (2010): 4).

**2'.1** If the speaker were commanding, then she would not be saying that she is concluding.

**2'.2** The speaker is saying she is concluding.

∴ **2'.3** The speaker is not commanding.

∴ **2'.4** The speaker is not both concluding and commanding.

Vranas thinks that 2'.1 is false – it is not the case that if the speaker were commanding, then she would not be saying that she is concluding. She may say she is concluding to make it clear that she is not *only* commanding, in (p<sup>4</sup>) and (r<sup>2</sup>). Vranas then anticipates a possible response from Hansen: that (Vranas's) (r<sup>2</sup>) is not a paraphrase of (r) because an utterance of (r<sup>2</sup>) could conclude without commanding, whereas an utterance of (r) must be commanding. He then replies to this anticipated response – asking us to consider:

(r<sup>3</sup>) John, watch TV if and only if you finish your homework. So, as a logical consequence:  
John, if you don't finish your homework, don't watch TV.

There is no difference between (r) and (r<sup>3</sup>) that is analogous to that between (r) and (r<sup>2</sup>), so (r<sup>3</sup>) passes (a modified version of) Hansen's test. Thus, the use of “so” before an imperative is evidence that imperative inferences are found in ordinary language.

### 4.3 The Argument From The Deduction Theorem

Wedeking's third argument against the existence of imperative arguments relies on the deduction theorem: the claim that whenever we have a valid argument, we can construct an analytically true conditional that has as its antecedent a conjunction of the premises of the argument, and as its consequent the conclusion of the argument. For example, take the following argument:

**D1** Ruby has a car and a hat.

∴ **D2** Ruby has a car.

We can construct an analytically true conditional as follows:

(s) If Ruby has a car and a hat, then Ruby has a car.

The antecedent is the premise of the argument, and the consequent is the conclusion of the argument. Another example:

- E1** Susie has a car or a hat.
- E2** Susie does not have a car.
- $\therefore$  **E3** Susie has a hat.

Again, we can construct a conditional:

- (t)** If Susie has a car or a hat and Susie does not have a car, then Susie has a hat.

Again, this conditional is analytically true, and has a conjunction of the premises as the antecedent and the conclusion as the consequent.

If the deduction theorem is a universal rule that works for every valid argument, and we assume we can have valid arguments containing imperatives, then it follows that it will work for an argument containing imperatives. So here is an example (from Hare (1952): 27):

- F1** Take all the boxes to the station!
- F2** This is one of the boxes.
- $\therefore$  **F3** Take this to the station!

Following the same method, we should construct the conditional as follows:

- (u)** If take all the boxes to the station and this is one of the boxes, then take this to the station!

However, not only is this not analytic, but it is a piece of nonsense. It makes no grammatical sense – it is not a well-formed sentence, let alone an analytically true one. It is never grammatical to place “if” before an imperative, so (the argument goes), it is grammatically impossible for imperatives to be antecedents of conditionals, and thus premises of arguments. Here is a reconstruction of this argument:

- 3.1** If an argument is valid, then an analytic conditional can be constructed with the conjunction of the premises as the antecedent and the conclusion as the consequent.
- 3.2** Imperatives cannot be antecedents of conditionals.
- ∴ **3.3** If the premises of an argument are imperatives, then an analytic conditional cannot be constructed with the conjunction of the premises as the antecedent and the conclusion as the consequent.
- ∴ **3.4** Imperatives cannot be premises of valid arguments.
- 3.5** No imperative conclusion can be validly drawn from a set of premises that does not contain at least one imperative.
- ∴ **3.6** Imperatives cannot be conclusions of valid arguments.
- ∴ **3.7** Imperatives can be neither premises nor conclusions of valid arguments.

Put 3.1 (the deduction theorem) aside for a moment. 3.2 is false. It is true that imperatives cannot be preceded by “if,” but that is not the only way to indicate an antecedent. Consider the following sentences:

(v) If open the door then it is cold outside!

(v<sup>1</sup>) Open the door only if it is cold outside!

While it is true that (v) is ungrammatical, (v<sup>1</sup>) is perfectly fine. Assuming that “if A then B” is, in general, equivalent to “A only if B,” which is contentious (although even if you deny this, it won’t ultimately save the argument), then (v<sup>1</sup>) is a counterexample to (3.2). This can be patched up, however:

**3.2\*** A conditional cannot have imperatives as both its antecedent and its consequent.

That is, imperatives can be antecedents of conditionals, and they can be consequents, but they cannot be both the antecedent and the consequent of the same conditional.

So we can have both of these:

(w) Open the door only if it is cold outside!

(w<sup>1</sup>) Open the door if it is cold outside!

But we cannot have either of these:

(**x**) Open the door if shut the window!

(**x**<sup>1</sup>) Open the door only if shut the window!

Then, we must change 3.3 to:

**3.3\*** If an argument has imperatives as both its conclusion and its premises, then an analytic conditional cannot be constructed with the conjunction of the premises as the antecedent and the conclusion as the consequent.

However, now we cannot conclude that it is impossible to have imperative premises. This new argument now only gets us to this modified conclusion:

∴ **3.4\*** There cannot be valid arguments with imperatives as both the conclusion and the premises.

So, we now have:

**3.1** If an argument is valid, then an analytic conditional can be constructed with the conjunction of the premises as the antecedent and the conclusion as the consequent.

**3.2\*** A conditional cannot have imperatives as both its antecedent and its consequent.

**3.3\*** If an argument has imperatives as both its conclusion and its premises, then an analytic conditional cannot be constructed with the conjunction of the premises as the antecedent and the conclusion as the consequent.

∴ **3.4\*** There cannot be valid arguments with imperatives as both the conclusion and the premises.

Hare's principle already says that it is impossible to have a valid argument with an imperative conclusion and (entirely) indicative premises. However, our modified argument does not rule out arguments with imperative premises but an indicative conclusion, such as:

**G1** Open the door only if it is cold outside!

**G2** Open the door!

∴ **G3** It is cold outside.

Likewise, it does not rule out the famous example from (Geach (1958): 52):

**H1** If you are a faithful subject, rise up, Sir George!

**H2** But do not rise!

∴ **H3** You are not a faithful subject.

So, this argument has not ruled out *all* kinds of imperative argument. However, it is unsatisfactory to accept that we can only have one sort of imperative argument. It still rules out our original examples (arguments A and B in chapter 1), which do seem to be imperative arguments.

So, now consider the so far unanalysed premise, the deduction theorem, 3.1 (if an argument is valid, then an analytic conditional can be constructed with the conjunction of the premises as the antecedent and the conclusion as the consequent.) The first thing to note is that the deduction theorem (as presented by Wedeking) fails to hold in some non-classical logics, such as relevant logic.

However, as I am assuming classical logic for the purpose of this thesis, we can assume the deduction theorem holds (for the declarative-only part of logic at least). We can, then, get to Wedeking's conclusion from this premise more easily together with the fact that no conditional with imperatives can be true, let alone analytically true:

**3.1** If an argument is valid, then an analytic conditional can be constructed with the conjunction of the premises as the antecedent and the conclusion as the consequent.

**3'.2** No conditional with imperatives as either the antecedent or the consequent can be analytically true.

∴ **3'.3** No argument with imperatives can be valid.

However, this begs the question against type two solutions (which hold that imperatives are truth-apt) and type three solutions (which hold that validity does not require the premises and conclusions of arguments to be truth-apt). 3'.2 just follows directly from the second claim in the inconsistent triad (that imperatives are not truth-apt, see chapter 1), and 3.1, tying validity to truth, is assuming T3 (that the relation of the consequence relation must be truth-apt).

If imperatives are not truth-apt, then by definition they cannot be true, and so by definition they cannot be analytically true. Assuming the truth-value of a conditional is determined by the truth-values of its antecedent and consequent, it cannot be true (or false) when either the antecedent or the consequent has no truth-value. So, no conditional with imperatives as either the antecedent or the consequent can be analytically true.

Someone would think it true that all valid arguments are capable of forming an analytically true conditional from their premises and conclusion *only if* they accepted from the outset that all of the component parts (the premises and the conclusion) of every valid argument are truth-apt. That is, it only makes sense to think that valid arguments can always be made into analytically true conditionals if it is already accepted that it is a necessary condition of validity that the premises and conclusions of valid arguments are truth-apt. This is because a true conditional is only true in virtue of the truth-values of its antecedent and consequent, and so these component parts by definition must be truth-apt.

Effectively, then, the argument goes: T3, T2, therefore not-T1. This is a cheap argument, and any of the three solutions can present their own version of this. Argument 3' gives us no reason to prefer solution type 1 over solution type 2 or solution type 3 that is independent of the claims in the trilemma itself, and so it can be of no use in persuading us to choose solution type 1.

Claim 3.1 (that if an argument is valid, then an analytic conditional can be constructed with the conjunction of the premises as the antecedent and the conclusion as the consequent) appears to have no motivation other than the fact that it works when considering a restricted class of arguments (those made up of purely declarative sentences). Tying valid arguments to analytically true conditionals in this way just begs the question against the existence of valid imperative arguments.

## 4.4 The Argument From The Grammar of Subordinate Clauses

There is a sense in which the argument from premise and conclusion indicating words and the argument from the deduction theorem are instances of the same, more general, argument. This is because premise-indicating words cannot precede imperatives for the same reason as imperatives cannot be antecedents of conditionals. Premise-indicating words (such as “for,” “because” and “since”) are subordinating conjunctions: they introduce subordinate clauses. So is “if.” So, really, they are instances of the same phenomenon: imperative sentences cannot appear as subordinate clauses. It is really this fact that is behind both arguments.

We can, then, reduce these two arguments to the following argument:

- 4.1 Imperatives cannot appear as subordinate clauses.
- 4.2 If imperatives cannot appear as subordinate clauses then they cannot be premises of arguments.
- ∴ 4.3 Imperatives cannot be premises of arguments.
- 4.4 No imperative conclusion can be validly drawn from a set of premises that does not contain at least one imperative.
- ∴ 4.5 Imperatives can be neither premises nor conclusions of valid arguments.

As I mentioned above, 4.1 is relatively uncontroversial (at least in English). However, the significance of this fact for the existence of imperative arguments is questionable. 4.2 (if imperatives cannot appear as subordinate clauses then they cannot be premises of arguments) appears to have no motivation whatsoever.

The connection between subordinate clauses of sentences and premises of arguments is that a common way (perhaps the most common way) for arguments to be expressed in English is with the use of a subordinating conjunction; the premise appears as a subordinate clause. However, this connection is contingent, even in English. We can express arguments in English without the use of subordinating conjunctions, and even if this were not possible, arguments are not necessarily English entities, but can be expressed in other languages: both natural and constructed. So, these arguments rest on the mistaken claim that arguments must be expressible a) in English, and b) with the use of a subordinating conjunction.

In this chapter, I have presented a series of arguments from rules of grammar to the conclusion that imperative arguments are impossible (that imperatives cannot be the relata of the consequence relation). I discussed the argument from premise- and conclusion-indicating words, the argument from motivating reasons, and the argument from analytic conditionals. I then showed that the first and third of these arguments both appeal to instances of the same phenomenon: that imperatives cannot appear in a subordinate clause.

All the arguments for eliminativism about imperative consequence I have considered have failed. So, our original examples (in chapter 1) of imperative consequence are strong *prima facie* evidence of imperative consequence, and in the absence of any better arguments to the contrary, we can cautiously move on to considering the next type of solution: Imperative Cognitivism.

## Part III

# Imperative Cognitivism

# Chapter 5

## Translation Schemas

In this chapter, I will outline several versions of Imperative Cognitivism. Recall that T2 of the trilemma was that imperatives are not truth-apt. Type two solutions (Imperative Cognitivism) reject T2 of the trilemma, and instead claim that imperatives *are* truth-apt. Several attempts to offer an account of how this is possible have been made, which I will outline in this chapter.

So, what is the problem here? Intuitively, imperatives are not in the business of describing how the world is. Imperatives are in the business of getting people to do things. It makes no sense to ask whether a command is true. If someone were to say, “Shut the door!” it would make no sense for you to reply, “yes, that is true,” or “no, I disagree.” It appears to be a category error; nonsensical in a similar way as replying to someone who had said “I rather like it when it rains” with “I refuse!”

So, what would it mean to deny T2? To deny T2 is to assert that imperatives *do* have truth-values. The most common form that this solution has taken is to claim that any imperative can be translated into a declarative without a change in meaning. So, the T2-denier must propose a plausible translation, and argue that the imperative and its translation really do have the same meaning. If this is done convincingly, then arguments containing imperatives can be treated the same as declarative-only arguments.

Before I discuss the candidates for this translation, though, it is important to distinguish two separate questions: 1) are these translations identical in meaning? 2) do they get the logic right? The first is stronger than the second; that is, if the translations do not get the logic right, then they must not be identical in meaning. It may be possible for the translations to not be identical in meaning and yet get the logic right. However, this would not amount to a rejection of T2. Rather it would be an admission that imperatives, as

they stand, are not truth-apt and cannot themselves stand in logical relations, but they can perhaps approximate logical relations insofar as they can be translated into declarative sentences that do stand in such relations. As such, it would (depending on how exactly it was fleshed out), be a version of either solution 1 or 3. So, for these solutions to count as type 2 solutions, they must hold that imperatives and their favoured translation are identical in meaning.

The five translation schemas I discuss translate imperatives into declaratives in various ways, thus making them truth-apt. They were proposed as translations that preserve meaning – that is, the translations are meant to mean the very same thing as their corresponding imperative. However, that is possibly a bigger ask than strictly necessary for our purposes. At the very least, a successful translation schema in this context must get the logic right. And it is at this level that I make most of my points.

## 5.1 Schema 1: Reports of Commands Theory

The first theory holds that commands are identical to reports of commands. The theory holds that any imperative “do x!” can be expressed with the corresponding performative utterance “I command that you do x!” That is, anything you say using the imperative mood can be said using the corresponding performative, and the performatives have truth-values. David Lewis, for example, proposes that imperatives

ought to be treated as paraphrases of the corresponding performatives, having the same base structure, meaning, intension, and truth-value at an index or on an occasion (Lewis (1983): 222).

Lewis maintains that

there is no difference in kind between the meanings of these performatives and non-declaratives and the meanings of the ordinary declarative sentences considered previously (Lewis (1983): 222).

Similarly, Frank Jackson and Philip Pettit hold this view:

Every competent English speaker knows that producing the words “Shut the door” in the right circumstances is *ipso facto* to command that the door be shut; that’s what is being done. And that is to say that it makes no difference whether

I say “Shut the door” or I say “I command that the door be shut” – a point which is independently plausible. But “I command that the door be shut” obviously has truth conditions. And so it follows that “Shut the door” has truth conditions too: it is true in *S*’s mouth at *t* just if *S* did indeed command at *t* that the door be shut (Jackson and Pettit (1998): 248).

So, this translation schema takes, for example, “get off my lawn!” and gives us “I command that you get off my lawn.” This is now a declarative sentence and as such can be true or false. It is true if I (the speaker) command that you get off my lawn, and false if I do not so command it. The claim here is that this is both a report of a command, and a performative – a command in itself.

## 5.2 Schema 2: Desires Theory

R. M. Hare put forward the Desires theory (and promptly rejected it): commands express “statements about the mind of the speaker” (Hare (1952): 5). In particular, “it might be held that “shut the door” means “I want you to shut the door” (Hare (1952): 5). So the translation would be from, for example, “get off my lawn!” to “I want you to get off my lawn.” This second statement is a straightforwardly declarative statement, which is true if I want you to get off my lawn, and false if I do not so want it.

This theory may seem immediately implausible, as it is clear that we do not need to desire an action be carried out in order to command that it be so. Also, it has the consequence that if one person says “shut the door!” and another person says (to the same agent) “do not shut the door,” they do not contradict each other (I will come back to this objection in chapter 6). However, it does have some merit. In at least most real-life uses of imperatives, the person commanding does so because he or she, on some level or other, desires the addressee to carry out the action.

## 5.3 Schema 3: Deontic Theory

Under the Deontic theory, imperatives are equivalent to sentences beginning with “you should...” For example, “get off my lawn!” is translated to “you should get off my lawn.” This theory looks very much like the Prescriptivism held by Hare (1952), but it is the reverse. Hare held that all moral statements, like “you should do the dishes” or “you ought to mow the lawn” are really disguised imperatives. So when we say, for example, “you ought to keep your

promises,” what we are actually doing is issuing a command: “keep your promises!” This view, in Ethics, is called Prescriptivism. That is, moral statements do not express facts, but *prescriptions*.

The Deontic Theory (set out notably by C.L. Hamblin (1987): 113), however, proposes the reverse translation. This view can be traced back to Kant, who says:

All imperatives are expressed by the word *ought* [or *shall*] and thereby indicate the relation of an objective law of reason to a will,... They say that something would be good to do or to forbear,.. (Kant (1962): 35-36).

It is worth pointing out, here, that neither Prescriptivism nor this Deontic Theory of Imperatives entails the other. Under Hare’s view, moral statements are a proper subset of all imperatives. In a parallel way, this Deontic view can hold that imperatives are a proper subset of “you should...” statements. That is, it does not require that all “you should...” statements are equivalent to some imperative sentence. So, objections to one theory will not necessarily be objections to the other — they will need to be applied independently.

## 5.4 Schema 4: Predictions Theory

P. C. Gibbons (1960) argues that all uses of language depend fundamentally on the declarative mood. He says: it is not “just one among a number of co-ordinate and relatively autonomous ways of [using language], but is the basic, fundamental way of using it upon which all its other uses depend” (Gibbons (1960): 107). In particular, he claims that “there is no separate logic of imperatives, that from the point of view of logic imperatives are indicatives” (Gibbons (1960): 107).

Gibbons proposes that imperatives are essentially predictions. That is, they are identical in meaning with their corresponding second-person future-tense indicative sentence. The translation of “do the dishes!” would be “you will do the dishes.” This second sentence is, again, straightforwardly declarative – it is true if the second-person future-tense sentence is true, and false if it is not. A command is true if the prediction is correct, or the addressee obeys the command. A command is false if the prediction is incorrect, or the addressee does not obey the command.

Gibbons points out that there are many grammatical forms of commands. Sometimes they will be in the imperative mood, for example “lay the table!” However, sometimes indicatives can have the same function: “you will lay the table,” and even sometimes questions:

“will you lay the table?” So, he argues, one can be replaced with either of the other two without remainder. That is, nothing is lost in the translation. If this is true, then there would appear to be nothing to stop us from singling out the declarative form as the most convenient for certain purposes, particularly for using them in logic. So, we can replace every imperative sentence with its corresponding declarative one (a prediction) and we don’t need a whole new imperative logic.

What is required, here, is that they really do have the same meaning, and that they can have the same function. I think it is clear that they *can* have the same function. But this is no major feat. Given an appropriate context, the same function can come from many apparently different sentences. For example, “Emily, Sam laid the table yesterday” or “Harry, someone needs to lay the table.” Even, given the right context, a nod of the head, a glance or a glare can perform the same function. So, the important claim is that the prediction-sentences have the same *meaning* as the imperative-sentences.

So, the Predictions theory holds that every imperative is identical in meaning to its corresponding declarative sentence to the effect that the command will be complied with. The translation of “get off my lawn!” would be “you will get off my lawn” – a prediction. This second sentence is, again, straightforwardly declarative: it is true if you get off my lawn, and false if you do not.

This view is, at best, one-directional. Even if every command is equivalent to an indicative sentence (a prediction that the command will be obeyed), it is not the case that every prediction is equivalent to a command. Of course (under normal circumstances) “it will rain” is not equivalent to any command, and (under normal circumstances) neither is “he will win.” Gibbons specifies that imperatives are second person future tense declaratives, but second person future tense declaratives cannot *all* be imperatives. What about “you’ll definitely get an A” or “you will win, I’m sure of it.” Of course, in the right context, these could be uttered as commands, but it is certainly not the case that they will always be commands. Because not every instance of a prediction is a command, there must be something different going on in those cases where they are commands. There appears to be, in fact, an ambiguity in, for example, “you will eat a banana for breakfast” between a command and a plain prediction: an ambiguity that is (normally) resolved by context. There can only be an ambiguity of meaning if there is more than one meaning to be had.

Gibbons, in his paper, points out that commands are a strange breed of predictions, in which the giving of the prediction is a necessary condition for the prediction to be fulfilled. Perhaps he can use this as a dividing line between plain predictions and commands. The

predictions that are also commands are those that are required to be made in order for the prediction to come true.

There still appears to be some further meaning tied up in the predictions, namely a belief (real or just intended to be communicated) that what is predicted will happen. This is what it means to predict. There is no requirement to believe your command will be obeyed in order to sincerely command. Perhaps you have disobedient soldiers or children. If it is not implicit in an imperative sentence that the command will be obeyed, then we have found a significant difference between the “meanings” of imperatives and predictions, and so imperatives cannot be equivalent to predictions (I will come back to this objection in chapter 6).

## 5.5 Schema 5: Elliptic Theory

This theory was proposed by H. G. Bohnert (1945), and also held by A. R. Anderson and Moore (1957). Bohnert thought that it is the best way to account for the motivating force behind commands. That is, when a command is issued to a person, it (under normal or ideal circumstances) holds some motivating force - the person will feel some motivating “pull” towards doing the specified action. Bohnert thinks that:

A situation  $M$  motivates a behaviour  $B$  of  $x$  only if  $x$  believes  $M$  and  $x$  can derive from  $M$  with the help of other personally believed sentences “ $P \vee B(x)$ ”, but not “ $P$ ” alone, where  $P$  is some future situation of directly unpleasant character (Bohnert (1945): 303).

Bohnert goes on to give an explanation of why the imperative mood has evolved in language (away from this motivating disjunction (or conditional) sentence):

Under conditions of urgency all sentences tend to become as short as possible. E.g. “The house is on fire!” becomes “Fire!”, and without losing its functional quality of being true or false... That is, it is an ellipsis. A command is here regarded as simply the more serious ellipsis of omitting the entire penalty clause of the described class of disjunctions, again without sacrificing the truth rules of a disjunction, which, however, remain clear, of course, only if the penalty *is* understood. “Run!”, then, is permitted to be considered true if the exhorted one stays and burns, where the indicative “you run” would be false (Bohnert (1945): 303).

That is, imperative sentences are ellipses of their corresponding disjunctive sentences, including the implicit threat or other unpleasant consequence, and they exist because they are commonly used in situations of urgency. So, the theory holds that commands are ellipses of disjunctive sentences of the form “you do x or else...” The translation of “get off my lawn!” would be, perhaps, “either you get off my lawn or I’ll call the police,” or more generally “either you get off my lawn or y” where y stands for some unspecified unpleasant consequence. This will perhaps be a threat, but it might not be. This means that commands are true if either you obey them, or if the bad thing that was implicit in the command happens, or, presumably, both.

## 5.6 Supposed Advantages of Schema 4

Gibbons claims that there are two advantages to his translation (Gibbons (1960): 114). First, that it is a form that we do actually use to explicitly give orders. In some circumstances, a person saying “you will do the dishes!” or “you will report to me in the morning!” can be taken straightforwardly as a command. It is not a particularly complicated or obscure type of command – it is usually quite obvious. It would also be a display of (probably intentional) misunderstanding if a person was to reply, “I don’t think you’re right about that” or, “yes, I think so too.”

Gibbons is correct that it is a good thing that his theory has this feature. However, it cannot be thought of as an advantage over the rival theories, because all the other theories I outlined also have this feature, so it is not an advantage over any of them. “I command you to do the dishes,” “I want you to do the dishes,” “you should do the dishes,” and “either you do the dishes or you’ll go to bed early,” can all legitimately be ways of commanding someone to do the dishes. Suppose we offer a translation that does not have this feature – that is, it is not a form that can be used to issue commands; for example, “Simon commands you to do the dishes!” Even in a case where this were true, and Simon did command you to do the dishes, I would be reporting that he so commanded it and not actually commanding it myself when uttering it. So, comparing this case with Gibbons’ one, the fact that his has the feature of being used to actually issue commands is an advantage over the Simon translation. In fact, it looks as though it might be a necessary condition for any translation. So, it is an advantage in the sense that it meets one minimum requirement, but all five of the theories meet this requirement.

The second advantage Gibbons suggests is that it can be rejoined to threats and promises

by means of “either-ors” and “if-thens.” It is easy to add a threat to “you will do the dishes.” It simply becomes “you will do the dishes or else I won’t give you any chocolate” or “you will do the dishes or else I won’t talk to you for a week.” Similarly, “you will report to me in the morning!” can become “either you will report to me in the morning, or there will be trouble!”

As for this advantage, if he just means that it is grammatically correct to add an “or else” clause to a command, then again it appears merely to be a necessary condition for a translation. However, there is a more interesting point lurking here. If we understand the disjunctive operator “or” as truth-functional, then it is difficult to make sense of it when one of the disjuncts is a command. Roughly the same problem is true of other operators too. So, for example, consider the perfectly meaningful sentence “either do the dishes or you’ll go to bed early!” If we understand the “or” in that sentence truth-functionally, then it is impossible to make sense of the sentence, because the first disjunct is “do the dishes,” which is not truth-apt. Note that this problem does not apply only to classical disjunction. It makes no difference how complicated your rules are for assigning a truth-value to the disjunction out of the truth-values of its disjuncts. You will still have a problem when one of the disjuncts has no truth-value.

Concerning this understanding of the problem, Gibbons’ theory does appear to have an advantage in that it can keep a truth-functional account of disjunction. If “do the dishes!” just means “you will do the dishes,” then both of the disjuncts are truth-apt. This is also an advantage over three of the four other theories I outlined, but I will discuss this point in section 6.4.

## 5.7 Cognitivist Pluralism

Henry S. Leonard (1959) holds a cognitivist *pluralism* about imperatives. That is, he believes that imperatives are truth-apt, but not always translated in the same way. In particular, he thinks that some sorts of imperatives should be translated as “you will...” statements, and others as “you should...” statements.

Leonard claims that “not only declarative sentences, but also interrogatives and imperatives, may be classified as true or false” (Leonard (1959): 172). His theory, as such, is another version of imperative cognitivism. He considers the following three sentences:

A Johnny will jump in the lake.

B Johnny, go jump in the lake!

C Will Johnny jump in the lake?

As I am not interested in interrogatives, I will only discuss declarative and imperative sentences. He explains that we can break sentences down into their *ultimate concern* and their *ultimate topic of concern*. In the case of the declarative, A, the ultimate topic of concern is that Johnny will jump in the lake, and the ultimate concern is that the addressee believe the ultimate topic of concern.

For the imperative, B, the ultimate topic of concern is the same: that Johnny will jump in the lake. However, the ultimate concern is different: it is that the addressee (Johnny) *makes it the case* that the ultimate topic of concern is true. In general, what is *expressed* by an imperative is the ultimate concern that the addressee bring about the ultimate topic of concern.

These three sentences all have the same ultimate topic of concern, namely *that Johnny will jump into the lake*. However, they have different ultimate concerns. Thus, Leonard explains,

We see the *same* ultimate topic of concern is *indicated* in all three speeches and we see *different* ultimate concerns *expressed* in all three speeches (Leonard (1959): 179).

The topic of concern is always a proposition, and is thus true or false. Leonard says:

Imperatives might be said to be true or false according as the propositions which are their ultimate topics of concern are true or false (Leonard (1959): 184).

In other words, an imperative sentence is true if and only if the proposition constituting the ultimate topic of concern describes the actual state of affairs. Although imperatives express the will of the speaker, their truth or falsity depends on the will of the addressee. Notice that this does not come apart in the case of declarative sentences. Or maybe this implies that declarative sentences express the ultimate concern that the addressee believe the proposition. Doesn't that mean that declarative sentences express the will of the speaker too?

Leonard explains that:

The addressee may conform his will to that of the speaker and act in the manner indicated, or he may disregard the speaker's will and act in a contrary fashion. According as he does the one or the other, he makes the imperative to be true or false (Leonard (1959): 184).

Incidentally, Leonard also thinks that imperatives can be lies, but this is not (necessarily) when they are false. It is when the speaker utters an imperative but *does not* desire that the ultimate topic of concern be brought about. His example is "Oh do come to spend Christmas with us! Bring all six of the children and the darling dog, and plan to stay until after New Year's!" (Leonard (1959): 184). What is going on in this example is that the speaker is uttering the imperative that expresses his desire for the addressee to come and spend Christmas with him, but he does not, in reality, have such a desire. It is, thus, a lie. However, whether it is true or false is still to be decided. If the addressee takes him up on his offer, and does bring the six children and the darling dog for Christmas, and does plan to stay until after New Year's, then he has made the imperative true, even though it was a lie.

Leonard also distinguishes between *personal* and *impersonal* imperatives. Personal imperatives express the will of the speaker, as we have already seen. Impersonal imperatives, on the other hand, express a *judgement* made by the speaker (Leonard (1959): 185). For example:

- (1) Thou shalt not steal.
- (2) Do unto others as you would that they do unto you.
- (3) Act only on that maxim that you can at the same time will it to become a universal law.
- (4) If you want to fry an egg, put it on the stove.

The judgement expressed in each case is:

- (1) You ought not to steal.
- (2) You ought to do unto others as you would that they do unto you.
- (3) You ought to act only on that maxim that you can at the same time will it to become a universal law.

(4) The stove is hot enough to fry an egg.

The truth or falsity of a moral imperative does not depend on human conformity with the injunction, so Leonard analyses them as *disguised declaratives*. As such, the ultimate topic of concern (of (1)) is not “you steal” or “you do not steal,” but “you ought not to steal.” So, impersonal imperatives are really disguised “you should” statements, and behave as such.

Leonard’s view, then, is a pluralist account of imperative cognitivism, where there are personal imperatives (directed at a particular person) and impersonal imperatives (directed at nobody in particular but just meant to be general rules). Personal imperatives are to be translated as “you will...” declarative sentences, or predictions, and impersonal imperatives are to be translated as “you should...” declarative sentences, or deontic statements.

So, Leonard’s theory is a combination of the Predictions Theory and the Deontic Theory. As such, it can be thought of as equivalent to the Predictions Theory where personal imperatives are concerned and equivalent to the Deontic Theory where impersonal imperatives are concerned. Thus, problems for those views will also be problems for Leonard’s view.

In summary, here are the five translation schemas I consider:

<b>Name of Theory</b>	<b>Translation of “Do x!”</b>
Reports Theory	“I command that you do x.”
Desires Theory	“I want you to do x.”
Deontic Theory	“You should do x.”
Predictions Theory	“You will do x.”
Elliptic Theory	“You will do x or else *bad thing*.”

In the next chapter, I will outline a series of problems for the schemas, and consider how each schema fares on each of them in turn.

# Chapter 6

## Problems for the Schemas

In chapter 5, I outlined the five different versions of imperative cognitivism. In this chapter, I point out four problems for these theories, and show how each of these problems affects each of the five translation schemas. I call them “problems” and not “objections” because some are worse than others, and I do not claim that any of them on their own is enough to knock down imperative cognitivism decisively. However, I claim that together they accumulate into a pretty good case against it.

The four problems are:

- 1** The Problem of Unwanted Consistencies – some pairs of sentences that ought to be inconsistent come out as consistent after they have been translated.
- 2** The Problem of Unwanted Validities – similarly, some arguments that should be invalid, come out as valid when translated.

These first two problems are due to Josh Parsons (2012). In his paper “Cognitivism about Imperatives,” he argued that imperative cognitivism suffers from unwanted consistencies and unwanted validities. However, he only considered the Reports theory. I will apply it to all five of the versions of cognitivism I have introduced. I also introduce two new problems:

- 3** The Problem of Soft Imperatives – these are imperatives with a softer force than outright commands, such as pieces of advice, requests, and instructions. They cannot always be accommodated by the translation schemas.
- 4** The Problem of Disjunctive Threats – adding an “...or else...” clause to a command causes problems for the translation schemas.

I will focus on one problem at a time, discussing it with respect to each of the translation schemas.

## 6.1 The Problem of Unwanted Consistencies

Unwanted consistencies occur when you translate seemingly contradictory commands. For example:

- (1) Attack!
- (2) Do not attack!

Let's see how each of the schemas translates 1) and 2).

### Unwanted Consistencies: Reports Theory

( $1_{rep}$ ) I command that you attack.

( $2_{rep}$ ) I command that you do not attack.

These two reports are not themselves inconsistent, even when uttered by the same person, because it is quite possible to issue inconsistent commands and to report that this has happened.

To be sure, a sincere speaker who utters ( $1_{rep}$ ) and ( $2_{rep}$ ) is making a mistake, and their mistake is a logical one. But it lies in their commanding contradictory things, NOT in the reports of their doing so. You might think that the translation of (2) should be:

( $2_{rep}'$ ) It's not the case that I command you to attack.

This results in the desired inconsistency. But, ( $2_{rep}'$ ) is not an accurate translation of (2). This is because (2) is a command, and thus should be translated (by the reports theory) as ( $2_{rep}$ ). Otherwise, (2) would have to be something like "It is not the case that attack!" which is ungrammatical, but more importantly not (2).

So, ( $1_{rep}$ ) and ( $2_{rep}'$ ) are inconsistent, but ( $2_{rep}'$ ) is not an accurate translation of (2) by the theory's own lights. So, ( $2_{rep}$ ) must be the translation, and this is not inconsistent with ( $1_{rep}$ ).

## Unwanted Consistencies: Desires Theory

( $1_{des}$ ) I want you to attack.

( $2_{des}$ ) I want you not to attack.

Just as it is not impossible to command inconsistent things, even though the commands themselves are inconsistent, it is not inconsistent to desire inconsistent things, even though the desires themselves are inconsistent (perhaps for different reasons, or without realising that they are inconsistent). So, the commands in (1) and (2) are inconsistent, but ( $1_{des}$ ) and ( $2_{des}$ ) are not inconsistent, because they can both be true.

## Unwanted Consistencies: Deontics Theory

( $1_{deo}$ ) You should attack.

( $2_{deo}$ ) You should not attack.

Are these inconsistent? Not if the “should” is interpreted as *pro-tanto*, rather than all-things-considered. It is entirely possible that there be a *pro-tanto* obligation that you attack, and also a *pro-tanto* obligation that you do not attack, so ( $1_{deo}$ ) and ( $2_{deo}$ ) would both be true. On the other hand, if the “should” is meant to be all-things-considered, then it is more plausible that they are inconsistent. Bearing in mind that you would be ruling out genuine tragic dilemmas, it could be defensible that it cannot be the case that, all things considered, you ought both to attack and to not attack.

But, (1) and (2) are simply inconsistent. It is impossible to obey them both. It is impossible to be logically competent and seriously issue them both.

Because (1) and (2) are inconsistent, and if ( $1_{deo}$ ) and ( $2_{deo}$ ) are to be plausible translations, then imperatives must be thought of as expressing all-things-considered obligations. Otherwise they would not preserve the inconsistency. But now our translation schema has been made independently less plausible, as we will see when we discuss the other problems. For now, just note that the only way to keep contradictory commands contradictory is to think of them as all-things-considered obligations.

## Unwanted Consistencies: Predictions Theory

( $1_{pre}$ ) You will attack.

( $2_{pre}$ ) You will not attack.

This is not a problem for this view. The translations are also inconsistent. Just as the imperatives “attack!” and “do not attack!” are inconsistent, so are the predictions “you will attack” and “you will not attack.”

However, a related problem occurs, which is actually the reverse – an unwanted *inconsistency*. It is entirely consistent both to sincerely command, for example, “Harry, get out of bed!” and to sincerely assert or believe the prediction “Harry will not get out of bed.” However, according to this theory, this amounts to predicting both that Harry will get out of bed and that he will not get out of bed. Although it is not impossible that someone may predict contradictory things (they may be insane, forgetful, or logically inept), it certainly amounts to endorsing a contradiction. However, it is entirely consistent, plausible, and in fact common, that someone could command someone to do something and yet also predict that he will disobey. So, two sentences that are consistent come out as inconsistent after being translated by this schema.

## Unwanted Consistencies: Elliptic Theory

( $1_{ell}$ ) You will attack or \*bad thing\*.

( $2_{ell}$ ) You will not attack or \*bad thing\*.

These translations are not inconsistent. If the \*bad thing\* is the same in both ( $1_{ell}$ ) and ( $2_{ell}$ ), then when it is true, both ( $1_{ell}$ ) and ( $2_{ell}$ ) will be true, regardless of whether or not you attack.

If the \*bad things\* are different, it just takes one of the \*bad things\* to be true, as long as it is the correct one. Because you can't help but either attack or not attack, whichever you do, just make the other \*bad thing\* true and both ( $1_{ell}$ ) and ( $2_{ell}$ ) will be true. In fact, the only way to make them inconsistent is to ensure the \*bad thing\* in ( $2_{ell}$ ) contradicts “you will attack” and vice versa. But then you just end up with “you will attack or you will attack” and “you will not attack or you will not attack,” neither of which fit the translation

schema. The idea behind the translation schema was to make the \*bad thing\* a negative consequence of not obeying the command.

So, the Reports, the Desires, and the Elliptic theories fail this test, the Predictions theory passes but fails on a related consideration, and the Deontics theory can perhaps get out of it by making imperatives express all-things-considered obligations, but they have to give up genuine tragic dilemmas, and anyway this will cause problems later.

## 6.2 The Problem of Unwanted Validities

When translating imperatives into these declarative transforms, we get arguments coming out as valid that should not be valid.

### Unwanted Validities: Reports Theory

- (3) Bow to me!  
∴ (4<sub>rep</sub>)Someone commands something.

This should come out as invalid, because the imperative sentence “bow to me!” does not, by itself, imply that anybody has actually used the sentence (or any other sentence) to issue a command. However, the imperatives as reports of commands view would translate the premise to “I command that you bow to me,” so this translation schema makes this argument valid.

Parsons argues that this argument is analogous to the following argument:

- (3\*) You are bowing.  
∴ (4\*) Someone asserts something.

This argument is invalid. You could very well be bowing but nobody asserts anything at all.

There appears to be an irrelevance problem here. That is, there is an irrelevance between the premise and the conclusion of just the sort that we normally consider prevents validity – they are “about” different things. The premise is about bowing, and the conclusion is about asserting. Similarly, there is an irrelevance to the imperative version of the argument, the first is about bowing, and the conclusion is about commanding.

### Unwanted Validites: Desires Theory

(3) Bow to me!

∴ ( $4_{des}$ )I want you to do something.

This argument is also invalid because it suffers from the relevant sort of irrelevance. (3) is about bowing, whereas ( $4_{des}$ ) is about the speaker's desires.

### Unwanted Validities: Deontic Theory

(3) Bow to me!

∴ ( $4_{deo}$ )You should do something.

First, note that ( $4_{deo}$ ) requires that there is some obligation requiring some action. This is not required by (3). It is not required by the apparent content of (3), it is not required to sincerely utter (3), and it is not required to obey, or to agree to obey, (3). Even the acceptance that (3) is in some sense "appropriate" or "correct" does not require any kind of objective normative fact or reason, particularly if we recall that the "should" must be an all-things-considered obligation. There seems to be more information contained in ( $4_{deo}$ ) than in (3), namely, that some obligation exists, so the argument is invalid.

This is decidedly less straightforward than some of the other theories. If I issue a command, it is likely that I think there is a reason for you to obey it, even if it is just that I want you to. Likewise, if I accept a command, whatever that means, it must be because I, in some manner, endorse or agree to the command. It is likely that doing this requires there being a reason for me to do so, even if it is simply that I like the idea.

But, firstly, remember that these reasons must be all-things-considered reasons. Secondly, the fact that it is likely there exists a reason when a command is issued or accepted sounds like a contingent fact about conversational implicature, rather than something contained in the sentence itself.

### Unwanted Validities: Predictions Theory

(3) Bow to me!

∴ ( $4_{pre}$ )You will do something.

It is pretty clear that it does not follow from a command that it will be obeyed, or in-

deed that anything will happen. But the predictions theory would translate (3) as “you will bow to me,” so under this view this argument comes out valid.

Consider the following sentences:

- (a) Do the dishes!
- (b) You will do the dishes.
- (c) Someone will do the dishes.
- (d) Harry will do the dishes.
- (e) You will do something.
- (f) Do something!

It is, I think, intuitively correct that from (b) (addressed to Harry), (c), (d), and (e) can be concluded. (d), because the subject is the same as in (b), but it has changed from second person to third person. In fact, “I will do the dishes,” uttered by Harry, would also follow from (b) or (d), because all three express the same proposition – that Harry will do the dishes.

It is perhaps, but contentiously, correct that (a) entails (f). But it is entirely counter-intuitive that (a) (addressed to Harry) would entail any of (c), (d) or (e). However, if (a) and (b) are equivalent, as Gibbons claims, then (a) would indeed entail (c), (d) and (e). If this is as absurd as it appears, then the Predictions Theory is incorrect: imperatives are not equivalent to predictions.

### Unwanted Validities: Elliptic Theory

- (3) Bow to me!
- ∴ (4<sub>ell</sub>) You will do something or something bad will happen.

This argument is valid under this view, because (3) is translated to “you will bow to me or \*bad thing\*.”

(3) and (4<sub>ell</sub>) seem to, again, have the same sort of irrelevance that we normally think would prevent validity. (4<sub>ell</sub>) is, at best, a somewhat reasonable thing to assume from the fact that (3) is issued. Of course, however, this is not enough for validity. And even so, it is also not infallible, because the issuer could be inclined to issue empty threats.

Perhaps we should see (1) as presupposing that there is a reason for the hearer to obey it or accept it. Aside from all the other problems we have seen, this particular view also assumes that the reason must be in the form of “...or else a bad thing will happen.” That is, the reason must be some sort of personal or impersonal threat. This does not seem to be the case. The reason might be just that a normative (perhaps moral, legal, or pragmatic) fact exists.

Perhaps, then, this could be stretched so that the “bad thing” is that you will violate a normative law. But then, it comes down to something like “you will attack at dawn or else you will violate your obligation to attack at dawn,” or worse: “you will obey your obligation or you will violate your obligation.”

So, all five of the theories fail this test.

### **6.3 The Problem of Soft Imperatives**

Soft imperatives are those that have less force than that of an outright command. I will focus on three main kinds of soft imperative.

#### **Pieces of advice**

For example,

- (5) Don't accept their offer just yet. Wait a couple of days. Make them think you've got other offers coming in.

#### **Requests or pleas**

For example,

- (6) Give me another chance! Please! I'll prove to you I can do it!

#### **Instructions**

For example,

- (7) Preheat the oven to 180 degrees Celsius and grease your cake tin.

These are soft imperatives, in the sense that they are less forceful than commands. All of the proposed translation-schemas assume imperatives express commands. And, of course, they often do. But it is important to remember that this is just one use of the imperative mood. Often, the imperative mood is used to express requests and other things. Let's see

what happens when we consider soft imperatives.

### **Soft Imperatives: Reports Theory**

It is clear that it makes no sense to translate non-command imperatives with “I command that...” as the operator. Not only would this make them always false, even when uttered sincerely and appropriately, it’s also just clearly not what is meant or understood. So, they will have to be translated differently depending on the type of speech-act that they are being used for.

(5<sub>rep</sub>)I advise that you not accept their offer just yet. I advise you to wait a couple of days.

(6<sub>rep</sub>)I request that you give me another chance. Please!

(7<sub>rep</sub>)I instruct you to preheat the oven to 180 degrees Celsius and to grease your cake tin.

These seem like somewhat plausible translations, but it does require that we know what sort of speech-act is being performed when we translate imperatives. We have to know a lot about the context of the utterance. At the very least, this wasn’t clear from the theory as initially outlined. Worse, though, this represents a marked asymmetry with declarative sentences. With declarative sentences, we do not need to know very much context in order to know what logical properties they have.

Also, this has the peculiar consequence that, in some cases, it would be invalid to conclude “do x” from “do x.” For example, in some cases, “open the window” would not imply “open the window,” because one would be one kind of speech act and the other another. I think it’s pretty clear that if an imperative can have any logical consequences, we want it to have itself as one. But, according to the Reports view, we can’t do the logic until we have translated the imperatives based on the context of their utterance. So, it might be that one is translated as “I request that you open the window” and the other as “I command that you open the window,” which would not be consequences of one another.

Of course, this objection seems a little unfair given that we usually keep the context of utterance fixed when we do logic. That is, this objection looks similar to the following analogous case:

Suppose Sam truthfully says “I am hungry” and Emily truthfully says “I am not hungry.” Now we have a case where “I am hungry and I am not hungry” is true - a true contradiction!

This is, quite clearly, either not a contradiction (if the “I” in each conjunct refers to a different person), or not true (if the “I” refers to the same person, whether Sam or Emily). This is an example of why we fix the context of indexicals – words like “I,” “here,” and “now.” The case of the Reports Theory, though, is a little different. It is not the indexical that is causing the problem. We can fix the speaker as the same person, and the problem remains. The problem is that there is a difference between requesting and commanding, and we can do one without the other but both with the use of the same imperative sentence. So, in order to get out of this problem for the Reports Theory we must also fix the *type of speech act* that is being made in any context.

### Soft Imperatives: Desires Theory

(5<sub>des</sub>) I want you not to accept their offer just yet. I want you to wait a couple of days.

(6<sub>des</sub>) I want you to give me another chance. Please!

(7<sub>des</sub>) I want you to preheat the oven to 180 degrees Celsius and I want you to grease your cake tin.

The translation of (6) is almost certainly fine - it closely approximates what is meant, except, perhaps, it is lacking the urgency of the imperative version.

The translation of (5) may not strike you immediately as wrong, but there is a worry that it is possible to give good, genuine, and appropriate advice without desiring that it be followed. The adviser may not have any desires either way, or they could desire that the advisee do just the opposite. I could conscientiously advise someone for their own good, even though for selfish reasons I would prefer them not to follow my advice. For example, I could advise a good friend to take up a wonderful career opportunity on the other side of the world even though I would prefer them to stay near me.

The translation of (7) is also wrong. Unless the author of the recipe book knows me personally, it is quite implausible that he or she really wants me to preheat the oven and grease my cake tin. If we interpret the “you” as plural, the author might have some general desire for people to follow his instructions and get to enjoy delicious cakes. But this is by no means required in order to write a recipe book or otherwise give instructions. It is not contained in the meaning of the instruction.

### Soft Imperatives: Deontic Theory

(5<sub>deo</sub>)You should not accept their offer yet. You should wait a couple of days.

(6<sub>deo</sub>)You should give me another chance. Please!

(7<sub>deo</sub>)You should preheat the oven to 180 degrees Celsius and you should grease your cake tin.

Keeping in mind that the “shoulds” must be thought of as all-things-considered shoulds, all three of these translations fail dismally. Advice can be given, it can be good and appropriate and genuine, without there being an all-things-considered obligation that you follow it. The same holds for instructions.

The translation of (6) is even less plausible. We can beg hard for another chance while privately believing that we should not be given one. It is almost *required* that there be no obligation, pro-tanto *or* all-things-considered, in order for (6) to make sense. The speaker realises that it is a plea to do something you *really don't*, in any way, have an obligation to do.

### Soft Imperatives: Predictions Theory

(5<sub>pre</sub>)You will not accept their offer just yet. You will wait a couple of days.

(6<sub>pre</sub>)You will give me another chance. Please!

(7<sub>pre</sub>)You will preheat the oven to 180 degrees Celsius and grease your cake tin.

These translations also fail dismally. Even if we grant that you can't genuinely command someone to do something if you don't believe they will obey you, as soon as we consider soft imperatives it becomes clear that this approach will not work. It is entirely possible to give advice, make requests, and issue instructions without holding any belief that the addressee will follow your advice or instructions or obey your request.

### Soft Imperatives: Elliptic Theory

Maybe the bad things will be:

(5<sub>ell</sub>) You will not accept their offer just yet or else they will think you are too eager and you will not have any bargaining power.

(6<sub>ell</sub>) You will give me another chance or else I'll be really sad. Please!

(7<sub>ell</sub>) You will preheat the oven to 180 degrees Celsius and grease your cake tin or else your cake will be suboptimal.

These translations seem, on initial consideration, to be rather attractive, particularly those of (5) and (7). They do seem to be plausible explications of what the speaker communicates to the hearer. But are they really *translations* of what is contained in the original (5)-(7)? The translations seem to have a lot more information than the originals; information that comes from the context, not contained in the sentence itself. This is, of course, a general problem with this theory. How can we ever be sure, when it is not explicated, what the correct \*bad thing\* is? It doesn't seem to be contained in any imperative that there is any negative consequence to not obeying it.

So, it looks like the first four theories fail this test spectacularly, and the Elliptic theory, at the least, looks even more implausible than it already did.

## 6.4 The Problem of Disjunctive Threats

The translation schemas run into difficulties when they try to make sense of disjunctive threats. Disjunctive threats are threats that are joined to an imperative with a disjunction.

Consider the following case:

(8) Finish your vegetables!

(9) Finish your vegetables or you'll have no dessert!

(9) has added a threat to (8), that is, a negative consequence of disobeying (8). (9) is, admittedly, somewhat unclear. Is it a real disjunction? Is it really a conditional – something like “if you don't finish your vegetables then you will have no dessert”? The weirdness of (9) may turn out to be responsible for the weirdness of the translations of (9), but I think they're worth looking at anyway, because they bring out some interesting features of each theory.

## Disjunctive Threats: Reports Theory

( $8_{rep}$ ) I command that you finish your vegetables.

First, there is a scope decision to be made, between:

( $9_{rep}^a$ ) Either I command that you finish your vegetables or you'll have no dessert.

and:

( $9_{rep}^b$ ) I command that either you finish your vegetables or you have no dessert.

( $9_{rep}^a$ ) doesn't seem right. The commander, when he adds a negative consequence of disobeying the command, doesn't want to suggest that he might not be commanding that you finish your vegetables, as long as you don't have any dessert.

( $9_{rep}^b$ ) initially looks a little better, but this does not capture what is meant either. It is not the case that the commander is giving you a genuine choice between finishing your vegetables and having no dessert. Rather, he is issuing an unconditional command to finish your vegetables, and then specifying the consequences of disobeying that unconditional command.

Perhaps we must translate (9) as:

( $9_{rep}^c$ ) I command that you finish your vegetables and if you do not finish your vegetables then you will have no dessert.

This is rather more plausible. However, surely it is a problem that the theory can't translate (9) directly – it has to break it into a conjunction of the command and the threat, and also to convert the threat into a conditional.

## Disjunctive Threats: Desires Theory

This theory faces much the same problems as the Reports theory.

( $8_{des}$ ) I want you to finish your vegetables.

( $9_{des}^a$ ) Either I want you to finish your vegetables or you'll have no dessert.

( $9_{des}^b$ ) I want you to either finish your vegetables or have no dessert.

I think this is even clearer than the Reports case, that is, neither  $(9_{des}^a)$  nor  $(9_{des}^b)$  are accurate translations of (9).

Similarly, it makes sense to do something similar to  $(9_{rep}^c)$ , that is:

$(9_{des}^c)$  I want you to finish your vegetables and if you do not finish your vegetables then you will have no dessert.

But again, we have to make the sentence rather more complex in order to understand it correctly.

### Disjunctive Threats: Deontic Theory

$(8_{deo})$  You should finish your vegetables.

$(9_{deo}^a)$  Either you should finish your vegetables or you'll have no dessert.

$(9_{deo}^b)$  You should either finish your vegetables or have no dessert.

The deontic theory cannot make sense of threats added onto imperatives with a disjunction, so it fails this test.  $(9_{deo}^a)$  is a disjunction of a normative fact and a prediction. So it entails that if you do have dessert, then a normative fact exists such that you should finish your vegetables. It also entails that if no such fact exists then you will not have dessert, which is a particularly strange consequence.

$(9_{deo}^b)$  is not much better, as it says that there is a normative reason to obey a disjunction – that is, one or other of the disjuncts. Presumably, then, if you do not have dessert, you will have complied with the normative fact. However, this is not what was originally meant.  $(8_{deo})$  gives an unconditional normative fact: you ought to finish your vegetables. You can't obey it by having no dessert; that is *what will happen if you disobey*.

### Disjunctive Threats: Predictions Theory

$(8_{deo})$  You will finish your vegetables.

$(9_{deo})$  Either you will finish your vegetables or you will have no dessert.

The predictions theory does not obviously fail the disjunctive threats test. However, it does seem to miss the point of (9), in that it somewhat loses the force of a threat. It looks more like a disjunctive prediction: a prediction that one of two things will happen.

It is built into the Predictions theory that the commander asserts that the command will be obeyed. That is, (8) requires that the commander asserts that the addressee will actually finish their vegetables. It doesn't seem that adding a threat to a command could change that. In fact, threats don't entirely make sense in this theory, given the commander can't genuinely issue a command without believing that command will be obeyed.

### Disjunctive Threats: Elliptic Theory

(8<sub>ell</sub>) You will finish your vegetables or else \*bad thing\*.

(9<sub>ell</sub><sup>a</sup>) Either you will finish your vegetables or you will have no dessert.

This translation schema, I think, does the best on this consideration, as (9<sub>ell</sub><sup>a</sup>) is simply (8<sub>ell</sub>) with the \*bad thing\* made explicit. The Elliptic Theory was actually designed, in part, around taking threats into account. It is designed to retain the motivating force that accompanies the imperative mood.

Now, it might be tempting to think we should translate (9) as:

(9<sub>ell</sub><sup>b</sup>) Either you will finish your vegetables or else \*bad thing\*, or you will have no dessert.

However, this is to misunderstand how the schema works. The \*bad thing\* just is the negative consequence of disobeying the command, whatever that is in the context. Sometimes it is an unspecified \*bad thing\*, but in cases like (9) the negative consequence is made clear, so it becomes the second disjunct.

So the correct translation, according to this view, is (9<sub>ell</sub><sup>a</sup>). It is, still, difficult to see how, under this view, (9) is distinct from the first disjunct of (9). Also, as in the Predictions case, it now looks like a disjunctive prediction rather than containing a threat.

All of the versions of imperative cognitivism, the Reports theory, the Desires theory, the Deontic theory, the Predictions theory, and the Elliptic theory, fail some combination of these considerations: the problem of unwanted consistencies, the problem of unwanted validities, the problem of soft imperatives, and the problem of disjunctive threats. They are, then, inadequate as translations of imperatives. So, all the proposed versions of imperative cognitivism are implausible.

We have now seen that the arguments for eliminativism about imperative consequence (solution type 1) are unconvincing, and imperative cognitivism (solution type 2) is implau-

sible. So, we now turn to solution type 3: formal systems of imperative logic. First, I will discuss several candidate formal logics of imperatives that are based on first-order logic, and demonstrate why none of these proposed approaches is successful. Then I will move on to two that are based on modal logic, and show that they are both unsuccessful as they stand. However, I claim that fixing each of them results in the same theory, which I outline and defend in chapter 12: Relational Preposcription Semantics and KD45.

## Part IV

# First-Order Imperative Logics

# Chapter 7

## Early Attempts

In this chapter, I consider two early attempts to formulate a logic of imperatives. First, I outline R. M. Hare’s Neustic-Phrastic analysis, and point out some problems with it. Then, I outline A. Hofstadter and J. C. C. McKinsey’s Operator analysis, again pointing out some serious problems with it.

### 7.1 The Neustic-Phrastic Analysis

R. M. Hare (1952) proposed one of the earliest accounts of imperative logic, based on a distinction he makes between what he calls the “neustic” and the “phrastic” elements of sentences. He then defines entailment between sentences, and proposes two rules for imperative validity.

#### 7.1.1 Indicative and Imperative Sentences

Hare uses the term “command” to refer to all the sorts of things that the imperative mood expresses, and the term “statement” to refer to what *typical* indicative sentences, if there are such, express. He points out what he takes to be the key difference between indicative and imperative sentences: “an indicative sentence is used for telling someone that something is the case; an imperative is not— it is used for telling someone to make something the case” (Hare (1952): 5). Telling someone that something is the case is answering the question “what are the facts?”, whereas telling someone to do something is answering the question “what shall I do?” (Hare (1952): 15). He goes further, asserting that commands (like statements) are *essentially intended for* answering questions asked by rational agents. It is this similarity

that makes it clear that commands (like statements) are governed by logical rules (Hare (1952): 15-16).

It is clear, though, that imperatives like “shut the door!” and indicatives like “you will shut the door” have something in common. They both seem to be about *your shutting the door*. Hare specifies that the common part of imperative and indicative sentences, for example “shut the door!” and “you will shut the door”, is called the *phrastic*, and is in this case: “your shutting the door in the immediate future.” The indicative and imperative sentences, although both about your shutting the door, are not identical in meaning, so they must have a part that is unique to each. Hare calls this part the *neustic*, and in this case, the neustic part of “shut the door!” is “please”, and the neustic part of “you are going to shut the door” is “yes.” These parts can be presented in a table for clarity:

	<b>English Sentence</b>	<b>Phrastic</b>	<b>Neustic</b>
<b>Indicative</b>	You are going to shut the door	Your shutting the door in the immediate future	yes
<b>Imperative</b>	Shut the door!	Your shutting the door in the immediate future	please

It is obvious, then, that when looking for the essential difference between the imperative mood and the indicative mood, we must look to the neustic, not to the phrastic. However, as Hare’s use of the single word “neustic” (from the Greek, meaning “to nod assent”) indicates, he thinks there is still something in common between the imperative neustic and the indicative neustic. This is “the common notion of, so to speak, “nodding” a sentence. It is something that is done by anyone who uses a sentence in earnest, and does not merely mention it or quote it in inverted commas; something essential to *saying* (and meaning) anything” (Hare (1952): 18). So, in the case of both imperatives and indicatives, to genuinely utter them, the speaker must “nod assent” to them. This is the neustic part of the sentence. The difference, then, resides in the different *ways* in which a speaker nods assent to the different types of sentence. In particular:

A hearer (sincerely) assents to a statement if and only if she believes that what the speaker said is true.

A hearer (sincerely) assents to a command if and only if she does, or resolves to do, what the speaker told her to do.

“Sincerely,” here, is merely meant to rule out sarcasm, lies, metaphor, jokes, mentioning rather than using sentences, and other such non-assertional uses of language.

It seems that defining what it takes for a *speaker* to assent to a sentence is important, but Hare does not mention this. Presumably, a speaker (sincerely) assents to a statement if and only if she believes that it is true. A speaker (sincerely) assents to a command if and only if she intends that the agent does, or resolves to do, what she says.

### 7.1.2 Logic

Hare thought that the logical connectives (conjunction, disjunction, conditional), as well as operators (quantifiers, modal operators), all belong in the *phrastic* part of the sentence. Negation, too, generally belongs in the phrastic part. However, he thought it was possible to *negate the neustic* of an imperative, and this in fact gives us a *permissive* sentence, such as “you may open the door,” or “I don’t tell you not to shut the door.” We can also negate the neustic of an indicative sentence, when, for example, we do not wish to assert a sentence. We thus get something like “it may (or may not) be that you will open the door.”

Hare thinks *entailment* can be defined “accurately enough” as follows (Hare (1952): 25):

A sentence  $P$  entails a sentence  $Q$  if and only if the fact that a person assents to  $P$  but dissents from  $Q$  is a sufficient criterion for saying that he has misunderstood one or other of the sentences.

It is not clear what he means by “dissents” here; he could mean merely “does not assent,” or something more active. For example, dissenting from a statement could mean “not believing that what the speaker said is true” (call this *weakly dissenting*), or it could mean “believing that what the speaker said is false” (call this *strongly dissenting*). Similarly, a hearer weakly dissents from a command if and only if she does not, and *does not resolve* to do, what the speaker told her to do. A hearer strongly dissents from a command if and only if she does not, and *resolves not* to do, what the speaker told her to do. However, we can charitably assume that Hare meant strong dissention, because if he meant weak dissention then very few, if any, entailments would hold. Someone could easily believe  $P$  but fail to believe  $Q$ , without misunderstanding either of them, simply because  $Q$  had never occurred to her.

He points out that this definition, along with the point that logical words (such as “all”) can occur in commands, implies that there must be entailment-relations between commands. His example is:

**A1** Take all the boxes to the station!

**A2** This is one of the boxes.

∴ **A3** Take this to the station!

Anyone who assents to A1 and A2, but (strongly) dissents from A3, clearly misunderstands (at least) one of the three sentences. It is pretty clear how this will work when all the parts of an argument are indicatives, or all the parts are imperatives. If an argument has exclusively indicative premises, then it will have an indicative conclusion, and if an argument has exclusively imperative premises, then it will have an imperative conclusion. What, though, of an argument with at least one indicative premise and at least one imperative premise? How do we know in which mood the conclusion is to be?

Hare introduces two important rules to deal with such cases (Hare (1952): 28):

- (1) No indicative conclusion can be validly drawn from a set of premises which cannot be validly drawn from the indicatives among them alone.
- (2) No imperative conclusion can be validly drawn from a set of premises which does not contain at least one imperative.

He points out that rule (2) has important implications for ethics, and has been implicitly appealed to by Aristotle, Hume, Kant, G.E. Moore, Prichard, and Ayer. It has, further, been explicitly appealed to by Henri Poincaré (1917) (225), and Karl Popper (1948) (154). There may well be important implications for ethics, but I will not discuss these implications in this thesis as it is outside the scope of the project.

### 7.1.3 Problems

First, Hare's account of imperative entailment is not a full formal definition, but merely a "rough-and-ready" test for entailment. His definition of entailment, recall, is that "a sentence  $P$  entails a sentence  $Q$  if and only if the fact that a person assents to  $P$  but dissents from  $Q$  is a sufficient criterion for saying that he has misunderstood one or other of the sentences." This cannot be a formal definition of entailment because, strictly speaking, there would never be entailment of any kind (imperative *or* indicative). A person assenting to  $P$  and dissenting from  $Q$  is never a sufficient criterion for saying that she has misunderstood one of the sentences. Even if  $P$  does entail  $Q$ , a person might understand both  $P$  and  $Q$  perfectly well but simply make a logical mistake in assenting to  $P$  while also dissenting from  $Q$ . She might be logically inept, or insane, or forgetful, or tired, or distracted. This is the only

account of entailment Hare gives in this discussion, but it cannot be thought of as a formal *definition* of entailment, as entailment is not dependent on our mental states - our assentation, dissentation, and understanding. Hare's theory is, then, not a fully developed theory: his "definition" of entailment is, then, just a rough-and-ready test for entailment.

There are also problems with Hare's rules. For instance, there is Peter Geach's notable counterexample (which we first met in section 4.3) to rule (1) (Geach (1958)):

**B1** If you are a faithful subject, rise up Sir George!

**B2** Do not rise up, stay on your knees fellow!

∴ **B3** George is not a faithful subject.

This is a case of a valid argument which has an indicative conclusion that cannot be validly drawn from its indicative premises alone, because there are no indicative premises. It requires both of the premises in order for the entailment to hold, but they are in the imperative mood. So, it violates rule (1), yet it is valid according to Hare's own account of entailment: under normal circumstances (or *ceteris paribus*), if a person assented to B1 and B2 but dissented from B3, this would be sufficient to conclude that she had misunderstood one or more of B1, B2, or B3.

Similarly, Geach (1958) also provides a counterexample to rule (2):

**C1** Grimby Hughes is the largest grocer in Oxford.

∴ **C2** Either do not go to the largest grocer in Oxford, or go to Grimby Hughes!

This is a case of a valid argument with an imperative conclusion but premises that do not contain at least one imperative. It violates rule (2), yet it is valid, again even by Hare's own test for entailment.

Because Hare's theory is not fully developed (he does not propose a formal definition of entailment but rather just an account of a rough-and-ready test for entailment), and because both of his rules to deal with mixed imperative and indicative arguments have glaring counterexamples, we can safely reject it as (probably) on the wrong track, and even if not, it is at least too far away from an actual account of imperative entailment to be worth pursuing given that there are other, more fully developed, theories that can resolve these counterexamples.

## 7.2 The Operator Analysis

A. Hofstadter and J. C. C. McKinsey (1939) put forward an early proposal for an imperative logic. They introduce a sentential operator, “!,” that turns any sentence into an imperative, as well as a set of connectives that also produce imperative sentences.

First, they draw a distinction between fiats and directives. A *directive* has a specified agent (much like Leonard’s “personal” imperatives, see section 5.7), whereas a *fiat* does not (much like Leonard’s “impersonal” imperatives. Again, see section 5.7). For example, “Emily, set the table!” is a directive, as it is directed at Emily, while “please don’t let it rain today!” is a fiat, as it has no reference to an agent; it is not directed at anybody in particular. Any directive can be converted into a fiat, but (perhaps) not without a change in meaning. In this example, the fiat-counterpart would be “let it be the case that Emily sets the table!” Hofstadter and McKinsey focus only on fiats in their syntactic analysis of imperatives, and so every directive must be turned into its fiat-counterpart. Hofstadter and McKinsey think that this exclusion of directives “involves a great limitation of subject matter and excludes topics of great interest” (Hofstadter and McKinsey (1939): 446). However, consider the following:

Take the directive:

(a) Emily, set the table!

and its fiat-counterpart:

(b) Let it be the case that Emily sets the table!

The same state of affairs will satisfy both (a) and (b). Both cases require Emily to set the table, they are satisfied (or “complied with”) if Emily sets the table, and violated if she does not do so. In both cases, it is critical that it is Emily who sets the table, not Sam or God or anyone else.

Hofstadter and McKinsey then go on to draw a second distinction: between the *satisfaction* and the *correctness* of imperatives. An imperative is satisfied when what is commanded is the case. An imperative is correct when what it requires ought to be the case, and incorrect when what it requires ought not to be the case. Satisfaction conditions (not correctness conditions) are taken to be analogous with the truth conditions of declarative sentences.

## 7.2.1 Syntax

Hofstadter and McKinsey take as their starting point Carnap's Language  $I$ , and add six new primitive symbols to get Language  $I_c$  (the first five of which result in the construction of imperative sentences, while the sixth results in the construction of a declarative sentence):

$!\phi$	imperative-creating operator	"let it be the case that $\phi$ !"
$\bar{\phi}$	imperative negation	"let it be the case that not- $\phi$ !" For example, if $\phi$ is "every place is blue," or "let it be the case that every place is blue" $\bar{\phi}$ will be "not every place is blue," or "let it be the case that not every place is blue."
$\phi + \psi$	imperative disjunction	"let it be the case that $\phi$ or $\psi$ !"
$\phi \times \psi$	imperative conjunction	"let it be the case that $\phi$ and $\psi$ !"
$\phi \rightarrow \psi$	imperative conditional	"if $\phi$ then let it be the case that $\psi$ !"
$\phi > \psi$	material inclusion	" $\phi$ materially includes $\psi$ ." This sentence is true either if $\phi$ is not satisfied or if $\psi$ is satisfied. This could be read as "either the imperative $\phi$ is not satisfied or the imperative $\psi$ is satisfied" (It constructs declarative sentences out of imperative sentences).

In addition to these new symbols, they use some of the existing symbols of Carnap's Language  $I$  in new ways:

$\phi = \psi$	material equality	" $\phi$ is satisfied if and only if $\psi$ is satisfied."
---------------	-------------------	--

The language also has universal and existential operators, which can stand before either an indicative sentence or an imperative sentence.

Syntactic definitions, for example (taking "S" to stand for statements and "C" to stand for commands), "sentence  $S_1$  is derivable from sentences  $S_2, \dots, S_n$ " and "sentence  $S_1$  is provable" remain the same as in Language  $I$ . It follows from this that every provable sentence of Language  $I$  is also a provable sentence of Language  $I_c$ .

Similarly, the definitions of terms such as "provable", "refutable", "analytic", "synthetic", and "contradictory" remain the same when denoting properties of, or relations between,

declarative sentences. They are later defined separately for imperative sentences as being parasitic on the imperative's corresponding declarative sentence. That is, if  $C_1 = !S_1$ , then  $C_1$  is provable, refutable, analytic, synthetic, or contradictory, if and only if  $S_1$  is correspondingly provable, refutable, analytic, synthetic, or contradictory.

From these theorems, and some new special theorems that come from the new primitive sentences, Hofstadter and McKinsey can demonstrate the following important theorem (where  $\equiv$  means "is equivalent to"):

**Theorem I:** If  $S_1$  is any (declarative) sentence of Language  $I_c$ , then there exists a (declarative) sentence  $S_2$ , of Language  $I$ , such that  $S_1 \equiv S_2$  is provable in Language  $I_c$ . If  $C_1$  is any imperative of Language  $I_c$ , then there exists a (declarative) sentence  $S_1$ , of Language  $I$ , such that  $C_1 = !S_1$  is provable in Language  $I_c$  (Hofstadter and McKinsey (1939): 452).

In other words, all imperative-connectives can be eliminated from a declarative sentence, and all imperative-connectives except "!" can be eliminated from an imperative sentence.

Also, it can be shown that if  $C_1 = !S_1$  and  $C_1 = !S_2$  are both provable, then  $S_1 \equiv S_2$  is also provable (so  $S_1$  is unique with respect to derivability).

## 7.2.2 Problems

The first problem with Hofstadter and McKinsey's approach is that they fail to explain what truth-functional connectives and their imperative counterparts have in common. They outline a set of connectives that only connect imperative sentences, as distinct from the regular set of connectives that connect declarative sentences. Their system, then, has two distinct negations, conjunctions, disjunctions, and conditionals. They have, for example, a declarative conjunction and an imperative conjunction, and they are defined separately. In natural language, however, both of these conjunctions are (usually) expressed using the word "and." We say, for example,

(c) "Emily will clear the table *and* do the dishes," and also

(d) "Emily, clear the table *and* do the dishes!"

The same word, "and," appears in both the declarative sentence (c) and the imperative sentence (d). It also appears to have a similar function in each case – there is something about the word "and" that makes sentences (c) and (d) more similar than, say, (c) and (e):

(e) “Emily, clear the table *or* do the dishes!”

That is, in (c) and (d), the “and” seems to be connecting what comes before and after it in much the same way, in a *conjunction-like* way. It says “*both of these.*” On the other hand, the “or” in (e) is performing a different function. My point is that in Hofstadter and McKinsey’s theory, the “and” in (c) and the “and” in (d) are completely different connectives, which happen to use the same word in English. However, this fails to explain why (c) and (d) *do* use the same word in English, given the two instances of “and” also appear to have the same meaning and function. Hofstadter and McKinsey do not explain what they have in common. The same objection applies to their distinguishing between declarative and imperative negations, disjunctions, and conditionals.

### 7.2.3 Imperative Derivability and Consequence

The really important part of Hofstadter and McKinsey’s theory of the logic of imperatives is how they define imperative derivability and imperative consequence. As we have seen, the definitions of imperative provability, refutability, and so forth are parasitic on the imperative sentence’s counterpart declarative sentence. However, next we hear that derivability and consequence are as well. It works as follows:

Suppose that  $C_1 = !S_1$  and  $C_2 = !S_2$  are provable. Then we call  $C_2$  *derivable from*  $C_1$ , if  $S_2$  is derivable from  $S_1$ , and  $C_2$  is a consequence of  $C_1$ , if  $S_2$  is a consequence of  $S_1$ ; if  $C_2$  is a consequence of  $C_1$  and  $C_1$  is a consequence of  $C_2$  then  $C_1$  and  $C_2$  are said to be *equipollent*. (Hofstadter and McKinsey (1939): 452).

It can easily be seen, because imperative consequence is parasitic on declarative derivability in this way, that all the various syntactic concepts (defined for declarative sentences) can be extended to cover imperatives as well. Hofstadter and McKinsey go on to admit that their theory shows that the introduction of the “!” is superfluous. That is,  $\phi \leftrightarrow !\phi$ , which is clearly undesirable. However, they do not see this as a problem, but rather a positive feature: “we are inclined to believe that this result, far from being a defect in our theory, is rather a recommendation” (Hofstadter and McKinsey (1939): 453). They do not offer any argument for this, or elaborate on the claim except to say “we feel that it would be desirable, when introducing the mark “!” into almost any sort of language, to assume sufficiently

strong primitive sentences and rules, in order to be able to prove such a theorem as ours” (Hofstadter and McKinsey (1939): 453).

So Hofstadter and McKinsey’s theory renders their own imperative connectives, and then even their imperative operator, superfluous. In effect, this leaves their logic as not a logic of imperatives at all. Rather, it is a logic of declaratives with a supplementary principle about the imperative mood: that an imperative is “true” if it is complied with and “false” if it is not complied with.

This amounts to a version of solution two – that imperatives can be truth-apt. In particular, it amounts to the predictions theory, which, recall, is a version of cognitivism where an imperative is thought to simply be a second person future declarative sentence. For example, “do x!” is thought to be translatable to “you will do x” without a change in meaning. Because these sentences are held to be identical in meaning, it can be said that an imperative is true if and only if its corresponding second person future declarative sentence is true. This, of course, amounts to the same thing as saying an imperative is true if and only if it is complied with.

This all highlights two important things about Hofstadter and McKinsey’s theory. First, it is not really a logic of *imperatives* at all, and so it is not a version of solution 3. It is, in an important sense, *eliminativist* with respect to the logic of imperatives. That is, Hofstadter and McKinsey begin by taking seriously the idea that there can be logical relations between imperatives, but in attempting to formalise these relations, they come to the result that imperatives have no place in logic. This is not an objection to the theory, as such, but rather it is an objection to this theory as a theory of imperative logic. In particular, a theory of imperative logic that claims to take seriously the idea that logical relations can hold between imperative sentences.

Second, because their theory can be seen to be in effect identical to the prediction theory, all the objections to the prediction theory can also be held against Hofstadter and McKinsey’s theory. This, then, is an objection (or rather, a series of objections) to Hofstadter and McKinsey’s theory. Even if we ignore the fact that sentences in the imperative mood are being eliminated from their logic of imperatives, it is still inadequate because it gets the logic (of these imperative sentences) wrong. Refer to chapters 5 and 6 for a full discussion of these objections, but (in short) it renders as valid arguments such as:

- D1** Do the dishes, Emily!
- $\therefore$  **D2** Emily will do something.

In the case of Hofstadter and McKinsey's theory, we can show the hidden steps that make it valid as follows:

**D1** Do the dishes, Emily!

**D1'** Let it be the case that Emily will do the dishes!

**D1''** Emily will do the dishes.

∴ **D2** Emily will do something.

The step from D1 to D1' comes from the translation from a directive into its fiat-counterpart, which is how Hofstadter and McKinsey propose to treat imperatives. The step from D1' to D1'' comes from the result that the "!" operator is superfluous and !p can always be translated to simply p without a change in logical properties.

On first glance, this conclusion, D2, appears not to follow from the command, D1. And because there is no argument given that either a) our initial intuitions are wrong and this should be valid, or b) there are sufficient advantages in treating imperatives in this way to justify such a counterintuitive result, there is not sufficient reason to reject our initial assessment.

Hofstadter and McKinsey's theory, then, fails to account for the common functions of the connectives in declarative and imperative sentences. It collapses into a form of cognitivism, the Predictions theory, and as such is not a version of solution 3; it is not really a logic of *imperatives* at all.

Next, we will consider formal logics of imperatives that introduce obedience-conditions as an imperative analogue for truth-conditions.

# Chapter 8

## Obedience and Termination Validity

There is an obvious candidate for an account of imperative consequence that may be immediately appealing. It seems that, although imperatives do not have *truth*-conditions, they do have *obedience*-conditions. That is, they can be obeyed or they can fail to be obeyed. So, we could formulate a logic of imperatives with obedience-values as an analogue of truth-values. I will discuss two versions of this theory.

### 8.1 Obedience-Validity

Charles Pigden (unpublished) provides an account that is based on that of Jack Smart (1984). Pigden summarises Smart's account as follows:

The basic idea behind imperative consequence is that an imperative **X** is the consequence of a set of imperatives **K** iff the premises cannot be obeyed without obeying the imperative that constitutes the conclusion **X** (Pigden (unpublished): 3. See Smart (1984): 14-19).

This covers imperative arguments where all the premises and the conclusion are imperatives, such as:

- A1** Open the window!
- A2** Close the door!
- $\therefore$  **A3** Open the window and close the door!

However, it doesn't cover other kinds of arguments, such as those with an imperative conclusion and a mix of imperative and indicative premises, such as:

- B1** If the weather is fine, attack at dawn!
- B2** The weather is fine.
- $\therefore$  **B3** Attack at dawn!

So, Pigden gives a more formal definition that allows for indicative premises, the set of which he calls the *circumstances* of the imperative argument. The new definition is:

An imperative sentence **X** is the logical consequence of a set of imperatives **K** and a (possibly empty) set of factual sentences **C**, if and only if, however the non-logical vocabulary is interpreted, the imperatives expressed at **K** cannot be obeyed under the circumstances described in **C** unless the imperative expressed by **X** is obeyed too (Pigden (unpublished): 3).

Pigden then refines Smart's conception of imperative validity as obedience-preserving. He points out that it has the "bizarre consequence that any command whatsoever is a consequence of a set of commands **K** if that set is empty," or an inconsistent set of commands, or a set of commands that cannot be obeyed under circumstances **C**. For example, consider the following argument:

- C1** Attack at dawn!
- C2** Do not attack at dawn!
- $\therefore$  **C3** Dance a jig!

Pigden is unhappy with this type of argument, call it "imperative explosion." That is, he is unhappy with the kind of argument that goes from an inconsistent set of (imperative) premises to any conclusion whatsoever. So, he makes his version of imperative consequence *non-monotonic*, to safeguard against this type of argument coming out as valid. He suggests "adding the proviso that a set of imperatives **K** only implies another imperative **X** if it logically *possible* (under the circumstances **C**) to obey the commands **K**" (Pigden (unpublished): 4).

Finally, Pigden considers Geach's king's modus tollens (which we originally encountered in section 4.3):

- D1** If you are a faithful subject, rise up Sir George!
- D2** Do not rise, stay on your knees fellow!
- $\therefore$  **D3** You are not a faithful subject.

This is a compelling case of an argument with an indicative conclusion that appears to follow validly from a set of entirely imperative premises. So, Pigden modifies his definition of imperative consequence one last time. First, he points out that it won't do to simply say that "an indicative  $\mathbf{X}$  is the consequence of a set of imperatives  $\mathbf{K}$  if  $\mathbf{K}$  cannot be obeyed unless  $\mathbf{X}$  is true" (Pigden (unpublished): 5), because this would have the unfortunate implication that every command has the consequence that it is obeyed. A command cannot be obeyed without the proposition *describing* the state of affairs of its obedience becoming true.

To combat this result, Pigden first adds the proviso that the conclusion must be deemed to be within the agent's power. This is to ensure we don't conclude, from D1 and D2:

∴ **D3\*** Do not be a faithful subject!

He then adds a definition of imperative-to-indicative consequence:

A set of imperatives  $\mathbf{K}$  implies an imperative  $\mathbf{X}$  iff a) the imperatives in  $\mathbf{K}$  can be obeyed, b) the imperatives in  $\mathbf{K}$  cannot be obeyed without obeying  $\mathbf{X}$  and [c)] obeying  $\mathbf{X}$  is deemed possible with respect to the agent.

A set of imperatives  $\mathbf{K}$  implies an *indicative*  $\mathbf{X}'$  iff a)  $\mathbf{K}$  can be obeyed, b)  $\mathbf{K}$  cannot be obeyed unless  $\mathbf{X}'$  is true and c) bringing about the state of affairs described by  $\mathbf{X}'$  is *not* deemed to be possible for the agent (Pigden (unpublished): 5).

Because it doesn't have to just be *logically* impossible, it could just be *physically* impossible for the addressee, Pigden adds a further proviso that  $\mathbf{X}$  is a consequence of  $\mathbf{K}$  *with respect to some agent (or agents)*. Here is his final definition of imperative-to-indicative and imperative and indicative-to-imperative validity:

### **Smart/Pigden Theory - two definitions of validity**

#### **Part 1 (Imperative-to-indicative validity):**

An indicative conclusion  $\mathbf{Y}$  is the consequence of a set of imperative premises  $\mathbf{K}$  with respect to some agent (or agents)  $\mathbf{P}$  if and only if i) it is logically possible to obey the imperatives in  $\mathbf{K}$ , and ii) the imperatives in the premises  $\mathbf{K}$  cannot be obeyed unless the indicative  $\mathbf{Y}$  is true and iii) it is not deemed to be within the agent's power to bring about the indicative conclusion (Pigden (unpublished): 6).

**Part 2 (Imperative and indicative-to-imperative validity):**

An imperative conclusion **X** is the consequence of a set of imperatives **K** with respect to some agent (or agents) **P**, and a possibly empty set of indicative premises **C**, if and only if i) the premises **K** cannot be obeyed under the circumstances **C** without obeying the imperative that constitutes the conclusion **X**, ii) it is logically possible to obey the commands **X** under the circumstances **C**, and iii) obeying **X** is deemed to be within the power of the agent or agents **P** (Pigden (unpublished): 6).

## 8.2 Problems

### Valid arguments with imperative and indicative premises and indicative conclusions

Consider the following argument:

- E1** If the weather is fine and you see the enemy, attack!
- E2** You see the enemy.
- E3** Do not attack!
- $\therefore$  **E4** The weather is not fine.

Pigden's definition either doesn't cover this argument, or it makes it come out as invalid, because it has a mix of both imperative and indicative premises leading to a indicative conclusion. But there is no reason not to suppose that this argument is valid. So, we must modify the first part of his definition to:

**Part 1' (Imperative and indicative-to-indicative validity):**

An indicative conclusion **Y** is the consequence of a set of imperative premises **K** with respect to some agent (or agents) **P**, and a possibly empty set of indicative premises **C** if and only if i) it is logically possible to obey the imperatives in **K** under the circumstances **C**, and ii) the imperatives in the premises **K** cannot be obeyed under the circumstances **C** unless the indicative **Y** is true and iii) it is not deemed to be within the agent's power to bring about the indicative conclusion.

This validates argument E. However, now consider argument F:

- F1** The window is open.
- F2** Make it rain!
- ∴ **F3** The window is open.

This argument should come out as valid. The conclusion F3 is just the same sentence as the premise F1. It is irrelevant what other premises are added, provided they do not contradict the other premises, as Pigden’s nonmonotonicity condition dictates. F2 does not contradict F1, so it should not affect the validity of the argument. Consider, for comparison, the following (imperative-less) analogue:

- G1** The window is open.
- G2** The flowers are blooming.
- ∴ **G3** The window is open.

Any plausible definition of (regular, indicative) validity would tell us that argument G is valid. Is argument F not analogous to argument G? Argument F, just like argument G, should be valid. However, according to the new definition of imperative and indicative to indicative validity, it is *invalid*, because it does not fulfil requirement iii); that is, it *is* entirely within the agent’s power to bring about the indicative conclusion.

Now consider argument H:

- H1** The weather is fine.
- H2** Attack and don’t attack!
- ∴ **H3** The weather is fine.

This example meets requirement iii): the conclusion H3 is not deemed to be within the agent’s power to bring about. It *does* defy the nonmonotonicity condition. That is, it fails to fulfil requirement i) of Part 1’ of the definition of validity (imperative and indicative to indicative validity). So, it also comes out as invalid. However, although it is logically impossible to obey H2, this premise is not doing any work in getting from H1 to H3. So, it shouldn’t affect the validity of the argument. So, the modified definition (Part 1’) that allows a indicative conclusion to follow from mixed imperative and indicative premises gives the wrong results in arguments F and H. If we go back to the original definition (Part 1), which effectively disallows this kind of argument, then not only will arguments F and H *still* not be valid, but neither will argument E, which doesn’t have redundant premises and should, as surely as any case of imperative validity, come out as valid.

## Valid arguments with premises that are logically impossible to obey

Argument H is one example of this, but it is also a problem for the *second* part of Pigden's final definition. Consider the following argument:

- I1** Open the window!
- I2** Attack and don't attack!
- $\therefore$  **I3** Open the window!

In this case, again, the second premise (I2) is logically impossible to obey, and yet it does no work in getting to the conclusion (it is redundant). It is odd, and undesirable, that the addition of a redundant premise can render a valid argument invalid.

## Valid arguments where the agent is deemed to be incapable of obeying the imperative conclusion

Consider the following argument:

- J1** Whenever a woman enters the room, stand up!
- J2** A woman has entered the room.
- $\therefore$  **J3** Stand up!

This appears to be a valid argument. But suppose that we learn that the addressee is paraplegic, confined to a wheelchair, and as such it is not within his power to stand up. The issuer of J1, to be sure, is probably being obnoxious, but this should not affect the validity of the argument. Part 2 of Pigden's definition implies that this is the case; that is, this argument is valid when the addressee (agent) is able-bodied, but invalid when the agent is paraplegic. Worse, if we accept the modified version (Part 1') of the definition then when the addressee is paraplegic the following argument is valid instead:

- K1** Whenever a woman enters the room, stand up!
- K2** A woman has entered the room.
- $\therefore$  **K3** You will stand up.

Argument K meets the three requirements of the modified definition of Part 1' (imperative and indicative to indicative validity):

- i) it is logically possible to obey the imperative in K1 under the circumstance K2.

- ii) the imperative in the premise K1 cannot be obeyed under the circumstances K2 unless the indicative K3 is true.
- iii) it is not deemed to be within the agent's power to bring about the indicative conclusion, K3.

In fact, an even simpler example will suffice. Consider:

- L1 Stand up!
- ∴ L2 Stand up!

And:

- M1 Stand up!
- ∴ M2 You will stand up.

When the agent is paraplegic, argument L is invalid while argument M is valid. The problem occurs because Pigden does not specify that the imperative *premises* must also be within the power of the agent to obey. So, here is a further modified definition:

**Part 1'' (imperative and indicative to indicative validity):**

An indicative conclusion **Y** is the consequence of a set of imperative premises **K** with respect to some agent (or agents) **P**, and a possibly empty set of indicative premises **C** if and only if i) it is logically possible, and deemed to be within the agent's power, to obey the imperatives in **K** under the circumstances **C**, and ii) the imperatives in the premises **K** cannot be obeyed under the circumstances **C** unless the indicative **Y** is true and iii) it is not deemed to be within the agent's power to bring about the indicative conclusion.

**Part 2' (imperative and indicative to imperative validity):**

An imperative conclusion **X** is the consequence of a set of imperatives **K** with respect to some agent (or agents) **P**, and a possibly empty set of indicative premises **C**, if and only if i) the premises **K** cannot be obeyed under the circumstances **C** without obeying the imperative that constitutes the conclusion **X**, ii) it is logically possible, and deemed to be within the agent's power, to obey the commands **X** under the circumstances **C**, and iii) obeying **X** is deemed to be within the power of the agent or agents **P**.

However, now we just generate more problems. The following arguments all come out as *invalid* (except when addressed to God, or at least some agent with the power to make it rain):

**N1** Make it rain and open the window!  
 $\therefore$  **N2** Open the window!

**O1** Make it rain and open the window!  
 $\therefore$  **O2** Make it rain!

**P1** Make it rain!  
 $\therefore$  **P2** Make it rain!

Quibbling about the details of the requirements aside, the Smart/Pigden theory suffers from two more fundamental problems. These problems are shared with the second version of this type of theory, that of Nicholas Rescher, which I will outline now.

### 8.3 Termination-Validity

Nicholas Rescher (1966) offers an account of imperative validity in terms of *termination*. Termination is very similar to obedience, but this theory was designed to apply to computer programs, hence the term “termination” instead of “obedience.” His definition of termination relies on what he calls a command’s “command termination statement,”  $T_t(C)$ . This is the statement that describes the state of affairs when a command has been obeyed. Where  $X$  is an agent,  $A$  is an action, and  $P$  is any condition in place for when the action is to be carried out (for example, “always,” “tomorrow after work,” or “whenever you get a paper cut”), we can represent a command,  $C$ , as:

$$C = [X ! A/P]$$

This says “Agent  $X$ , do action  $A$  under condition  $P$ !” Then, Rescher’s command termination statement is defined, precisely, as follows:

If a command

$$C = [X ! A/P]$$

is considered to be given at some time  $t$ , one can form the corresponding (purely assertoric) *command termination statement*:

“ $T_t(C)$ ” for “*has (always) realized A whenever P obtained after time t*”  
(Rescher (1966): 52).

For example, suppose the following command  $C$  is given at time  $t$ :

“Tom, tell Jim to see me when (next) you see him!”

$X$  is Tom,  $A$  is telling Jim to see me, and  $P$  is Tom sees Jim. Thus, the corresponding  $T_t(C)$  is:

“Tom did tell Jim to see me when next (after  $t$ ) he saw him.”

Rescher then defines termination (and explains the concept further) as follows:

Given a command  $C$ , we shall say that this command is *terminated* as of time  $t$  whenever  $T_t(C)$  is true. A conditional command whose execution precondition never arises will be regarded as automatically terminated (in a trivial way, if you please). Also, we shall say that “a person *terminates* a command” (as of  $t$ ) when his behaviour is such as to render the appropriate termination statement true. (Note that a standing order only becomes terminated posthumously.) While commands themselves are *prospective* and *ante eventum*, their termination statements are always *retrospective* and *post eventum* (Rescher (1966): 53).

He distinguishes between *homogenous* imperative arguments (when the premises contain only commands) and *heterogenous* imperative arguments (when the premises also contain assertoric sentences), and then he outlines three considerations for a definition of validity to have:

- (i) *Anyone* who overtly *gives* the premise commands may legitimately claim (or be claimed) to have implicitly given the command conclusion.
- (ii) *Anyone* who overtly *receives* the premise commands may legitimately claim (or be claimed) to have implicitly received the command conclusion.

- (iii) Any course of action on the part of their command recipient which terminates the premise commands cannot fail to terminate the command conclusion (Rescher (1966): 77-78).

He then defines invalidity, or what he calls the “rule of rejection”:

A command inference is *patently invalid* if it is possible for all its assertoric premises to be true and for all of its command premises to be terminated, and yet its command conclusion to remain unterminated (Rescher (1966): 78).

And, finally, Rescher defines his notion of patent validity:

The inference whose conclusion is the command  $C$  and whose premises include the commands  $C_1, C_2, \dots, C_n$  is *patently valid* if the command  $C$  can be decomposed into the set of commands  $C_1^\#, C_2^\#, \dots, C_m^\#$  in such a way that each  $C_i^\#$  is covered by some  $C_j$  (Rescher (1966): 79).

For example, take the argument:

- Q1** Do A & B!  
**Q2** Do C & D when P!  
 $\therefore$  **Q3** Do A & C when P!

Q3 can be decomposed into:

- a) Do A when P!  
b) Do C when P!

Then, a) is “covered by” Q1 and b) is “covered by” Q2.

Rescher then goes on to extend this definition so that it also covers heterogenous imperative arguments. He defines his so-called *contextual validity* as follows:

The inference whose conclusion is the command  $C$  and whose premises include the commands  $C_1, C_2, \dots, C_n$  is *valid in the context of* (or read: *contextually valid given*) the assertoric premises  $S_1, S_2, \dots, S_m$  if the command conclusion  $C$  can be decomposed — either absolutely or in the context of  $S_1, S_2, \dots, S_m$  — into the set of commands  $C_1^\#, C_2^\#, \dots, C_k^\#$  in such a way that each  $C_i^\#$  is covered by some of the  $C_j$  either by simple coverage or by contextual coverage given the assertoric statements  $S_1, S_2, \dots, S_m$  (Rescher (1966): 80).

### 8.3.1 Problems

First, this theory does not cover all types of argument. In particular, it does not cover arguments with an indicative (assertoric) conclusion but at least one imperative premise. Recall (again) argument D:

- D1** If you are a faithful subject, rise up Sir George!
- D2** Do not rise, stay on your knees fellow!
- $\therefore$  **D3** You are not a faithful subject.

This argument is an example of this; it has two imperative premises but an indicative conclusion. Rescher's theory is silent on arguments like argument D.

## 8.4 Deeper Problems with both Obedience- and Termination-Validity

Rescher's theory of validity has "termination-conditions" as an analogue of truth-conditions for imperative sentences. That is, in Rescher's theory, sentences are classified as either imperative or declarative ("commands" or "assertoric sentences"), and then arguments are analysed in terms of the sentences' termination-conditions and their truth-conditions, respectively. This is very similar to the Smart/Pigden theory, that has "obedience-conditions" as the analogue of truth-conditions for imperative sentences. "Termination" occurs when the command has been obeyed, so it is really another term for the same idea. So, although the details of the two theories differ, they have enough in common to both suffer from the same two fundamental problems:

### 8.4.1 Partitioning Language into "Imperative" and "Indicative"

First, both Rescher's and the Smart/Pigden view require that all sentences be either categorically imperative or categorically indicative, which raises the problem of how to classify *mixed* sentences. By "mixed" sentences, I mean complex sentences (containing connectives), with at least one imperative part and at least one indicative part. For example:

- (a) It is raining and the door is closed and the window is open and set the table!
- (b) It is raining and close the door and open the window and set the table!

- (c) It is raining or the door is closed or the window is open or set the table!
- (d) It is raining or close the door or open the window or set the table!

Because there are distinct definitions of validity in both theories (that is, when there are no imperatives in the argument, we are to revert to classical logic's definition of validity in terms of truth-conditions, and when there are imperatives the theories each give a different definition of validity to use), it matters how sentences like these are classified so that we know which definition to apply when they are used as parts of arguments. For example, consider the following argument:

**R1** It is raining.

**R2** The door is closed.

**R3** The window is open.

**R4** Set the table!

∴ **R5** It is raining and the door is closed and the window is open and set the table!

If we categorise R5 as indicative, this argument is an indicative-and-imperative to indicative argument, but if we classify it as an imperative, then the argument is an indicative-and-imperative to imperative argument. If it is indicative, then (taking Part 1'' of the Smart/Pidgen definition), although it is (presumably) within the agent's power to set the table under the circumstances that it is raining, the door is closed and the window is open, it is invalid because it does not meet criterion ii). That is, the imperative premises can be obeyed under the circumstances described in R1-R3 without making the conclusion true, because the conclusion has an imperative part and as such cannot be true. Rescher's theory is silent on argument R, if R5 is taken to be indicative.

If, on the other hand, we categorise R5 as imperative, then (taking Part 2' of the Smart/Pigden definition), because the imperative premise R4 cannot be obeyed under the circumstances R1-R3 without also obeying the conclusion R5, and because it is logically possible and deemed within the agent or agents' power to obey the conclusion R5 under the circumstances R1-R3, this argument is valid. Although it would be strange to say that R5 can be obeyed at all (because it has indicative parts which can no more be obeyed or not obeyed than imperatives can be true or false), I will leave this issue for now so as to be able to discuss one issue at a time. Similarly, according to Rescher's definition, all the imperatives in the premises (in this case, only R4) are covered by the conclusion (in conjunction with the "contextual" indicative premises), so it is valid.

So, if R5 is categorised as indicative then this argument is invalid (or ignored), but if it is categorised as imperative then it is valid. We need a principled way to categorise sentences of this kind, otherwise they will have to be categorised on a case-by-case basis and this will render each logic even less usable than it currently is.

Because being valid is clearly the correct answer for argument R, let's say that a sentence is imperative whenever it has even one imperative element in it. This will make R5 come out as imperative, and so argument R come out as valid. However, now consider the following arguments:

**S1** It is raining and the door is closed and the window is open and set the table!  
∴ **S2** It is raining.

**T1** It is raining and the door is closed and the window is open and set the table!  
∴ **T2** The door is closed.

S1 and T1 (which are both identical to R5), have an imperative element so must be categorised as imperatives. Rescher's definition is, again, silent on both argument R and argument S because they have indicative conclusions. As for the Smart/Pigden view, we must look to Part 1'' of the definition, dealing with arguments with imperative and (a possibly empty set of) indicative premises and an indicative conclusion. Now though, argument S comes out as valid (unless addressed to God) while argument T comes out as invalid (unless, perhaps, addressed to a baby or other person with suitably limited agency), because of criterion iii): it can only be valid if it is not deemed to be within the agent's power to bring about the indicative conclusion. S2 meets this condition; it is indeed outside the agent's power to bring it about that it is raining, unless the agent is God, in which case the argument would be invalid. T2 does not meet this condition, as it (usually) *is* within the agent's power to make it the case that the door is closed, unless the agent is (for some reason) incapable of bringing this about.

These are very disturbing results, partly because validity isn't the sort of thing that should be affected by contextual facts such as to whom the sentences are uttered, but mostly (I think), because the imperative conjunct in "It is raining and the door is closed and the window is open and set the table!" ("set the table!"), is playing no part in the concluding of S2 or T2, so the arguments *should* be valid or invalid just as the following arguments are valid or invalid:

**U1** It is raining and the door is closed and the window is open.  
∴ **U2** It is raining.

**V1** It is raining and the door is closed and the window is open.  
∴ **V2** The door is closed.

Arguments U and V have no imperative elements in any of the parts, so are subject (according to both Rescher's and the Smart/Pigden view) to standard classical logic. As such, they are both valid (no matter who is speaking or listening). Argument U is just argument S, and argument V is just argument T, each with the imperative element in the premise removed. The imperative elements in arguments S and T play no role whatsoever in getting to the respective conclusions, yet are inexplicably transformed into arguments that are valid or invalid depending on to whom it is addressed.

#### 8.4.2 Validity as Nothing More than Disjunctive

Finally, both Rescher's and the Smart/Pigden view suffer from a deeper theoretical problem: they fail the criterion of generality. That is, they give two distinct definitions of validity, and they don't explain what the different kinds of validity have in common. Recall two of our motivating examples of valid imperative arguments:

**A1** Open the window!  
**A2** Close the door!  
∴ **A3** Open the window and close the door!

**B1** If the weather is fine, attack at dawn!  
**B2** The weather is fine.  
∴ **B3** Attack at dawn!

Initially, it was thought that arguments like A and B are valid *in the same way* as their purely indicative counterparts. They are of the same form as the familiar indicative instances of conjunction introduction and modus ponens. For example:

- W1** The window is open.  
**W2** The door is closed.  
∴ **W3** The window is open and the door is closed.

- X1** If the weather is fine, you will attack at dawn.  
**X2** The weather is fine.  
∴ **X3** You will attack at dawn.

We thought that their conclusions (A3 and B3) *follow from* their premises (A1 & A2, and B1 & B2) *in the same sense* that the conclusions of arguments W and X follow from their respective premises. In all four arguments, if we accept their premises, then we are *compelled* to accept their conclusion. We are compelled, in each case, *in the same way*. We are *logically* compelled. A purely disjunctive definition, such as Rescher's or the Smart/Pigden definition, fails to explain this commonality in the different types of argument. The relation of "being a consequence of" is a different relation in each case, because it is defined in a different, non-unified, way.

No doubt they would reply that consequence is a family resemblance concept and that an imperative need only follow from a set of imperatives in *roughly* the same way that a declarative follows from a set of declaratives. Perhaps, if we cannot get a unified conception of consequence then that would be a sufficient answer. However, if a unified conception *can* be constructed then it is to be preferred to a disjunctive conception of consequence bearing only a family resemblance to one another. It is just this kind of unified conception that I attempt to construct in chapter 12.

# Chapter 9

## Validity in Terms of Reasons for Obeying

Peter Vranas (2008, 2010, 2012, 2013) puts forward a theory of imperative logic that is based on *reasons for obeying* commands, rather than obedience-conditions. In order to understand his theory, we first must become clear on his terminology.

### 9.1 Terminology

First, he distinguishes between *propositions*, which are the (familiar) things that declarative sentences typically express, and *prescriptions*, which are what imperative sentences typically express. Specifically, for Vranas, a prescription is an ordered pair of (logically) incompatible propositions; a *satisfaction* proposition (a proposition that is true when the prescription is satisfied) and a *violation* proposition (a proposition that is true when the prescription is violated) (Vranas (2008): 6, (2012): 3). The *disjunction* of the satisfaction proposition and the violation proposition is called the *context* of the prescription, and the *negation* of the context is called the *avoidance proposition* (Vranas (2008): 10). For example, take the following two imperative sentences:

- (a) Open the window!
- (b) If it rains, bring in the washing!

The first is just a simple imperative (that is, it has no connectives), and the second is complex: a conditional imperative. Here are each of the elements as Vranas outlines them:

<b>Imperative Sentence</b>	Open the window!
<b>Satisfaction Proposition</b>	You open the window.
<b>Violation Proposition</b>	You do not open the window.
<b>Context</b>	Either you open the window or you do not open the window.
<b>Avoidance Proposition</b>	It is not the case that either you open the window or you do not open the window.

<b>Imperative Sentence</b>	If it rains, bring in the washing!
<b>Satisfaction Proposition</b>	It rains and you bring in the washing.
<b>Violation Proposition</b>	It rains and you do not bring in the washing.
<b>Context</b>	Either it rains and you bring in the washing, or it rains and you do not bring in the washing.
<b>Avoidance Proposition</b>	It is not the case that either it rains and you bring in the washing, or it rains and you do not bring in the washing.

Notice that for the first sentence, “open the window!”, the avoidance proposition is necessarily false. This is true of all simple imperatives, and so according to Vranas they can only ever be satisfied or violated, never avoided. On the other hand, the avoidance proposition for the conditional, “if it rains, bring in the washing!”, will be true when it does not rain. Thus, conditional imperatives are avoided when their antecedent is false. This has theoretical appeal: conditional commands can be thought of as commands that only become “activated” when their antecedent is true. If it does not rain, then it is not activated, so there is no command in force for you to satisfy or violate. So, you avoid having to do anything.

It is worth noting that this theory does not deviate from the usual result that conditionals are equivalent to disjunctions. For example, the disjunction that corresponds to (b) is:

(c) Either it does not rain or bring in the washing!

This can be decomposed into its satisfaction proposition, context, and avoidance proposition as follows:

<b>Imperative Sentence</b>	Either it does not rain or bring in the washing!
<b>Satisfaction Proposition</b>	It rains and you bring in the washing.
<b>Violation Proposition</b>	It rains and you do not bring in the washing.
<b>Context</b>	Either it rains and you bring in the washing, or it rains and you do not bring in the washing.
<b>Avoidance Proposition</b>	It is not the case that either it rains and you bring in the washing, or it rains and you do not bring in the washing.

These are, of course, identical to those of the conditional command.

As his theory is based on *reasons for obeying*, we also need to be clear on what he means by “reasons,” and some related concepts. Generally, a *reason* is some fact that counts in favour of - in short, that *favours* - some proposition over some other one (Vranas (2012): 4). For example, a *reason to bring in the washing* is a fact that favours the proposition that you bring in the washing over the proposition that you do not bring in the washing. A reason *supports* a prescription if and only if it favours the satisfaction proposition over the violation proposition of the prescription. A fact *guarantees* a proposition iff, necessarily, the proposition is true if the fact exists. A fact *sustains a proposition* iff it guarantees it. A fact *sustains a prescription* iff it supports it (Vranas (2012): 5). A fact *conclusively*—or *indefeasibly*—supports a prescription exactly if, necessarily, if the fact exists then it undefeatedly supports the prescription. This is what Vranas means by having a “conclusive” reason to do something (Vranas (2012): 7-8).

This idea of *sustaining* is important because, according to Vranas, propositions and prescriptions each merit (either pro-tanto or all-things-considered) endorsement when they are sustained by some fact. That is, a proposition *merits endorsement* iff it is true - equivalently, iff it is *sustained*, i.e. *guaranteed* by some fact. A prescription *merits endorsement* iff it is supported by reasons - equivalently, iff it is *sustained*, i.e. *supported* by some fact (Vranas (2012): 5-6). A prescription merits *pro-tanto* (i.e. *prima facie*) endorsement iff it is supported by *some* reason. A prescription merits *all-things-considered* endorsement iff it is supported by *a conclusive* reason (Vranas (2012): 7).

A final terminological distinction is that between *strong* and *weak* support:

A fact (1) *strongly supports* a prescription exactly if it favours every proposition which entails the satisfaction proposition of the prescription over every proposition which entails the violation proposition of the prescription, and (2) *weakly supports* a prescription  $I$  exactly if it strongly supports some  $I^*$  whose satisfaction proposition entails the satisfaction proposition of  $I$  and whose context is the same as the context of  $I$  (Vranas (2012): 28).

For example, if you have promised to buy me lunch today, then that fact strongly and weakly supports the prescription “buy me lunch today!”, and weakly but not strongly supports the prescription “buy me lunch!” This distinction is important because Vranas defines many concepts in terms of support: he defines sustaining, being a reason for, meriting endorsement, being valid, entailing, following from, and being consistent all in terms of supporting. Thus, there are strong and weak versions of each of these. For prescriptions, he defines *strong sustaining* as strong support and *weak sustaining* as weak support (Vranas (2012): 28). Similar definitions can be devised for the other concepts. Validity and entailment are important ones, which I will get to in section 9.3.

## 9.2 Types of Argument

Vranas makes a distinction between 6 different kinds of argument. There are *declarative* arguments (those with a declarative conclusion) and *imperative* arguments (those with imperative conclusions). Then, within each of these kinds of arguments there can be *pure* arguments (those with premises and conclusions of the same sort), and *mixed* arguments (those with premises of a different kind than their conclusion). Finally, within mixed arguments, there are *mixed-premise* arguments (mixed arguments with at least one declarative premise and at least one imperative premise), and *cross-species* arguments (those with premises entirely of the opposite kind as the conclusion). Here are the six kinds of argument, in a table for clarity, with Vranas’s examples (Vranas (2012): 2):

	<b>Declarative arguments</b> (the conclusion is a proposition)	<b>Imperative arguments</b> (the conclusion is a prescription)
<b>Pure</b>	<i>Pure declarative arguments</i> e.g. He sang two songs ∴ He sang at least two songs	<i>Pure imperative arguments</i> e.g. Honour your mother and your father ∴ Honour your mother
<b>Mixed-premise</b>	<i>Mixed-premise declarative arguments</i> e.g. If he loves you, marry him You will not marry him ∴ He doesn't love you	<i>Mixed-premise imperative arguments</i> e.g. If you love him, marry him You love him ∴ Marry him
<b>Cross-species</b>	<i>Cross-species declarative arguments</i> e.g. Don't touch my sister ∴ I have a sister	<i>Cross-species imperative arguments</i> e.g. You have a conclusive reason to dance ∴ Dance

Vranas points out that every pure argument (no matter how many premises) and every cross-species argument can be reduced to a single-premise argument by forming a conjunction of all the premises. Similarly, every mixed-premise argument can be reduced to a two-premise argument – one will be the conjunction of all the declarative premises and one will be the conjunction of all the imperative premises (Vranas (2012): 3-4). Consequently, he can treat all pure and cross-species arguments as single-premise arguments, and all mixed-premise arguments as two-premise arguments.

### 9.3 Definitions of Validity

Vranas has six sub-definitions of validity, one for each type of argument. All of these definitions can be generalised into one *General Definition of Argument Validity*:

An argument is (deductively) valid—i.e., its premises *entail* its conclusion; equivalently, its conclusion *follows* from its premises—exactly if, necessarily, every fact that sustains every premise of the argument also sustains the conclusion of the argument (Vranas (2012): 9).

Here are the six sub-definitions of validity:

	<b>Declarative arguments</b> (the conclusion is a proposition)	<b>Imperative arguments</b> (the conclusion is a prescription)
<b>Pure arguments</b>	A <i>pure declarative argument</i> is valid iff every fact that guarantees the premise of the argument also guarantees the conclusion of the argument (Vranas (2012): 10).	A <i>pure imperative argument</i> is valid iff every fact that supports the premise of the argument also supports the conclusion of the argument (Vranas (2011): 376-7, Vranas (2012): 10).
<b>Mixed-premise arguments</b>	A <i>mixed-premise declarative argument</i> is valid iff every fact that both supports the imperative premise and guarantees the declarative premise of the argument guarantees the (declarative) conclusion of the argument (Vranas (2012): 22).	A <i>mixed-premise imperative argument</i> is valid iff every fact that both supports the imperative premise and guarantees the declarative premise of the argument supports the (imperative) conclusion of the argument (Vranas (2012): 26).
<b>Cross-species arguments</b>	A <i>cross-species declarative argument</i> is valid iff every fact that supports the (imperative) premise of the argument guarantees the (declarative) conclusion of the argument (Vranas (2012): 15).	A <i>cross-species imperative argument</i> is valid iff every fact that guarantees the (declarative) premise of the argument supports the (imperative) conclusion of the argument (Vranas (2012): 10).

Vranas also distinguishes between *strong* semantic validity and *weak* semantic validity (for pure imperative arguments), which goes back to his distinction between strong and weak support. His definitions of these are as follows:

A pure imperative argument is (1) *strongly semantically valid* exactly if, for every interpretation  $m$ , every declarative sentence that *strongly* supports on  $m$  every conjunction of all premises of the argument also *strongly* supports on  $m$  the conclusion of the argument, and is (2) *weakly semantically valid* exactly if, for every interpretation  $m$ , every declarative sentence that *weakly* supports on  $m$  every conjunction of all premises of the argument also *weakly* supports on  $m$  the conclusion of the argument (Vranas (2013): 6-7).

## 9.4 Proof Theory

In *New Foundations for Imperative Logic IV*, Vranas (2013) sets out to provide a system of Natural Deduction that is sound and complete with respect to his general definition of argument validity. He lists all the rules of (declarative) natural deduction, then provides the imperative counterparts for each rule (where there are such), as well as new ones for imperatives and for how imperatives and declaratives interact. Excluding the quantified extension and the modal extension, there are 117 rules in total. However, a large proportion of these are redundant, so Vranas gives a shorter (complete) subset of rules in a footnote to an appendix (Vranas (2013): 37, footnote 33). Here is that *subset* of rules (where  $p, q$  stand for declarative sentences and  $i, j, !\phi$  stand for imperative sentences, and where  $\Leftrightarrow$  means “can be replaced with”):

### 9.4.1 Replacement Rules

Name of rule	Declarative logical equivalences	Imperative logical equivalences
Double Negation	$\neg\neg p \Leftrightarrow p$	
Idempotence	$p \& p \Leftrightarrow p$ $p \vee p \Leftrightarrow p$	
Commutativity	$p \& q \Leftrightarrow q \& p$ $p \vee q \Leftrightarrow q \vee p$ $p \leftrightarrow q \Leftrightarrow q \leftrightarrow p$	$i \& j \Leftrightarrow j \& i$ $i \vee j \Leftrightarrow j \vee i$ $p \leftrightarrow i \Leftrightarrow i \leftrightarrow p$
Associativity	$p \& (q \& r) \Leftrightarrow (p \& q) \& r$ $p \vee (q \vee r) \Leftrightarrow (p \vee q) \vee r$ $p \leftrightarrow (q \leftrightarrow r) \Leftrightarrow (p \leftrightarrow q) \leftrightarrow r$	
Distributivity	$p \vee (q \& q') \Leftrightarrow (p \vee q) \& (p \vee q')$ $p \& (q \vee q') \Leftrightarrow (p \& q) \vee (p \& q')$ $p \rightarrow (q \& q') \Leftrightarrow (p \rightarrow q) \& (p \rightarrow q')$ $p \rightarrow (q \vee q') \Leftrightarrow (p \rightarrow q) \vee (p \rightarrow q')$ $(p \vee p') \rightarrow q \Leftrightarrow (p \rightarrow q) \& (p' \rightarrow q)$ $(p \& p') \rightarrow q \Leftrightarrow (p \rightarrow q) \vee (p' \rightarrow q)$	
Transposition		$p \rightarrow i \Leftrightarrow \neg i \rightarrow \neg p$ $i \rightarrow p \Leftrightarrow \neg p \rightarrow \neg i$

Negated Conditional		$\neg(p \rightarrow i) \Leftrightarrow p \rightarrow \neg i$ $\neg(i \rightarrow p) \Leftrightarrow \neg i \rightarrow p$
Material Equivalence		$p \leftrightarrow i \Leftrightarrow (p \leftrightarrow i) \& (i \leftrightarrow p)$ $p \leftrightarrow i \Leftrightarrow (p \leftrightarrow i) (i \leftrightarrow p)$
De Morgan's Laws	$\neg(p \& q) \Leftrightarrow \neg p \vee \neg q$ $\neg(p \vee q) \Leftrightarrow \neg p \& \neg q$	
Exportation		$p \rightarrow (q \rightarrow i) \Leftrightarrow (p \& q) \rightarrow i$
Absorption		$p \rightarrow !q \Leftrightarrow p \rightarrow !(p \& q)$
Tautologous Conjunction	$(p \vee \neg p) \& q \Leftrightarrow q$	
Tautologous Antecedent		$(p \vee \neg p) \rightarrow i \Leftrightarrow i$
Unconditional Negation		$\neg!p \Leftrightarrow !\neg p$
Imperative Conjunction		$(p \rightarrow !q) \& (p' \rightarrow !q') \Leftrightarrow$ $(p \vee p') \rightarrow !((p \rightarrow q) \& (p' \rightarrow q'))$
Imperative Disjunction		$(p \rightarrow !q) \vee (p' \rightarrow !q') \Leftrightarrow$ $(p \vee p') \rightarrow !((p \& q) \vee (p' \& q'))$

### 9.4.2 (Pure Imperative) Inference Rules:

Name of rule	Rule
Strengthening the Antecedent (SA)	$\begin{array}{l} i \\ \text{---} \\ p \rightarrow i \end{array} \quad \begin{array}{l} \text{If } p' \text{ entails } p: \\ p \rightarrow i \\ \text{---} \\ p' \rightarrow i \end{array}$
Weakening the Consequent (WC)	$\begin{array}{l} \text{If } q \text{ entails } q': \\ !q \\ \text{---} \\ !q' \end{array} \quad \begin{array}{l} \text{If } q \text{ entails } q': \\ p \rightarrow !q \\ \text{---} \\ p \rightarrow !q' \end{array}$
Ex Contradictione Quodlibet (ECQ)	$\begin{array}{l} !(p \& \neg p) \\ \text{---} \\ i \end{array}$

Strong and weak semantic entailment have (slightly) different sets of syntactic rules. For strong semantic entailment, add SA and ECQ to the Replacement Rules. To get weak semantic entailment, add SA and WC to the Replacement Rules.

## 9.5 Problems

First, and this is not so much an objection as a complaint, both Vranas's semantics and his proof theory are extremely complicated. Although it is significantly shorter than the full list, Vranas's complete subset of syntactic rules is very long. There are 36 rules for strong semantic entailment, or 37 for weak semantic entailment. Similarly, for his semantics, Vranas introduces a lot of new terminology and several concepts and distinctions, many of which are philosophically problematic, as they rely on dubious and murky theories of truthmakers and of reasons for acting.

### 9.5.1 Truthmaker Theory

A fact *sustains* a proposition when the fact guarantees the proposition. A fact *guarantees* a proposition when, necessarily, the proposition is true if the fact exists. First, it is far from straightforward what it means for a "fact" to "exist." Vranas does not even give an account of what a fact *is*. Facts can't be sentences, as it's too easy for sentences to exist. False sentences exist, so if facts are sentences then false facts exist. Facts, then, must be something else. Perhaps they are sections of the world, or states of affairs. Plausibly, Vranas's theory can work if we take facts to be something like this. However, it is worth noting that it does require a rather substantial theory of facts. Also, there are well-established problems with truthmaker theory. For example, does a fact exist that makes it true that "no unicorns exist"? Perhaps, if we mean "no unicorns exist in the universe," then perhaps the universe in its entirety constitutes the fact that makes this true (because all objects in the universe are non-unicorns, and so the way the universe in its entirety is incompatible with the existence of a unicorn). However, this will not work if we mean "no unicorns exist *at all*" (see Parsons (2006): 594). That is, it will not work if we recognise the possibility that objects exist outside of our universe (perhaps in a co-existing, but causally distinct, universe). Further, it will not help to specify that this universe is all there is, because there is no fact that can make "there is only one universe," or "there is nothing other than this universe" true, either. So, Vranas's theory is based on an already problematic theory of truthmakers.

### 9.5.2 Reasons for Acting

What does it mean for a fact to sustain a prescription? A fact sustains a prescription when it supports it. A fact supports a prescription when it favours the satisfaction proposition over the violation proposition of the prescription expressed by the imperative. This is problematic because it is not clear what counts as “favouring.” Does my preference for Sauvignon Blanc over Merlot count as a reason for obeying “order a bottle of Sauvignon Blanc!”? That is, does it favour the proposition that I order a bottle of Sauvignon Blanc over the proposition that I do not order a bottle of Sauvignon Blanc? Does it, similarly, count as a reason for obeying “drink Sauvignon Blanc with your steak dinner!”? Does the fact that someone has commanded something count as a reason (in itself) to obey that command? The notion of reasons for obeying commands is very murky, and consequently so is the notion of facts supporting prescriptions.

There is a whole related literature on reasons for action (that is, reasons for doing things), into which I will not delve too deeply here. In short, if we accept a Humean conception of practical reason, then a fact only constitutes a reason for somebody to do something given a particular set of desires. For example, the fact that an oncoming bus will kill me constitutes a reason for me to step into its path if I am suicidal and to jump out of its way if I am not. We do sometimes employ non-Humean norms (norms that do not rely on certain desires to count as reasons for acting), but we often employ far too many and we employ them far too often, without being sufficiently discerning about which are Humean and which are not (Pigden (2010): 16-18). A given fact can constitute a reason for somebody to do one thing given one norm of rationality and another thing given another norm. This would not be a problem if Vranas could argue that one of these norms is correct and the others mistaken. Then, he could say that a fact “is a reason for” an action if a person who was aware of that fact *and conformed to the correct norm of practical rationality* would be motivated to do the action. However, the very principle that Vranas presupposes, that to every truth there is a truthmaker, seems to preclude this. There does not seem to be any fact that can make it the case that one norm of practical rationality is correct and the others are not.

Vranas also assumes that there are all-things-considered reasons for action, that is, facts that rationally favour an action when everything has been taken into account. This all-things-considered locution is very popular in the reasons-for-action industry but it is not clear that it corresponds to anything tangible. Given a fact and a norm of practical rationality such that given the norm, the fact favours some action on the part of somebody or other, it may be possible to adduce further facts such that the action in question is no longer favoured.

To suppose that there are facts which constitute all-things-considered reasons for action we must suppose that there is some plausible norm of rationality, some fact and some action such that given the norm the action would be the most rational thing to do and such that there is no further fact that would render that action irrational (or less rational) according to that norm. Characterized in this way, all-things-considered reasons for action are perhaps possible, except that we could not in general be sure that we had considered all things. If the facts are supposed to include our epistemic situations, then this suggests that there are no all-things-considered reasons for action, and thus no conclusive reasons for obeying imperatives. Vranas's theory will only work if these various problems with practical rationality can be solved.

### 9.5.3 Nonmonotonicity

Further, Vranas's account, if based on the relation of favouring, will not support a monotonic logic. This is because a fact can favour an action when taken in isolation but cease to do so when combined with other facts. So, for example, the fact that it is a lovely day might favour my going for a walk. That is, there is a plausible norm of rationality such that the fact that it is a lovely day constitutes a reason for you to go for a walk – it favours the proposition that you go for a walk over the proposition that you do not go for a walk. However, if we add the fact that there are assassins waiting outside to get me, then it is no longer so obvious that the fact that it is a lovely day constitutes a reason for me to go for a walk, as it is overridden by other facts. Worse, if we imagine a capricious God who rewards people for walking on nasty days and punishes them for walking on lovely days, then the loveliness of the day no longer constitutes a reason for walking *at all*. If this is correct, then we can have an intuitively valid imperative argument in which there is no reason to obey the imperative in the conclusion. Consider the following argument:

- A1** If it is a lovely day, go for a walk!
- A2** It is a lovely day.
- ∴ **A3** Go for a walk!

Argument A is intuitively valid. It is yet another instance of modus ponens. However, suppose we add a further premise, that God punishes people who go for walks on lovely days:

- B1** If it is a lovely day, go for a walk!
- B2** It is a lovely day.
- B3** God punishes with eternal torture those who go for walks on lovely days.
- $\therefore$  **B4** Go for a walk!

Argument B should still be valid, if imperative consequence is monotonic. However, on Vranas’s view it is either invalid or trivially valid. If B3 overrules any possibility of there being any reason to go for a walk on a lovely day, then argument B is invalid, making Vranas’s logic nonmonotonic. Alternatively, argument B is trivially valid, because B1 and B3 are inherently contradictory. In this case, Vranas’s theory is no longer nonmonotonic, but argument B being merely trivially valid is no better, as this would make the following argument also (trivially) valid:

- C1** If it is a lovely day, go for a walk!
- C2** It is a lovely day.
- C3** God punishes with eternal torture those who go for walks on lovely days.
- $\therefore$  **C4** Dance a jig!

Argument C is intuitively invalid. So, either Vranas’s theory is nonmonotonic, or argument C serves as a clear counterexample to it.

### 9.5.4 Validity as Disjunctive

Aside from having an unworkably complicated syntax and some confused semantic notions, Vranas’s theory suffers from (roughly) the same theoretical problem as the Smart/Pigden view and Rescher’s theory: it gives a merely disjunctive definition of validity.

In Vranas’s case, there are *six* different definitions of validity. They are all given in terms of different configurations of “supporting” and “guaranteeing.” Vranas attempts to unify them in terms of “sustaining.” However, “sustaining” is, itself, a merely disjunctive concept. What it is for a fact to sustain a proposition is quite different from what it is for a fact to sustain a prescription. “To sustain *X*” is “to guarantee *X* (if *X* is a proposition) *or* to support *X* (if *X* is a prescription).” Supporting is quite a different thing from guaranteeing. A fact guarantees a proposition when if the fact exists, the proposition is true. On the other hand, a fact supports a prescription when it *favours* the satisfaction proposition of the prescription over the violation proposition of the prescription.

Supporting and guaranteeing do not seem to have anything substantively in common.

Vranas unifies them by calling them both “sustaining.” But this term is defined merely disjunctively. So, the “general” definition of validity is just a trick: it is not general at all, because the concept of sustaining is not a general concept. In particular, supporting and guaranteeing are very disanalogous in that the relation of guaranteeing is a *necessary* relation, whereas the relation of supporting is not.

A fact guarantees a proposition whenever “necessarily, the proposition is true if the fact exists.” Note why it is important that it is a *necessitated* conditional. Suppose that a fact guarantees a proposition whenever “the proposition is true if the fact exists.” Conditionals are problematic. We could, for instance, take an arbitrary true proposition and an arbitrary (and unconnected) fact, and this would make a true conditional and thus the fact would guarantee the proposition. For example, take the proposition that the sky is blue, and the fact that John Key is the Prime Minister of New Zealand. John Key *is* the Prime Minister of New Zealand, so this fact exists. The proposition “the sky is blue” is true. So, the conditional ““the sky is blue” is true if John Key is the Prime Minister of New Zealand” is true. So, the fact that John Key is the Prime Minister of New Zealand would *guarantee* that the sky is blue. Similarly, any fact that does *not* exist would guarantee every proposition. For example, it is *not* a fact that Helen Clark is the Prime Minister of New Zealand. So, this fact (that does not exist) would guarantee any arbitrary proposition *p*, because the conditional “*p* is true if the fact that Helen Clark is the Prime Minister exists” is true. It is for these reasons that Vranas’s definition of guaranteeing must be a *necessitated* conditional. On the other hand, a fact supports a prescription whenever the fact “favours the satisfaction proposition over the violation proposition of the prescription.” Even if we assume that we have a satisfactory account of favouring, this is not a necessary relation, and so is essentially disanalogous to the relation of guaranteeing. Calling both guaranteeing and supporting “sustaining” is just to paper over the gulf between them – they are two very different things, and so Vranas’s account of imperative consequence is essentially disjunctive.

### 9.5.5 Counterexamples

Even if the theoretical problems do not concern you, or you think they can be worked out, Vranas’s theory also suffers from a series of counterexamples. First, take the following argument:

- D1** You have a conclusive reason to dance.
- ∴ **D2** Dance!

Vranas's theory makes this argument come out as valid, because any fact that guarantees D1 also supports D2. That is, any fact that makes D1 true is also a reason that favours the proposition that you dance over the proposition that you do not dance. However, this is intuitively incorrect. Argument D is intuitively invalid. The premise is a statement about your having a conclusive reason to do something (dance), but the conclusion is simply about dancing. Having a reason to do something (even if it is a conclusive reason) does not *entail* an imperative to the effect that you do that thing. It gives you a reason to obey the imperative if it were issued to you, or a reason to act (to do the thing) even if there is no command issued to you. It also, perhaps, makes it the case that if someone were to issue that command to you, they would be justified in doing so, but it does not in itself *imply* that such an imperative is in force or applies to you, or exists. Now consider the following two arguments:

- E1** Dance!
- ∴ **E2** There is a reason for you to dance.
- F1** Dance!
- ∴ **F2** A fact exists that favours the proposition that you dance over the proposition that you do not dance.

Of course, arguments E and F are really the same argument, as F2 just spells out the definition of "reason." Again, these arguments come out as valid on Vranas's definition. Every fact that supports the premise also guarantees the truth of the conclusion. That is, any fact that is a reason for you to dance (supports E1/F1) also, obviously, makes E2 true. Any fact that favours the proposition that you dance over the proposition that you do not dance (supports E1/F1) also, obviously, guarantees the truth of F2. However, these arguments also seem intuitively invalid. An imperative *in itself* does not entail that there is any reason for you to obey it. Even if an imperative has been issued, or uttered, or otherwise exists, this would not entail that there is any reason for you to obey it. An imperative *certainly* does not entail that any fact exists, in particular one that favours the proposition to the effect that you obey it over the proposition to the effect that you disobey it. All of these counterexamples point to the underlying theoretical problem with Vranas's theory - that he takes the analogue of truth conditions to be *reasons for obeying*.

So, Vranas's theory of imperative logic as based (on reasons for obeying) fails, along with the theories of Hare, Hofstadter and McKinsey, Smart and Pigden, and Rescher. We will, then, move onto imperative logics that are based on *modal* logic.

## Part V

# Modal Imperative Logics

# Chapter 10

## Validity in Terms of Imperative Obligation

In his 1971 paper, “Imperatives,” Brian Chellas presents a system of imperative logic that is based on the notion of *imperative obligation*. He makes use of the accessibility semantics of modal logic to define a relation of “imperative alternativeness” that holds between worlds, and then defines a notion of imperative obligation as being represented by the ideal alternatives.

### 10.1 Terminology

Chellas begins outlining his theory of imperative logic with a discussion of what we can take the values of imperatives to be. He explains that he thinks it is not necessary to decide this matter, as it is also not essential that the values of indicative sentences are truth-values. We divide indicative sentences up into true ones and false ones, but this could just as easily be any two-valued function, as long as it agrees with truth in its partition. Additionally, he sees no reason to suppose that imperatives are any more than two-valued, so we might as well keep it simple and opt for two values for imperatives. So, he proposes to simply introduce a two-valued function on the set of imperative sentences, without going into an analysis of what, if anything, it corresponds to in reality. He calls them, without committing to any deep meaning, “to hold” and “to fail” (Chellas (1971): 117).

Next, we would need to know how to interpret imperatives. Chellas introduces his notion of “contexts of utterance” of imperatives. He thinks that, under the right conditions, an imperative expresses an obligation. That is, the imperative “get Ned drunk and get the key off him!” conveys an obligation to get Ned drunk and get the key off him. However,

of course, not every instance of an utterance of “get Ned drunk and get the key off him!” conveys an obligation. The circumstances must be right. For example, a king can convey an obligation to his subjects, but perhaps not the other way around, because the king is in the relevant position of authority over them, but they are not in the relevant position of authority over him. So, Chellas believes, the context of utterance is essential in analysing imperatives. These contexts of utterance will be complicated and entirely unique to each situation, except in the rare case where orders – and a hierarchical system – are regulated completely, such as in the military. It is precisely because such well-established, regulated, situations are so rare that Chellas believes obedience to be a poor analogue of truth-values.

Chellas’ notion of obligation – “imperative obligation,” as he calls it, is not to be confused or identified with moral obligation. The concepts are (probably) related in complex ways, but certainly not identical. Moral obligations, if they exist, can (and mostly do) exist independently of any specific imperative being uttered by anyone. Imperative obligations, on the other hand, come into being when an imperative is uttered. They may, or may not, also coincide with a moral obligation, but what is important is that it is a distinct concept.

We must distinguish the notion of obligation from the dual notion of permission, such as “you may go now, Ned”, in cases where, for example, the commander does not intend for it to be mandatory for Ned to go. Take, for example, the “wait” and “walk” signs at a traffic intersection. The first expresses an obligation – it is obligatory to wait – whereas the second expresses a permission – it is permitted (but not obligatory) to walk (Chellas (1971): 118).

Chellas defines a “context of utterance” for any sentence as the state of the world at the time the sentence is uttered. Fully explicated, it will include a speaker, a hearer, a time and a place, at the least (Chellas (1971): 118). For simplicity, we can specify it entirely in terms of a world and a time. He then introduces the notion of “imperatively ideal” worlds. These are the worlds that differ from the actual world only in the respect that what is imperatively obligated in our world is true in them. So, to say that an imperative,  $!x$ , holds in  $W$  is to say that  $x$  is true in all of  $W$ ’s imperatively ideal worlds.

Imperatives, then, are evaluated in terms of their contexts of utterance, as well as a relation of “imperative alternativeness.” Chellas outlines two constraints on what it takes for one context of utterance to be an imperative alternative to another (Chellas (1971): 119-20). First, imperative alternatives to a world,  $W$ , must have the same past history as that world. That is, they can differ only in respects possible from the standpoint of the laws of nature. This formalises the idea that imperative obligations come into existence at a world as a function of events that occur at that world. Thus, the history of that world is of utmost

importance.

Second, if two worlds have their past histories in common then they necessarily must also have their imperative alternatives in common. This formalises the idea that imperative obligations come into existence at a world solely as a function of events that occur at that world.

These constraints, together with the assumption that at any given time every world has at least one imperative alternative, form the basis of Chellas' formal logic of imperatives.

## 10.2 Chellas' Formal System

The formal logic, language L, is as follows:

L:	P1, P2, P3...	atomic sentences
	$\neg A$	negation "not A"
	$A \rightarrow B$	conditional "if A then B"
	$!A$	imperative "let A be the case"
	$\text{!}A$	permission "it is permitted that A"

In addition to these 4 basic operators/connectives, Chellas also distinguishes the following modal operators:

$\Box A$	natural necessity "necessarily, A"
$\Diamond A$	natural possibility "possibly, A"
$[P]A$	"it always was the case that A"
$(-)A$	"at the moment just past it was the case that A"
$\langle P \rangle A$	"it (at least once) was the case that A"

Next, Chellas outlines the models of L. He introduces a (non-empty) set W of worlds, and a set T of times, then represents a context of utterance as a pair  $\langle w, t \rangle$ , where  $w \in W$  and  $t \in T$ .

T is discrete and open-ended;

T:  $\{\dots, -1, 0, +1, \dots\}$

Each world is a function from the set T of times into a (non-empty) set X of states of affairs;

$w: T \rightarrow X$

That is, each world is a series of (discrete) times (or moments), at each of which a set of propositions is true.

Additionally,  $R_t(w, w')$  means that the world  $w'$  is an imperative alternative to the world  $w$  at time  $t$ , and  $S_t(w, w')$  means that the worlds  $w$  and  $w'$  have the same past history at time  $t$ .

Chellas then states the three conditions on the relation of imperative alternativeness precisely (Chellas (1971): 122):

For each  $w, w', w'' \in W$ , and  $t \in T$ :

- I** there is a  $w' \in W$  such that  $R_t(w, w')$
- II** if  $R_t(w, w')$ , then  $S_t(w, w')$
- III** if  $S_t(w, w')$ , then  $R_t(w, w'')$  iff  $R_t(w', w'')$

In English, these three conditions are:

- I** Every world has at least one imperative alternative.
- II** If one world is an imperative alternative to another, then they share the same past.
- III** If two worlds share the same past history, then any further world that is an imperative alternative to one is also an imperative alternative to the other.

$W$  (worlds),  $T$  (times), and  $R$  (the relation of imperative alternativeness), along with  $\| \cdot \|$ , an interpretation which evaluates the sentences of  $L$  with respect to contexts of utterance  $\langle w, t \rangle \in W \times T$ . The interpretation is two-valued: 1 (designated) and 0. In a model for  $L$ ,  $\| \cdot \|$  is defined to meet the following conditions:

For each  $w \in W$ , and  $t \in T$ :

1.  $\|P_i\|(w, t) \in \{1, 0\}$ ,  $i = 1, 2, 3, \dots$ ;
2.  $\|\neg A\|(w, t) = 1$  iff  $\|A\|(w, t) = 0$ ;
3.  $\|A \rightarrow B\|(w, t) = 1$  iff  $\|A\|(w, t) \leq \|B\|(w, t)$ ;
4.  $\|\!|A\|\!(w, t) = 1$  iff  $\|A\|(w', t) = 1$ , for every  $w' \in W$  such that  $R_t(w, w')$ ;
5.  $\|_i A\|(w, t) = 1$  iff  $\|A\|(w', t) = 1$ , for some  $w' \in W$  such that  $R_t(w, w')$ ;

6.  $\|\Box A\|(w,t) = 1$  iff  $\|A\|(w',t) = 1$ , for every  $w' \in W$  such that  $S_t(w,w')$ ;
7.  $\|\Diamond A\|(w,t) = 1$  iff  $\|A\|(w',t) = 1$ , for some  $w' \in W$  such that  $S_t(w,w')$ ;
8.  $\|[P]A\|(w,t) = 1$  iff  $\|A\|(w,t') = 1$ , for every  $t' \in T$  such that  $t' < t$ ;
9.  $\|(-)A\|(w,t) = 1$  iff  $\|A\|(w,t-1) = 1$ ;
10.  $\|\langle P \rangle A\|(w,t) = 1$  iff  $\|A\|(w,t') = 1$ , for some  $t' \in T$  such that  $t' < t$ .

Translated (loosely) into English, these conditions are:

1. Every atomic sentence, indexed to a world and a time, has the value of 1 or 0;
2.  $\neg A$  iff not  $A$ ;
3.  $A \rightarrow B$  iff  $A$  is 0 or  $B$  is 1;
4.  $!A$  iff at all imperatively alternative worlds  $A = 1$ ;
5.  $¡A$  iff at some imperatively alternative world  $A = 1$ ;
6.  $\Box A$  ( $A$  is historically necessary) iff  $A = 1$  in every world with the same past history;
7.  $\Diamond A$  ( $A$  is historically possible) iff  $A = 1$  in some world with the same past history;
8.  $[P]A$  ( $A$  has always been 1) iff  $A = 1$  at every past time;
9.  $(-)A$  ( $A$  was 1 just now) iff  $A = 1$  at  $t-1$ ;
10.  $\langle P \rangle A$  ( $A$  was once 1) iff  $A = 1$  at some past time.

In addition,  $\|\ \|\$  is subject to the further assumption:

For each  $w, w' \in W$ ,  $t \in T$ , and atomic sentence  $P_t$  ( $t = 1, 2, 3, \dots$ ),  
if  $w(t) = w'(t)$ , then  $\|P_t\|(w,t) = \|P_t\|(w',t)$ .

That is, if two worlds are the same world, then every sentence has the same value at both of those worlds. In other words, the value of every atomic sentence is a function solely of the state of affairs, determined by the context of utterance.

So these,  $\langle W, T, R, \|\ \|\rangle$ , form a model of  $L$ .  $\models A$  means that  $A$  is valid – that is, it receives the value 1 at every context of utterance in every model  $\langle W, T, R, \|\ \|\rangle$  of  $L$ .

Logical consequence can be defined as follows:

An argument is valid if and only if, for every model, if the premises all receive the value 1 then so does the conclusion.

### 10.2.1 Some Results

Chellas then outlines some results of this account of imperatives (Chellas (1971): 123-5):

- (1) If  $A$  is a (classical propositional) tautology, then  $\models A$  (all (classical propositional) tautologies are valid).
- (2) If  $\models A$ , then  $\models !A$  (the imperative closure of a valid sentence is valid).
- (3)  $\models !(A \rightarrow B) \rightarrow (!A \rightarrow !B)$  (the set of (proper) imperatives is closed under modus ponens)
- (4) If  $\models (A_1 \wedge \dots \wedge A_n) \rightarrow B$ , then  $\models (!A_1 \wedge \dots \wedge !A_n) \rightarrow !B$  (the imperative mood is closed under all valid forms of argument)
- (5)  $\models !A \leftrightarrow \neg_j \neg A$   
and
- (6)  $\models \neg_j A \leftrightarrow \neg !\neg A$  (! and  $\neg_j$  are interdefinable)

(7) - (10) specify the ways in which the imperative operators distribute with respect to conjunction and disjunction:

- (7)  $\models !(A \wedge B) \leftrightarrow (!A \wedge !B)$
- (8)  $\models \neg_j (A \vee B) \leftrightarrow (\neg_j A \vee \neg_j B)$
- (9)  $\models !(A \vee !B) \rightarrow !(A \vee B)$
- (10)  $\models \neg_j (A \wedge B) \rightarrow (\neg_j A \wedge \neg_j B)$

So far, these results, (1) - (10), hold even if the relation of imperative alternativeness satisfies no special conditions. But because this relation satisfies (I) (all worlds access some world), imperative obligations entails imperative permission:

- (11)  $\models !A \rightarrow \neg_j A$

In other words, contrary imperatives  $!A$  and  $!\neg A$  are contradictory. However, while they cannot both be true, they can both be false (i.e. while  $!A$  implies  $\neg!\neg A$ ,  $\neg!A$  does *not* imply  $!\neg A$ ).

This also means that (11) does *not* imply that  $!\neg A$  and  $\neg!A$  are equivalent.

Further, because the relation of imperative alternativeness satisfies (II) and (III) (which together make the relation transitive), any iteration of the imperative operators is vacuous:

$$(12) \models!A \leftrightarrow !!A$$

$$(13) \models!A \leftrightarrow !iA$$

$$(14) \models_iA \leftrightarrow !iA$$

$$(15) \models_iA \leftrightarrow !_iA$$

### 10.2.2 Similar to KD45

Chellas' formal system of imperatives is, it turns out, equivalent to the normal modal logic KD45, with an extra requirement that all worlds must have the same past history as their ideal worlds. Specifically, all Chellas-models are KD45-models, but not vice versa. On the other hand, all KD45-validities are Chellas-validities, but not vice versa. This makes KD45 a *sublogic* of Chellas-logic.

Chellas' constraints on the (accessibility) relation of imperative alternativeness again:

For each  $w, w', w'' \in W$ , and  $t \in T$ :

**I** there is a  $w' \in W$  such that  $R_t(w, w')$

**II** if  $R_t(w, w')$ , then  $S_t(w, w')$

**III** if  $S_t(w, w')$ , then  $R_t(w, w'')$  iff  $R_t(w', w'')$

Constraint **II** is the extra constraint that worlds and their ideal worlds must have the same past history. However, we can combine **II** and **III**, to give:

**II-III** if  $R_t(w, w')$ , then  $R_t(w, w'')$  iff  $R_t(w', w'')$

KD45 is the normal modal logic K, with the additional axioms D, 4, and 5 (KD45 will

be outlined in full in section 12.2):

**D**  $!\neg\phi \rightarrow \neg!\phi$

**4**  $!\phi \rightarrow !!\phi$

**5**  $!_i\phi \rightarrow !_i!\phi$

These axioms correspond to an accessibility relation that is extensible, transitive, and euclidean:

$\eta$  for all  $w$ , there is a  $w'$  such that  $Rww'$  (R is *extensible*)

$\tau$  if  $Rww'$  and  $Rw'w''$ , then  $Rww''$  (R is *transitive*)

$\epsilon$  if  $Rww'$  and  $Rww''$ , then  $Rw'w''$  (R is *euclidean*)

The second and third constraints can be rewritten less elegantly as follows:

$\tau$  if  $Rww'$ , then if  $Rw'w''$ , then  $Rww''$

$\epsilon$  if  $Rww'$ , then if  $Rww''$ , then  $Rw'w''$

We can then combine  $\tau$  and  $\epsilon$ , to give:

$\tau\text{-}\epsilon$  if  $Rww'$ , then  $Rww''$  iff  $Rw'w''$

This constraint,  $\tau\text{-}\epsilon$ , is identical to Chellas' constraint, **II-III**. So, Chellas' system is essentially KD45, with the additional requirements regarding past histories.

### 10.3 Chellas' Theory on the Success Criteria

The criteria for a successful imperative logic, recall, are that it must be general, conservative, adequate, and non-cognitivist about imperatives.

Chellas' theory is general. His definition of validity is that an argument is valid whenever the conclusion receives the value 1 in every model in which the premises all also receive the value 1. Because the (same) values 1 and 0 are given to imperatives and declaratives alike, this definition of validity applies in the same way to pure declarative arguments, pure imperative arguments, and arguments with a combination of declaratives and imperatives.

The theory is also conservative, because the values 1 and 0, although not meant to be interpreted as truth-values, agree with truth in every case, for declarative sentences. That is, for any declarative sentence, whenever it is true it gets the value 1, and whenever it is false

it gets the value 0. So, every premise and conclusion in any purely declarative argument (that is, any argument with no imperative elements) will get the value 1 iff it is true. Thus, every purely declarative argument will be Chellas-valid exactly when it is classically valid.

Chellas' theory is also adequate, but I will not argue for this here. Instead, in chapter 11, I will outline another theory (that of Parsons), show that it is inadequate, then show that when we fix Parsons' theory we arrive back at KD45.

The problem with this theory is that Chellas does not attempt to explain how imperatives are not truth-apt. So, it fails the criterion of non-cognitivism about imperatives. He dismisses this question entirely, by calling the two (binary) values "to hold" and "to fail," instead of "to be true" and "to be false." However, it really is nothing more than a name change, because he does not give any account of what these properties *are*. In particular, he does not give any account of how, if at all, "holding" and "failing" differ between imperatives and declaratives. That is, even if we do not call them "truth" and "falsity," there are two binary values that sentences can possess and these values seem to be the same, for Chellas, whether the sentences are imperatives or declaratives. For declaratives, usually, we take these binary values to correspond to the natural properties of truth and falsity, respectively. That is, a sentence gets the value T (or 1) if and only if it is true; if it accurately describes the world – if what it says is the case really is the case. A (declarative) sentence gets the value F (or 0) if and only if it is false; if it *inaccurately* describes the world – if what it says is the case is *not* really the case. So, there is a natural interpretation of these values as *truth*-values. Chellas even specifies that 1 and 0, whatever they are, must "agree with truth" in how they partition up the sentences (Chellas (1971): 116). However, he wants to take away this natural interpretation, and imagine the values as interpretation-less 1s and 0s. These binary values can be given to declarative and imperative sentences alike. This is philosophically unsatisfying, for two reasons. First, we do not want these values to be allocated arbitrarily. We want them to *mean something*. They should be allocated to sentences when those sentences really do possess the corresponding properties. Chellas does not attempt to offer an account of what these properties are; what it means for a sentence (either imperative or declarative) to "hold" or to "fail to hold." Secondly, in particular, he does not explain how these properties differ for imperatives and declaratives. If we interpret the values to correspond to the properties "truth" and "falsity," then imperatives as well as declaratives would come out as truth-apt. We have already seen that no version of cognitivism that tries to explain how this is possible is satisfactory. Imperatives do not say that something is the case, they do not describe the world, so they cannot succeed or fail to do so. They may

“hold” and “fail to hold” in another way, but Chellas does not describe how this works, or explain how this is possible, so his theory is not non-cognitivist about imperatives, because it offers no explanation of to what properties the 1s and 0s correspond, so he gives no account of truth or of any imperative analogue.

In chapter 12, I will show that we can provide an account of this for Chellas (at least, for KD45). First, though, as I mentioned, I will outline Josh Parsons’ theory of imperative logic and point out some counterexamples. Then, in chapter 12, I will show that when we fix Parsons’ logic we, again, arrive at KD45.

# Chapter 11

## Functional Preposcription Semantics and KDDc4

Josh Parsons, in his paper “Command and Consequence” (2013), formulates a new account of imperative validity. First, Parsons asks us to consider the following three arguments:

**A1** Attack at dawn if the weather is fine!  
**A2** The weather is fine.  
 $\therefore$  **A3** Attack at dawn!

**B1** Attack at dawn unless the weather is fine!  
**B2** The weather is not fine.  
 $\therefore$  **B3** Attack at dawn!

**C1** Attack at dawn if the weather is fine!  
**C2** Attack at dawn!  
 $\therefore$  **C3** The weather is fine.

Argument A is an instance of modus ponens, argument B is (arguably) an instance of disjunctive syllogism (assuming “unless” is translated as a simple “or,” which is contentious), and argument C is an instance of the fallacy of affirming the consequent. Because imperatives (it is normally thought) are not truth-apt, Parsons argues that we need to find a new criterion of validity.

## 11.1 Content-Validity

To understand Parsons' theory of prescriptions, we must first understand what he calls the *standard view* (Parsons (2013): 63): that imperatives and declaratives both have propositions as their contents, and that it is between these propositions that the relation of logical entailment holds. The standard view is that the content of a declarative sentence is its *truth-conditions*: the set of worlds at which it is true. Similarly, the content of an imperative sentence is its *compliance-conditions*: the set of worlds at which it is complied with. For example, consider the following two sentences:

- (1) You attack at dawn.
- (2) Attack at dawn!

The content of (1), being in the declarative mood, is the proposition that you attack at dawn; that is, the set of worlds at which you attack at dawn. The content of (2), being in the imperative mood, is the proposition that describes the world when you have complied with it; that is, the set of worlds at which you attack at dawn. Thus, we can see that (1) and (2) have the same content.

The difference between them is of course still apparent, and must be explained in some other way. Examples of attempts to do this include, for example, those of R. M. Hare (1952) (see chapter 7) and H. S. Leonard (1959) (see chapter 5). Hare says that they have the same *phrastic* part but different *neustic* parts, the neustic part being the way in which the speaker “nods assent” to the phrastic. Leonard says that they have the same *ultimate topics of concern* but different *ultimate concerns*, where the ultimate concern of one is that the addressee *believe* the topic of concern, and the ultimate concern of the other is that that addressee *bring about* the topic of concern.

The idea behind content-validity is that it is these contents of sentences that are the relata of the relation of logical entailment. So, for example, we could restate the traditional (Tarskian, see section 1.3.2) criterion of validity in terms of propositions instead of truth (assuming that propositions are sets of possible worlds):

A collection of propositions  $p_1 \dots p_n$  *entail* a proposition  $q$  iff there is no possible world that is a member of each of  $p_1 \dots p_n$  and not a member of  $q$  (Parsons (2013): 64).

As contents are propositions, we can thus define content-validity:

An argument is *content-valid* iff the contents of its premises jointly entail the content of its conclusion (Parsons (2013): 65).

However, if the content of *both* imperative sentences and declarative sentences are thought to be *propositions*, and hence sets of possible worlds, then we run into what Parsons calls the “fallacies of mood” (Parsons (2013): 66) which are content-valid arguments that are intuitively invalid because they do not preserve the mood of the sentences accurately. Consider the following examples:

**D1** You attack at dawn.  
∴ **D2** Attack at dawn!

**E1** Attack at dawn if the weather is fine!  
**E2** Let the weather be fine!  
∴ **E3** You attack at dawn.

Both arguments fail the rough-and-ready tests that I outlined in section 2.3.3. In the case of argument D, it is perfectly consistent to predict that you will attack at dawn but refuse to command that you attack at dawn. So, someone could sincerely utter D1 and sincerely refuse to utter D2 without making a logical mistake. In the case of argument E, similarly, it is possible to consistently (that is, without logical error) sincerely utter E1 and E2 and yet sincerely refuse to utter E3. That is, someone could command you to attack at dawn if the weather is fine, and also command that the weather be fine, but simultaneously believe that the weather will not be fine, or that you will disobey my command. This is a perfectly consistent set of attitudes, and thus she could refuse to assert that you will attack at dawn, as she does not believe that you will. Again, there is no logical mistake.

## 11.2 Preprescription Semantics

To solve the problem of the fallacies of mood for the notion of content-validity, Parsons proposes that the contents of sentences are not propositions, but *preprescriptions* (Parsons (2013): 71). That is, the contents of neither declarative nor imperative sentences are propositions, because they are not fine-grained enough to give us the right results when we consider arguments such as D and E. Propositions do not carry the information that is needed to determine whether a sentence is declarative or imperative. So, Parsons introduces a new,

two-dimensional set-theoretic construction out of possible worlds: preprescriptions. The inspiration for these two-dimensional preprescriptions came from Robert Stalnaker (1999), but they are very different to Stalnaker’s two-dimensional semantics, and the two should not be confused.

A preprescription is a set of *ordered pairs* of possible worlds. Suppose, for illustration, that there are just four possible worlds:

$w_{fa}$  in which the weather is fine and you attack at dawn

$w_{f*}$  in which the weather is fine and you do not attack at dawn

$w_{*a}$  in which the weather is not fine and you attack at dawn

$w_{**}$  in which the weather is not fine and you do not attack at dawn

A preprescription is a set of pairs of worlds, where the first member of the pair has to do with truth-conditions, and the second member has to do with compliance conditions. So how do we know which pairs are in a preprescription and which are not? Instead of listing all possible pairs or worlds, we can represent the full list in a matrix, where each cell of the matrix represents a pair, with the first member being the world that heads the cell’s row, and the second member being the world that heads the cell’s column. Here are the 16 possible ordered pairs:

	$w_{fa}$	$w_{f*}$	$w_{*a}$	$w_{**}$
$w_{fa}$				
$w_{f*}$				
$w_{*a}$				
$w_{**}$				

The worlds listed down the left-hand side of the matrix represent the first members of the pairs, the ones that are to do with truth-conditions. The worlds listed across the top of the matrix represent the second members of the pairs, the ones that are to do with compliance conditions. So, to illustrate in the following matrix, the cell marked with an x represents the ordered pair of worlds ( $w_{fa}, w_{*a}$ ):

	$w_{fa}$	$w_{f*}$	$w_{*a}$	$w_{**}$
$w_{fa}$			×	
$w_{f*}$				
$w_{*a}$				
$w_{**}$				

We can then indicate which pairs of worlds are in the content of a given preprescription by ticking the cells that correspond to the pairs that are in the preprescription.

### 11.2.1 Simple Prescriptions

The following matrix represents the simple (that is, logical connective-less) preprescription “the weather is fine”:

	$w_{fa}$	$w_{f*}$	$w_{*a}$	$w_{**}$
$w_{fa}$	✓	✓	✓	✓
$w_{f*}$	✓	✓	✓	✓
$w_{*a}$				
$w_{**}$				

(3) “The weather is fine.”

All the cells in the rows which are headed by  $w_{fa}$  and  $w_{f*}$  are ticked. This is because (3) is a simple declarative sentence, so all that matters is the truth-conditions (the first member of the pair). So, all the cells which represent a pair where the first member is a world at which “the weather is fine” is true are ticked. That means that all such pairs of worlds are in the content of the preprescription, regardless of what is going on at the second world in the pair. Similarly, we can represent simple imperative sentences. For example, the following matrix represents the preprescription “attack at dawn!”:

	$w_{fa}$	$w_{f*}$	$w_{*a}$	$w_{**}$
$w_{fa}$	✓		✓	
$w_{f*}$	✓		✓	
$w_{*a}$	✓		✓	
$w_{**}$	✓		✓	

(2) “Attack at dawn!”

Compare this to the matrix representing the preprescription “You will attack at dawn”:

	$wfa$	$wf*$	$w*a$	$w**$
$wfa$	✓	✓	✓	✓
$wf*$				
$w*a$	✓	✓	✓	✓
$w**$				

(1) “You attack at dawn.”

In the simple declarative case, what matters is the truth-conditions, so we only look at the first member of each ordered pair. Similarly, in the simple imperative case, what matters is the compliance-conditions, so we only look at the *second* member of each ordered pair, and all the pairs with you-attacking-at-dawn worlds as the second members are in the content of the preproscription, regardless what is happening at their first-member world.

That is, where  $V(\phi)$  is “the content of  $\phi$ ”:

$V(\phi)$ , where  $\phi$  is a simple declarative, is the set of all pairs  $(w,w')$  such that  $\phi$  is true at  $w$ .

$V(\phi)$ , where  $\phi$  is a simple imperative, is the set of all pairs  $(w,w')$  such that  $\phi$  is complied with at  $w'$  (Parsons (2013): 74).

### 11.2.2 Complex Preproscriptions

Complex preproscriptions are those that contain logical connectives. Firstly, there are the straightforward cases. Here are some examples:

	$wfa$	$wf*$	$w*a$	$w**$
$wfa$				
$wf*$				
$w*a$	✓	✓	✓	✓
$w**$	✓	✓	✓	✓

(4) “The weather is not fine.”

Here, we have a declarative negation. So the content of the preproscription  $\neg\phi$  contains all the pairs  $(w,w')$  such that  $\phi$  is not true at  $w$ .

	$w_{fa}$	$w_{f*}$	$w_{*a}$	$w_{**}$
$w_{fa}$	✓			
$w_{f*}$	✓			
$w_{*a}$	✓			
$w_{**}$	✓			

(5) “Let it be that the weather is fine and attack at dawn!”

This is an imperative conjunction. So, the content of the preprescription  $!\phi \wedge !\psi$  contains all the pairs  $(w, w')$  such that both  $!\phi$  and  $!\psi$  are complied with at  $w'$ .

	$w_{fa}$	$w_{f*}$	$w_{*a}$	$w_{**}$
$w_{fa}$	✓	✓	✓	✓
$w_{f*}$	✓	✓	✓	✓
$w_{*a}$	✓	✓	✓	✓
$w_{**}$				

(6) “The weather is fine or you attack at dawn.”

This is a declarative disjunction. The content of the preprescription  $\phi \vee \psi$  contains all the pairs  $(w, w')$  such that either  $\phi$  or  $\psi$  is true at  $w$ .

These examples are cases where two subsentences *of the same kind* are connected with a logical connective. Less straightforward are cases where the connectives connect two subsentences *of different kinds*. For example:

	$w_{fa}$	$w_{f*}$	$w_{*a}$	$w_{**}$
$w_{fa}$	✓		✓	
$w_{f*}$	✓		✓	
$w_{*a}$				
$w_{**}$				

(7) “The weather is fine and attack at dawn!”

In this case, the pairs of worlds  $(w, w')$  in the content of the preprescription are the worlds in which “the weather is fine” is true at  $w$  *and* “attack at dawn!” is complied with at  $w'$ .

	$w_{fa}$	$w_{f*}$	$w_{*a}$	$w_{**}$
$w_{fa}$	✓	✓		
$w_{f*}$				
$w_{*a}$	✓	✓		
$w_{**}$				

(8) “You attack at dawn, and let the weather be fine!”

Similarly, in this case, the pairs of worlds  $(w, w')$  in the content of the preposcription are the worlds in which “you attack at dawn” is true at  $w$  and “let the weather be fine” is complied with at  $w'$ .

### 11.2.3 Conditional Imperatives

Finally, the least straightforward: the conditional command, or equivalently (because we are treating the English conditional as the material conditional), disjunctions where one disjunct is imperative and the other is declarative.

One view of conditional commands is that they are simply a command to make a conditional true. So, for example,

(9) “If our plans are betrayed, flee for your life!”

would, under this view, be equivalent to:

(10) “Let it be that if our plans are betrayed, you flee for your life!”

which is, in turn, equivalent to:

(11) “Let it be that either our plans are not betrayed, or you flee for your life!”

However, this implies that we could comply with the command *either* by fleeing for your life, *or* by ensuring that our plans are not betrayed. This doesn’t seem to capture what is meant by the conditional command “If our plans are betrayed, flee for your life!”. What it means is that, under the condition that our plans are betrayed, you are commanded to flee. Relatedly, this analysis implies that this conditional command is equivalent to the following one:

(12) “Let our plans be betrayed only if you flee for your life!”

This is equivalent, under this analysis, to:

(10) “Let it be that if our plans are betrayed, you flee for your life!”

which is, again, equivalent to:

(11) “Let it be that either our plans are not betrayed, or you flee for your life!”

So, (9) and (12) are equivalent, under this view. However, they are not equivalent, as (9) is telling us to flee for our life (in the case that our plans are betrayed), whereas (12) is telling us to let our plans be betrayed (but only in the case where we flee for our life). Parsons' preposcription semantics respects this non-equivalence. Here are the contents of (9) and (12) (where  $p$  is "our plans are betrayed" and  $f$  is "you flee for your life"):

	$W_{pf}$	$W_{p*}$	$W_{*f}$	$W_{**}$
$W_{pf}$	✓		✓	
$W_{p*}$	✓		✓	
$W_{*f}$	✓	✓	✓	✓
$W_{**}$	✓	✓	✓	✓

(9) "If our plans are betrayed, flee for your life!"

	$W_{pf}$	$W_{p*}$	$W_{*f}$	$W_{**}$
$W_{pf}$	✓	✓	✓	✓
$W_{p*}$			✓	✓
$W_{*f}$	✓	✓	✓	✓
$W_{**}$			✓	✓

(12) "Let our plans be betrayed only if you flee for your life!"

Now, we can formally define the content of various complex preposcriptions (where, again,  $V(\phi)$  means "the content of  $\phi$ "):

$V(\ulcorner \phi \wedge \psi \urcorner)$  is the set of all pairs of worlds  $(w, w')$  such that  $(w, w') \in V(\phi)$  and  $(w, w') \in V(\psi)$ .

$V(\ulcorner \phi \vee \psi \urcorner)$  is the set of all pairs of worlds  $(w, w')$  such that  $(w, w') \in V(\phi)$  or  $(w, w') \in V(\psi)$ .

$V(\ulcorner \neg \phi \urcorner)$  is the set of all pairs of worlds  $(w, w')$  such that  $(w, w') \notin V(\phi)$ .

$V(\ulcorner \phi \rightarrow \psi \urcorner)$  is the set of all pairs of worlds  $(w, w')$  such that  $(w, w') \notin V(\phi)$  or  $(w, w') \in V(\psi)$ .

We can also think of the imperative mood as an operator on preposcriptions. It converts a declarative preposcription,  $\phi$ , into an imperative one,  $!\phi$ , where  $!\phi$  is complied with at all and only those worlds that  $\phi$  is true at.

Recall the matrices for the preposcriptions (1) and (2):

	$w_{fa}$	$w_{f*}$	$w_{*a}$	$w_{**}$
$w_{fa}$	✓	✓	✓	✓
$w_{f*}$				
$w_{*a}$	✓	✓	✓	✓
$w_{**}$				

(1) “You attack at dawn.”

	$w_{fa}$	$w_{f*}$	$w_{*a}$	$w_{**}$
$w_{fa}$	✓		✓	
$w_{f*}$	✓		✓	
$w_{*a}$	✓		✓	
$w_{**}$	✓		✓	

(2) “Attack at dawn!”

Notice that the matrix for (2) is a “flipped” version of the matrix for (1). Parsons gives us a more precise description of how the imperative operator “flips” the matrix:

It takes the value of each cell across the diagonal of the input matrix, and reproduces those cells across each row of the output matrix (Parsons (2013): 77).

In the terminology of two-dimensional modal logic, Parsons’ imperative operator is the “Stalnaker dagger” (Stalnaker (1999): 82) applied to preproscriptions. To get to the matrix for (2), we take the matrix for (1), and we look to all the cells across the diagonal (indicated here by circles in the relevant cells):

	$w_{fa}$	$w_{f*}$	$w_{*a}$	$w_{**}$
$w_{fa}$	⊙	✓	✓	✓
$w_{f*}$		○		
$w_{*a}$	✓	✓	⊙	✓
$w_{**}$				○

(1) “You attack at dawn.”

We see the following pattern: tick, blank, tick, blank. Then, we reproduce the pattern across each of the rows:

	$w_{fa}$	$w_{f*}$	$w_{*a}$	$w_{**}$
$w_{fa}$	✓		✓	
$w_{f*}$	✓		✓	
$w_{*a}$	✓		✓	
$w_{**}$	✓		✓	

(2) "Attack at dawn!"

Each row in (2) has the pattern of tick, blank, tick, blank.

The imperative operator can also be attached to a preprescription that is already an imperative, but that this makes no difference to the content of the preprescription. Take, for example, the preprescription "let our plans be betrayed!":

	$w_{pf}$	$w_{p*}$	$w_{*f}$	$w_{**}$
$w_{pf}$	⊗	✓		
$w_{p*}$	✓	⊗		
$w_{*f}$	✓	✓	○	
$w_{**}$	✓	✓		○

(13) "Let our plans be betrayed!"

The diagonal pattern, here, is: tick, tick, blank, blank. So, we reproduce the 'new' matrix as follows:

	$w_{pf}$	$w_{p*}$	$w_{*f}$	$w_{**}$
$w_{pf}$	✓	✓		
$w_{p*}$	✓	✓		
$w_{*f}$	✓	✓		
$w_{**}$	✓	✓		

(13) "\*Let it be that let our plans be betrayed!"

The resulting matrix is identical, and it always will be in cases like this; that is, in general the matrix for  $!!\phi$  will be identical to that of  $!\phi$ . In general, then, the content of  $!!\phi$  is identical to the content of  $!\phi$ .

Consider a more complex example:

	$w_{fa}$	$w_{f*}$	$w_{*a}$	$w_{**}$
$w_{fa}$	⊗		✓	
$w_{f*}$	✓	○	✓	
$w_{*a}$			○	
$w_{**}$				○

(7) “The weather is fine and attack at dawn!”

The diagonal pattern is: tick, blank, blank, blank. So, the matrix for “\*let it be that: the weather is fine and attack at dawn!” is as follows:

	$w_{fa}$	$w_{f*}$	$w_{*a}$	$w_{**}$
$w_{fa}$	✓			
$w_{f*}$	✓			
$w_{*a}$	✓			
$w_{**}$	✓			

(14) “\*Let it be that: the weather is fine and attack at dawn!”

In this case, the only pairs of worlds in the content/preposcription are the ones that have worlds in which both “let the weather be fine!” and “attack at dawn!” are complied with as their second member.

Similarly, consider the conditional command:

	$w_{pf}$	$w_{p*}$	$w_{*f}$	$w_{**}$
$w_{pf}$	⊗		✓	
$w_{p*}$	✓	○	✓	
$w_{*f}$	✓	✓	⊗	✓
$w_{**}$	✓	✓	✓	⊗

(9) “If our plans are betrayed, flee for your life!”

This becomes:

	$w_{pf}$	$w_{p*}$	$w_{*f}$	$w_{**}$
$w_{pf}$	✓		✓	✓
$w_{p*}$	✓		✓	✓
$w_{*f}$	✓		✓	✓
$w_{**}$	✓		✓	✓

(9) “Let it be that if our plans are betrayed, flee for your life!”

Note that in this case, the content is all the pairs of worlds that have as their second member the worlds at which “let it be that: if our plans are betrayed, you flee for your life!” is complied with.

Note that the cells on the diagonal of the matrices correspond to the pairs of worlds that have the same world as member one and member two. Then, formally, the content of  $!\phi$  is:

$V(!\phi)$  is the set of pairs of worlds  $(w, w')$  such that  $(w', w') \in V(\phi)$ .

### 11.2.4 Entailment and Validity

Given all this, it is possible to define a notion of entailment for preproscriptions. Following Parsons (2013), we define entailment between a set of preproscriptions and a preproscription:

$p_1, \dots, p_n$  jointly *entail*  $q$  iff every member of each of  $p_1, \dots, p_n$  is also a member of  $q$  (i.e. every cell ticked in the matrices for all of  $p_1, \dots, p_n$  is also ticked in the matrix for  $q$  (Parsons (2013): 78).

We can also save the notion of content-validity:

An argument is *valid* iff the contents of its premises jointly entail the content of its conclusion (Parsons (2013): 79).

We may check argument A for validity (the squares indicate the cells that are ticked for all premises, whereas the circles on the previous pages indicated the cells on the diagonal):

	$wfa$	$wf*$	$w*a$	$w**$
$wfa$	☑		☑	
$wf*$	☑		☑	
$w*a$	✓	✓	✓	✓
$w**$	✓	✓	✓	✓

(A1) “Attack at dawn if the weather is fine!”

	$wfa$	$wf*$	$w*a$	$w**$
$wfa$	☑	✓	☑	✓
$wf*$	☑	✓	☑	✓
$w*a$				
$w**$				

(A2) “The weather is fine.”

	$w_{fa}$	$w_{f*}$	$w_{*a}$	$w_{**}$
$w_{fa}$	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>	
$w_{f*}$	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>	
$w_{*a}$	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>	
$w_{**}$	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>	

**(A3)** "Attack at dawn!"

There are only four cells ticked in the matrices for *both* A1 and A2, and all four of the same cells are ticked in the matrix for A3, the conclusion, so the argument is valid.

Similarly, we can check argument E for validity (recall that this was an example of a fallacy of mood):

- E1** Attack at dawn if the weather is fine!
- E2** Let the weather be fine!
- $\therefore$  **E3** You attack at dawn.

Here are the matrices:

	$w_{fa}$	$w_{f*}$	$w_{*a}$	$w_{**}$
$w_{fa}$	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>	
$w_{f*}$	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>	
$w_{*a}$	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
$w_{**}$	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

**(E1)** "Attack at dawn if the weather is fine!"

	$w_{fa}$	$w_{f*}$	$w_{*a}$	$w_{**}$
$w_{fa}$	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		
$w_{f*}$	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		
$w_{*a}$	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		
$w_{**}$	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		

**(E2)** "Let the weather be fine!"

	$w_{fa}$	$w_{f*}$	$w_{*a}$	$w_{**}$
$w_{fa}$	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
$w_{f*}$	<input type="checkbox"/>			
$w_{*a}$	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
$w_{**}$	<input type="checkbox"/>	<input type="checkbox"/>		

**(E3)** "You attack at dawn."

In this case, there are six cells that are ticked in the matrices of both E1 and E2. However, these six cells are not all ticked in the matrix for the conclusion, E3. So this argument is invalid.

### 11.3 Formal System: Accessibility Semantics and KDDc4

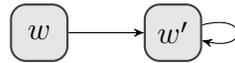
In “Preprescription Semantics and KDDc4” (unpublished), Parsons proves that the logic produced by this two-dimensional semantics is the normal modal logic KDDc4. That is, the normal propositional modal logic **K** with the following additional axioms (see chapter 12 for a full presentation of **K**):

- D**  $!\neg\phi \rightarrow \neg!\phi$
- Dc**  $\neg!\phi \rightarrow !\neg\phi$
- 4**  $!\phi \rightarrow !!\phi$

In addition to the preprescription semantics outlined above, this logic also has a standard one-dimensional possible-worlds semantics. The axioms of KDDc4 correspond, respectively, to the following constraints on the accessibility relation,  $R$  (Parsons (2014), 1):

- ( $\eta$ )  $R$  is *extensible* for all  $w$ , there is a  $w'$  such that  $Rww'$
- ( $\phi$ )  $R$  is *functional* for all  $w, w', w''$ , if  $Rww'$  and  $Rww''$ , then  $w' = w''$
- ( $\tau$ )  $R$  is *transitive* for all  $w, w', w''$ , if  $Rww'$  and  $Rw'w''$ , then  $Rww''$

This accessibility relation produces models that contain pairs of interconnected worlds like so (where an arrow from any  $w$  to any  $w'$  indicate that  $Rww'$  holds):



### 11.4 Parsons’ Theory on the Success Criteria

These preprescription semantics are two-dimensional, with truth-condition worlds and compliance-condition worlds. Because imperatives are evaluated in terms of the compliance-condition worlds, and not in terms of truth-condition worlds, they do not have truth-conditions, so imperatives are not truth-apt. Parsons’ theory is also conservative: suppose  $p_1, \dots, p_n$  and  $q$  are all declarative. Then, the first world in each pair solely determines whether it is in the preprescription, so it collapses into a proposition (rather, it collapses into a set of

worlds, each with a wholly impotent buddy). Thus, the preproscriptions  $p_1, \dots, p_n$  entail the preproscription  $q$  whenever the corresponding propositions  $p_1, \dots, p_n$  entail the corresponding proposition  $q$ . It is also general, because pure declarative sentences, pure imperative sentences, and complex sentences are all preproscriptions, so the same definition of entailment applies no matter what type  $p_1, \dots, p_n$  and  $q$  are.

However, Parsons' theory fails on the criterion of adequacy. That is, it sometimes predicts that there is entailment when there is none. In particular, this theory has the following undesirable results:

1.  $!\neg\phi \leftrightarrow \neg!\phi$
2. For every model and every formula  $\phi$ , either  $!\phi$  or  $!\neg\phi$  holds
3.  $!(\phi \vee \psi) \rightarrow (!\phi \vee !\psi)$

I will discuss them each in turn.

### 11.4.1 Equivalence of Internal and External Negation

First, axioms D and Dc together entail  $!\neg\phi \leftrightarrow \neg!\phi$ . That is, internal and external negation are equivalent. Roughly, the first is a command (to make it the case that  $\neg\phi$ ) and the second is a negation (of the command to make it the case that  $\phi$ ). Of course,  $!\phi$  is not a command (just as  $\phi$  is not an assertion: commands and assertions are speech-acts), but *using*  $!\neg\phi$  amounts to issuing a command (or request, etc) that  $\neg\phi$ , whereas *using*  $\neg!\phi$  does not. It is possible to endorse  $\neg!\phi$  without endorsing  $!\neg\phi$ .

It must be noted that  $\neg!\phi$  is not expressible in a single sentence of English, except as a *permissive* sentence. That is, if we take “ $\phi$ ” to be “you attack at dawn,” we can express  $!\neg\phi$  and  $\neg!\phi$  as follows:

	<b>Literal reading</b>	<b>More natural sentence</b>
$!\neg\phi$ :	“let it be that it is not the case that you attack at dawn!”	“Do not attack at dawn!”
$\neg!\phi$ :	“it is not the case that let it be that you attack at dawn!”	“You may refrain from attacking at dawn!”

In response to this, Parsons says that permissives are not the negations of imperatives, because the permissive sentence “doesn't even have the word “not” in it, and nor is any part

of it in the imperative mood.” He says that “there is no way in English (or any other natural language that I am aware of) of getting an act of permission by negating an imperative. So there is no reason to draw a semantically significant distinction between internal and external negation in relation to the imperative mood” (Parsons (2013): 82).

However,  $\neg!\phi$  is not a permissive. It is the negation of an imperative. It turns out (I think) that it is equivalent to a permissive, but that takes some extra steps. As an analogy, consider the “possibly” and “necessarily” of modal logic.  $\neg\Box\phi$  means “it is not the case that necessarily  $\phi$ .” As it turns out, it is equivalent to  $\Diamond\neg\phi$  (“possibly, it is not the case that  $\phi$ ”), but this must be built into the logic. It is incorrect to pronounce  $\neg\Box\phi$  as “possibly, it is not the case that  $\phi$ ,” even though they are equivalent. Similarly, it is incorrect to pronounce “7+5” as “twelve,” even though the two are equivalent (or even identical). There is an extra step required (the extra step being the derivation that they are equivalent).

So, by parity,  $\neg!\phi$  should not be pronounced as “you may refrain from doing  $p$ !” even though it turns out to be equivalent to this. Instead, the sentence “ $\neg!\phi$ ” is not expressible in English. However, the fact that  $\neg!\phi$  is not directly expressible in English should not count against its being distinct from  $!\neg\phi$ . In general, the fact that something is not expressible in (any) natural language should not count against its having semantic content. Also, in this case,  $\neg!\phi$  is *indirectly* expressible in English. We can express the semantic content by using a permissive sentence, as above. Alternatively, we can use several sentences to express it; this often happens in the “taking back” of commands:

### Scenario 1

Suppose a mother says to her daughter: “Do your homework before you watch TV!” and she responds “I only have Maths homework, and I already understand how to do differentiation. I got 100% on the practice test.” The mother then says, “oh, never mind then.”

It isn’t that the daughter has no homework (so it’s not a case of a conditional command with a false antecedent), or that she’s already done her homework (the command isn’t already complied with). The mother has revised her commands; specifically, she has “taken back” her command that the daughter do her homework. She would endorse  $\neg!\phi$ , and not  $!\neg\phi$ : she doesn’t mind if she wants to do her homework anyway. She just wants to make it clear that  $!\phi$  is *not* in force. Upon (a small amount of) further consideration, it is clear that endorsing  $\neg!\phi$  amounts to giving permission (to refrain from doing  $\phi$ ). Just as denying  $p$  is not the same as asserting  $\neg p$  but it amounts to (in some sense) endorsing it, denying  $!\phi$  is not the same as giving permission to refrain from doing  $\phi$  but it amounts to it.

Further, there are other ways to express commands (or command-like speech-acts) that do not use the imperative mood. That is, there are non-imperative sentences that sometimes approximate imperatives. For example, (the performative utterance) “I command that you do  $\phi$ ,” “I want you to do  $\phi$ ” (if said sternly, this can be a command, and even if not, it can be a request or a piece of advice), and “you ought to do  $\phi$ ” (perhaps if thought of as a piece of advice). Granted, none of these are *synonymous* with “do  $\phi$ !” (see chapters 5 and 6), but they are close approximations. In each of these cases, there is a semantic difference between internal and external negation:

<b>Internal Negation</b>	<b>External Negation</b>
“I command you not to do $\phi$ ”	$\neq$ “I do not command that you do $\phi$ ”
“I want you not to do $\phi$ ”	$\neq$ “I do not want you to do $\phi$ ”
“You ought not to do $\phi$ ”	$\neq$ “It is not the case that you ought to do $\phi$ ”

Because there are grammatically correct ways to express  $\neg!\phi$ , as distinct from  $!\neg\phi$ , it is clear that there is some semantic difference between them. There is a semantic distinction to be made, and so we should make it in our formal language. We could, in fact, employ the rough and ready test from section 2.3.3, and say that  $\neg!\phi \therefore !\neg\phi$  fails that rough and ready test. Because these formulas *can* be expressed in these ways, it is clear that someone could accept  $\neg!\phi$  and deny  $!\neg\phi$  without making a logical error. So, an adequate theory should not make this valid, but Parsons’ theory does.

## 11.4.2 Imperative Exhaustion

Second, for every model and every formula  $\phi$ , either  $!\phi$  holds or  $!\neg\phi$  holds. In Parsons’ models, we have pairs of worlds  $\langle w, w' \rangle$ , where  $w$  sees only  $w'$ , and  $w'$  sees only itself. We can think of the second member as an “imperatively ideal” world: imperatively as opposed to morally or any other measure of idealness. Non-ideal worlds see exactly one ideal world, and ideal worlds see only themselves. Intuitively, in a pair  $\langle w, w' \rangle$ ,  $w'$  is the world at which all the commands that are *in force* at  $w$  are complied with. “In force” may mean that it has been commanded, or that it follows from something else that has been commanded.

We can assume that these ideal worlds are complete, that is, at each world, for all  $\phi$ , either  $\phi$  is true or  $\neg\phi$  is true there. This is because if some worlds can be incomplete, then Parsons will face the usual sorts of incompleteness problems (see Rosen (1990)). So, it follows that at every world (including the actual world), for all  $\phi$ , either  $!\phi$  or  $!\neg\phi$  is in force. So for every preproscription,  $\phi$ , either we are commanded to make it true, or we are commanded to

make it false. For example, suppose Kirsten is deciding whether or not to take an umbrella to work today. There are two possibilities: she takes an umbrella, or she does not take an umbrella. According to Parsons’ semantics, there is only one ideal world,  $w'$ , accessible from our world,  $w$ . So, at  $w'$ , either she takes an umbrella or she does not take an umbrella. One of those things is determinately true at  $w'$ . Say at  $w'$  she *does* take an umbrella. That means that “Kirsten, take an umbrella to work today!” is complied with at  $w'$ . Because  $w'$  is her world’s imperatively ideal world, that means that “Kirsten, take an umbrella to work today!” is in force at her world. However, if at  $w'$  she does *not* take an umbrella, then “Kirsten, do not take an umbrella to work today!” is in force at her world.

This just doesn’t capture the nature of imperatives well at all. There are many things that are not commanded either way (in fact, the vast majority of things are not commanded either way). It does not matter, imperatively speaking, when no command has been promulgated, whether or not Kirsten takes an umbrella to work today. There is no command in force that she takes one, nor that she does not take one. It is entirely up to her. It is genuinely *open* whether or not she takes it: neither option is *imperatively ideal*.

### 11.4.3 Distribution of ! over Disjunction

Third, Parsons’ theory makes  $!(\phi \vee \psi)$  entail  $!\phi \vee !\psi$ . (For a proof, suppose  $!(\phi \vee \psi)$  holds at  $w$ . It follows (by definition of !) that  $\phi \vee \psi$  holds at  $w'$ , but then (by definition of  $\vee$ ), either  $\phi$  or  $\psi$  holds at  $w'$ . If  $\phi$  holds at  $w'$ , then  $!\phi$  holds at  $w$ , and thus (by disjunction introduction)  $!\phi \vee !\psi$  holds at  $w$ . On the other hand, if  $\psi$  holds at  $w'$ , then  $!\psi$  holds at  $w$ , and thus  $!\phi \vee !\psi$  holds at  $w$ ). We cannot express the difference between  $!(\phi \vee \psi)$  and  $!\phi \vee !\psi$  in two different English sentences, but this is not enough to conclude that there is no semantic difference there to be formalised. Where  $\phi$  is “you attack” and  $\psi$  is “you climb that hill,” both  $!\phi \vee !\psi$  and  $!(\phi \vee \psi)$  would be translated as “attack or climb that hill!”

Parsons argues that “there is no scope ambiguity in the natural language disjunctive imperative” (Parsons (2013): 82.) He says that there is only one meaning of disjunctions like “attack or climb that hill!”, not two meanings as would be required if there were a difference between  $!(\phi \vee \psi)$  and  $!\phi \vee !\psi$ . That is, there is no way of expressing the difference between  $!(\phi \vee \psi)$  and  $!\phi \vee !\psi$  in natural language, so there should be no semantic distinction made in imperative logic.

However, as with the internal/external negation case, so with this disjunction-scope case. We cannot express the difference in two different English sentences, but this is not enough to conclude that there is no semantic difference there to be formalised. Where  $p$  is “you attack”

and  $q$  is “you climb that hill,” both  $!p \vee !q$  and  $!(p \vee q)$  would be translated as “attack or climb that hill!” However, there are two different things the commander could be commanding. The first is a disjunction between commands; that is, either “attack!” is imperatively obligatory or “climb the hill!” is imperatively obligatory. The second is a *disjunctive command*; that is, it is obligatory to comply with the disjunction (by complying with *either* of the disjuncts).

Normally, when we command a disjunction such as “attack or climb that hill!”, we mean it as a truly disjunctive, or choice-giving, command. This is because, normally, what we command is entirely up to us, so there is no uncertainty to be had. The usual use of a declarative disjunction is when we don’t know or we can’t remember which of the two disjuncts is true, for example, when asked where the dairy is, responding that “the dairy is on either George St or Great King St” would be strange (and unhelpful, if not outright dishonest) if we know exactly where it is. It is conversationally implicated (see section 1.3.1, and also Grice (1989)) that we do not remember which of those two streets the dairy is on. However, in the case of imperatives, the analogous case is very uncommon. Because it is (generally) *up to you* what you command, it cannot be that you cannot *remember* which of  $!\phi$  or  $!\psi$  you are commanding. That is, although it is perfectly acceptable (and common) to say things like “the dairy is on either George St or Great King St, but I can’t remember which,” it would generally be odd to say “either attack or climb the hill, but I can’t remember which.” However, the fact that something is generally odd is not sufficient to say that the sentence has no semantic content. For that, it would have to be not just unusual, but impossible. But it is possible:

### **Scenario 2**

Suppose an army Lieutenant has the authority to command Privates, but also passes down commands from the Colonel. Suppose she is forgetful, or perhaps on this particular morning is sleep-deprived and a little hungover from the officers’ party last night. She turns up at the morning drill and says “either attack or climb the hill!: I forget which.”

The soldiers would be perplexed, having not been given enough information to go and comply with the command. They still don’t know which of “attack!” and “climb that hill!” they are (imperatively) obliged to comply with. No command has been promulgated. Contrast this with a different day:

### **Scenario 3**

Suppose the army Lieutenant is clear-headed and means to give the soldiers a

genuine choice as to whether they will attack or climb the hill. In that case, the Lieutenant could make this clear by saying “either attack or climb the hill!: it’s up to you.”

This time, they have enough information to go and comply with the (disjunctive) command: they can choose which disjunct to comply with. Parson’s theory can make no sense of someone who makes it clear that they mean the second: “either do X or do Y! It’s up to you.” They could even spell out their permissive presuppositions: “do X or Y! What I mean is, I am commanding you to do X or Y, but permitting you to refrain from doing X, and permitting you to refrain from doing Y: I am just not permitting you to refrain from doing *both* X and Y.” Using talk of imperatively ideal worlds, it seems that there are (at least) three ideal worlds: one at which you do X but not Y, one at which you do Y but not X, and one at which you do both X and Y. It is genuinely up to you which of these three options you choose (that is, none of them are obligatory). When a disjunctive command is issued, it is not the case that either disjunct is commanded. Consider the following case:

#### Scenario 4

Suppose the wind is causing a window in a house to bang open and closed. A teenager who lives in this house is told by his mother: “Oh, for heaven’s sake, either shut that window or latch it open!” So, he complies by shutting the window. Then, a few minutes later, his father comes into the room, and says “who shut that window? It’s boiling in here!” The teenager replies that his mother told him to.

The teenager has, in fact, said something false. His mother did *not* tell him to shut the window. She was satisfied when he did shut it, but that is because she didn’t mind how (of the two ways) he complied with the command.

So, there are possible scenarios (for example, scenario 2) where we express  $!\phi\vee!\psi$ , as distinct from  $!(\phi\vee\psi)$ . Also, there are scenarios (for example, scenario 4) where we express  $!(\phi\vee\psi)$  and it does not entail  $!\phi\vee!\psi$ . So,  $!\phi\vee!\psi$  is semantically distinct from, and is not entailed by,  $!(\phi\vee\psi)$ .

So, Parsons’ imperative logic, based on the modal logic KDDc4 along with his prepositional semantics, has three undesirable consequences:  $!\neg\phi \leftrightarrow \neg!\phi$ , for every model and every formula  $\phi$ , either  $!\phi$  or  $!\neg\phi$  holds, and  $!(\phi\vee\psi) \rightarrow (!\phi\vee!\psi)$ . It is, thus, inadequate, and so fails on the success criteria that Parsons himself endorses.

We have seen that the other theories of imperative logic, from that of R. M. Hare to that of Peter Vranas, all fail in quite major ways. Then, we saw that Brian Chellas' theory fails on the criterion of non-cognitivism about imperatives – he fails to give an explanation for how declaratives are truth-apt and imperatives are not. Josh Parsons' theory suffers from three counterexamples: it makes internal and external negation equivalent, it makes it so that for any arbitrary  $\phi$ , either  $!\phi$  holds or  $!\neg\phi$  holds, and it makes  $!(\phi \vee \psi)$  equivalent to  $!\phi \vee !\psi$ . In the next chapter, I propose a new account of imperative logic that solves these problems with both Chellas' and Parsons' view. It turns out that when we solve the problems with each one, we arrive at the same place.

## Chapter 12

# Relational Preprescription Semantics and KD45

We have seen that Parsons' preprescription semantics (and corresponding accessibility semantics and modal logic  $KDDc4$ ), though non-cognitivist about imperatives, general, and conservative, is inadequate due to three counterexamples. It makes external negation ( $\neg!\phi$ ) equivalent to internal negation ( $!\neg\phi$ ), it makes it the case that for all  $\phi$ , either  $!\phi$  or  $!\neg\phi$  holds, and it makes  $!(\phi \vee \psi)$  equivalent to  $(!\phi \vee !\psi)$  – that is,  $!$  distributes over disjunction. We have also seen that Chellas's logic of imperatives, related to the modal logic KD45, though general, conservative, and adequate, is not non-cognitivist about imperatives. It does not attempt to explain how imperatives “hold” or “fail to hold,” and without any such addition to his theory, he has not given a full semantics, and so there is no account of truth for declarative sentences nor any account of whatever the analogue is for imperative sentences. When we fix each of these theories, we get the same result: Relational Preprescription Semantics, which are equivalent to the modal logic KD45, and which I will outline in this chapter.

### 12.1 Relational Preprescription Semantics

The problems with Parsons' theory are all caused by the constraint of *functionality*. It is this, the constraint that specifies that each world has no more than one ideal world, that makes  $\neg!\phi$  equivalent to  $!\neg\phi$ , makes  $!(\phi \vee \psi)$  equivalent to  $(!\phi \vee !\psi)$ , and makes it the case that either  $!\phi$  or  $!\neg\phi$  hold for every arbitrary  $\phi$ . So, I propose that we remove this constraint, and allow worlds to have more than one ideal world. I keep the constraint of extensibility, so worlds must still have at least one ideal world, and I keep the constraint of transitivity,

so that if a world  $w$  has an ideal world  $w_1$ , and  $w_1$  has an ideal world  $w_2$ , then  $w_2$  is also one of  $w$ 's ideal worlds.

Now that we have worlds having multiple ideal worlds, rather than just one, we must also make the accessibility relation *euclidean*, so that these ideal worlds can all see each other. This is to ensure that we get the desirable result that any time we have an imperative operator (either imperative obligation or imperative permission), adding any number of either type has no effect. That is, any iteration of imperative operators is vacuous.

### 12.1.1 Functional Precriptions Again

Recall that, for Parsons, precriptions are sets of pairs of worlds, and his precription semantics are as follows:

$$\begin{aligned}
V(\phi) &= \text{where } \phi \text{ is a simple declarative, } \{\langle w, w_1 \rangle \mid \phi \text{ is true at } w\}. \\
V(\phi) &= \text{where } \phi \text{ is a simple imperative, } \{\langle w, w_1 \rangle \mid \phi \text{ is complied with at } w_1\}. \\
V(\phi \wedge \psi) &= \{\langle w, w_1 \rangle \mid \langle w, w_1 \rangle \in V(\phi) \text{ and } \langle w, w_1 \rangle \in V(\psi)\}. \\
V(\phi \vee \psi) &= \{\langle w, w_1 \rangle \mid \langle w, w_1 \rangle \in V(\phi) \text{ or } \langle w, w_1 \rangle \in V(\psi)\}. \\
V(\neg\phi) &= \{\langle w, w_1 \rangle \mid \langle w, w_1 \rangle \notin V(\phi)\}. \\
V(\phi \rightarrow \psi) &= \{\langle w, w_1 \rangle \mid \langle w, w_1 \rangle \notin V(\phi) \text{ or } \langle w, w_1 \rangle \in V(\psi)\}. \\
V(!\phi) &= \{\langle w, w_1 \rangle \mid \langle w_1, w_1 \rangle \in V(\phi)\}.
\end{aligned}$$

### 12.1.2 Relational Precriptions

To avoid the counterexamples of Parsons' theory, I propose a new account of precriptions, where worlds can have *more than one* ideal world. Relational precriptions are sets of pairs  $\langle w, U \rangle$ , where  $w$  is a possible world and  $U$  is a *set* of possible worlds. The semantics are as follows (where  $V(\phi)$  is "the content of  $\phi$ "):

$$\begin{aligned}
V(p) &= \{\langle w, U \rangle \mid p \text{ is true at } w\}. \\
V(\neg\phi) &= \{\langle w, U \rangle \mid \langle w, U \rangle \notin V(\phi)\}. \\
V(\phi \wedge \psi) &= \{\langle w, U \rangle \mid \langle w, U \rangle \in V(\phi) \text{ and } \langle w, U \rangle \in V(\psi)\}. \\
V(\phi \vee \psi) &= \{\langle w, U \rangle \mid \langle w, U \rangle \in V(\phi) \text{ or } \langle w, U \rangle \in V(\psi)\}. \\
V(\phi \rightarrow \psi) &= \{\langle w, U \rangle \mid \langle w, U \rangle \notin V(\phi) \text{ or } \langle w, U \rangle \in V(\psi)\}. \\
V(!\phi) &= \{\langle w, U \rangle \mid \langle u, U \rangle \in V(\phi) \text{ for every } u \in U\}. \\
V(i\phi) &= \{\langle w, U \rangle \mid \langle u, U \rangle \in V(\phi) \text{ for some } u \in U\}.
\end{aligned}$$

Entailment (between preproscriptions), and validity, are defined in the same way as Parsons. Where  $p_1, \dots, p_n$  and  $q$  are all preproscriptions:

**Entailment**

$p_1, \dots, p_n$  jointly *entail*  $q$  iff every member of all of  $p_1, \dots, p_n$  is also a member of  $q$ .

**Validity**

An argument  $p_1, \dots, p_n \therefore q$  is *valid* iff  $p_1, \dots, p_n$  jointly *entail*  $q$ .

### 12.1.3 RPS Interpretations

An *RPS – interpretation* is any pair  $(W, V)$ , where:

- $W$  is a set
- $V$  is a function from sentences to subsets of  $W \times 2^W$ , where  $W \times 2^W$  is the set of all pairs  $\langle w, U \rangle$  where  $x$  is a member of  $W$  and  $U$  is a (non-empty) set of members of  $W$ , and where  $V$  satisfies the following valuation clauses:

<b>I</b> <sub>RPS</sub> :	$\langle w, U \rangle \in V(\neg\phi)$	iff $\langle w, U \rangle \notin V(\phi)$
<b>II</b> <sub>RPS</sub> :	$\langle w, U \rangle \in V(\phi \wedge \psi)$	iff $\langle w, U \rangle \in V(\phi)$ and $\langle w, U \rangle \in V(\psi)$
<b>III</b> <sub>RPS</sub> :	$\langle w, U \rangle \in V(\phi \vee \psi)$	iff $\langle w, U \rangle \in V(\phi)$ or $\langle w, U \rangle \in V(\psi)$
<b>IV</b> <sub>RPS</sub> :	$\langle w, U \rangle \in V(\phi \rightarrow \psi)$	iff $\langle w, U \rangle \notin V(\phi)$ or $\langle w, U \rangle \in V(\psi)$
<b>V</b> <sub>RPS</sub> :	$\langle w, U \rangle \in V(!\phi)$	iff for all $u$ in $U$ , $\langle u, U \rangle \in V(\phi)$
<b>VI</b> <sub>RPS</sub> :	$\langle w, U \rangle \in V(j\phi)$	iff for some $u$ in $U$ , $\langle u, U \rangle \in V(\phi)$

A sentence  $\theta$  is *designated* (or, it *holds*) at  $\langle w, U \rangle$  in  $(W, V)$  iff  $\langle w, U \rangle \in V(\theta)$ .

A sentence  $\theta$  is an *RPS – consequence* of a set of sentences  $\Gamma$ , that is  $\Gamma \models_{RPS} \theta$ , iff, for every RPS-interpretation  $(W, V)$  and every  $\langle w, U \rangle \in W \times 2^W$ , if every member of  $\Gamma$  is designated at  $\langle w, U \rangle$  in  $(W, V)$ , then  $\theta$  is designated at  $\langle w, U \rangle$  in  $(W, V)$ .

## 12.2 KD45

To correspond with these Relational Preproscription Semantics, we can now outline the accessibility semantics of this functionality-less modal logic.

## 12.2.1 KD45 Accessibility Semantics

A *KD45-interpretation* is any triple  $(W, R, V)$ , where:

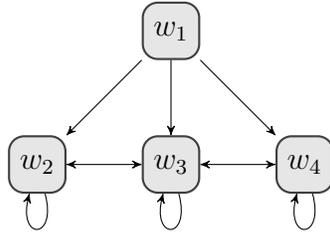
- $W$  is a set
- $R$  is an accessibility relation over  $W$  that satisfies the following constraints:

$R$  is *extensible* for all  $w \in W$ , there is a  $w_1 \in W$  such that  $Rww_1$

$R$  is *transitive* for all  $w, w_1, w_2 \in W$ , if  $Rww_1$  and  $Rw_1w_2$ , then  $Rww_2$

$R$  is *euclidean* for all  $w, w_1, w_2 \in W$ , if  $Rww_1$  and  $Rww_2$ , then  $Rw_1w_2$

This produces models with independent clusters of worlds arranged as follows, where an arrow from any  $w_n$  to any  $w_m$  indicates that  $Rw_nw_m$ :



- $V$  is a function from sentences to subsets of  $W$ , where  $V$  satisfies the following valuation clauses:

<b>I</b> <sub>KD45</sub> : $w \in V(\neg\phi)$	iff $w \notin V(\phi)$
<b>II</b> <sub>KD45</sub> : $w \in V(\phi \wedge \psi)$	iff $w \in V(\phi)$ and $w \in V(\psi)$
<b>III</b> <sub>KD45</sub> : $w \in V(\phi \vee \psi)$	iff $w \in V(\phi)$ or $w \in V(\psi)$
<b>IV</b> <sub>KD45</sub> : $w \in V(\phi \rightarrow \psi)$	iff $w \notin V(\phi)$ or $w \in V(\psi)$
<b>V</b> <sub>KD45</sub> : $w \in V(!\phi)$	iff for all $w_1$ s.t. $Rww_1$ , $w_1 \in V(\phi)$
<b>VI</b> <sub>KD45</sub> : $w \in V(i\phi)$	iff for some $w_1$ s.t. $Rww_1$ , $w_1 \in V(\phi)$

A sentence  $\theta$  is *designated* at  $w$  in  $(W, R, V)$  iff  $w \in V(\theta)$ .

A sentence  $\theta$  is a *KD45-consequence* of a set of sentences  $\Gamma$ , that is  $\Gamma \models_{KD45} \theta$ , iff, for every KD45-interpretation  $(W, R, V)$  and every  $w \in W$ , if every member of  $\Gamma$  is designated at  $w$  in  $(W, R, V)$ , then  $\theta$  is designated at  $w$  in  $(W, R, V)$ .

## 12.2.2 KD45 Proof Theory

These accessibility semantics are sound and complete with respect to the axioms of KD45 (for soundness and completeness proofs see section 12.2.3), which is the normal modal logic

K with the additional axioms D, 4, and 5. That is:

**Axioms:**

All classical tautologies

**K**  $!(\phi \rightarrow \psi) \rightarrow (!\phi \rightarrow !\psi)$

**D**  $!\neg\phi \rightarrow \neg!\phi$

**4**  $!\phi \rightarrow !!\phi$

**5**  $!j\phi \rightarrow !j!\phi$

**Rules of Inference:**

Modus Ponens      If  $\vdash \phi, \phi \rightarrow \psi$  then  $\vdash \psi$

Necessitation      If  $\vdash \phi$  then  $\vdash !\phi$

### 12.2.3 Soundness and Completeness Proofs

For soundness, we show that the rules of inference produce valid formulas from valid formulas, and that each of the axioms is valid. Then, it follows that every line of a proof must be a valid formula, so the last line (the formula being proved) is also valid.

For illustration, take Modus Ponens. Suppose  $\phi \rightarrow \psi$  is true at a world,  $w$ . Then, either  $\phi$  is not true at  $w$ , or  $\psi$  is true at  $w$ . So if  $\phi$  is also true at  $w$ , then  $\psi$  must be true at  $w$ . Thus, if  $\phi$  and  $\phi \rightarrow \psi$  are both valid in a model (that is, both are true at every world of the model), then  $\psi$  must also be valid in that model.

This can also be done in similar fashion for Necessitation, as well as each of the axioms (see Fitting and Mendelsohn (1998): 74-75 for details).

For completeness, we extend the completeness proofs for classical logic. First, we define consistency and maximal consistency: a finite set  $\{X_1, X_2, \dots, X_n\}$  of formulas is *KD45-consistent* iff  $(X_1 \wedge X_2 \wedge \dots \wedge X_n) \rightarrow \perp$  is not provable in the proof theory of KD45. An infinite set is KD45-consistent iff every finite subset is KD45-consistent. A set  $S$  of formulas is *maximally KD45-consistent* iff  $S$  is KD45-consistent, and no proper extension of it is KD-45-consistent.

Then, we prove Lindenbaum's theorem, which is an instance of the theorem that any consistent set may be extended to a maximally consistent set: if a set of formulas  $S$  is KD45-consistent, it can be extended to a maximally KD45-consistent set. We also prove that if the set  $\{\neg!Y, !X_1, !X_2, \dots\}$  is KD45-consistent, then so is  $\{\neg Y, X_1, X_2, \dots\}$  (see Fitting and Mendelsohn (1998): 76-77 for the details).

We then define a canonical model for KD45 as follows:  $\mathcal{G}$  is the set of *all* maximally

KD45-consistent sets of formulas. If  $\Gamma$  and  $\Delta$  are in  $\mathcal{G}$ , set  $\Gamma\mathcal{R}\Delta$  provided, for each formula in  $\Gamma$  of the form  $!x$ , the corresponding formula  $x$  is in  $\Delta$ . And finally, for each propositional letter  $p$  and each  $\Gamma \in \mathcal{G}$ , set  $\Gamma \Vdash p$  just in case  $p \in \Gamma$ . We have now completely defined a model  $\mathcal{M} = \langle \mathcal{G}, \mathcal{R}, \Vdash \rangle$ . This is the *canonical model* for KD45. We can prove the following result about canonical models: let  $\langle \mathcal{G}, \mathcal{R}, \Vdash \rangle$  be the *canonical model* for KD45. For every formula  $x$  and for every  $\Gamma \in \mathcal{G}$ ,  $\Gamma \Vdash x$  if and only if  $x \in \Gamma$  (see Fitting and Mendelsohn (1998): 78-79 for details).

Finally, using these definitions, we prove completeness by proving its contrapositive: that if  $\not\vdash \theta$  then  $\not\Vdash \theta$ :

Suppose  $\not\vdash X$ , that is,  $X$  is a formula that has no proof. Then we know that the set  $\{\neg X\}$  is KD45-consistent (because if it were not, then  $X \rightarrow \perp$  would be provable, but this is equivalent to  $X$ ). We can then extend  $\{\neg X\}$  to a maximally KD45-consistent set  $\Gamma_0$ . Let  $\langle \mathcal{G}, \mathcal{R}, \Vdash \rangle$  be the *canonical model* for KD45. We know  $\Gamma_0 \in \mathcal{G}$ . But also,  $\neg X \in \Gamma_0$ , since  $\Gamma_0$  extends  $\{\neg X\}$ , so  $X \notin \Gamma_0$ . But then,  $\Gamma_0 \not\Vdash X$ , so  $X$  is not KD45-valid (for details, see Fitting and Mendelsohn (1998): 79).

## 12.3 Equivalence of RP Semantics and KD45 Semantics

Relational Preprescription Semantics are equivalent to the accessibility semantics of the modal logic KD45. As such, because the proof theory of KD45 is sound and complete with respect to the accessibility semantics of KD45, we can conclude that the proof theory of KD45 is sound and complete with respect to Relational Preprescription Semantics. In what follows, I prove that the semantics of KD45 are equivalent to Relational Preprescription Semantics. That is:

**Theorem 1** *RPS is equivalent to KD45.*

### 12.3.1 Equivalence Proof

We want to prove that RPS semantics are equivalent to KD45 semantics, that is:

$$\Gamma \models_{RPS} \theta \text{ iff } \Gamma \models_{KD45} \theta$$

First, we prove:

$$\text{if } \Gamma \models_{KD45} \theta \text{ then } \Gamma \models_{RPS} \theta$$

We do this by taking the contrapositive, that is:

$$\text{if } \Gamma \not\#_{RPS} \theta \text{ then } \Gamma \not\#_{KD45} \theta$$

Suppose that  $\Gamma \not\#_{RPS} \theta$ . Then (by the definition of RPS-consequence), there is an RPS-interpretation  $(W, V)$  and a  $\langle w, U \rangle \in W \times 2^W$ , in which all members of  $\Gamma$  are designated but  $\theta$  is not designated. To show that  $\Gamma \not\#_{RPS} \theta$ , we can construct a KD45-interpretation  $(W', R, V')$  at which, for some  $w \in W'$ , all members of  $\Gamma$  are designated at  $w$  but  $\theta$  is not designated at  $w$ :

- Let  $W'$  be  $W \times 2^W$  (that is, sets of pairs  $\langle w, U \rangle$  where  $w$  is a member of  $W$  and  $U$  is a non-empty set of members of  $W$ )
- Let  $R\langle w, U \rangle\langle w_1, U_1 \rangle$  iff  $w_1 \in U$  and  $U = U_1$
- Let  $V'$  be identical to  $V$

We can check that  $R$  is extensible, transitive and euclidean:

$R$  is *extensible*, because  $U$  is by definition a non-empty set, so for all  $\langle w, U \rangle$ , there exists a  $w_1 \in U$ , and let  $U_1$  be  $U$ , so because  $U = U$ ,  $U = U_1$ . Thus, we know that there exists a  $\langle w_1, U_1 \rangle$  such that  $R\langle w, U \rangle\langle w_1, U_1 \rangle$ .

$R$  is *transitive*, because if  $R\langle w, U \rangle\langle w_1, U_1 \rangle$  and  $R\langle w_1, U_1 \rangle\langle w_2, U_2 \rangle$ , then we know that  $w_1 \in U$  and  $U = U_1$  and that  $w_2 \in U_1$  and  $U_1 = U_2$ . Thus, we know that  $w_2 \in U$  and  $U = U_2$ , so  $R\langle w, U \rangle\langle w_2, U_2 \rangle$ .

$R$  is *euclidean*, because if  $R\langle w, U \rangle\langle w_1, U_1 \rangle$  and  $R\langle w, U \rangle\langle w_2, U_2 \rangle$ , then we know that both  $w_1$  and  $w_2$  are in  $U$ , and  $U = U_1 = U_2$ . Thus, we know that  $w_2 \in U_1$  and  $U_1 = U_2$ , so  $R\langle w_1, U_1 \rangle\langle w_2, U_2 \rangle$ .

$V'$  satisfies the first four valuation clauses of a KD45-interpretation ( $\mathbf{I}_{KD45}$ ,  $\mathbf{II}_{KD45}$ ,  $\mathbf{III}_{KD45}$ , and  $\mathbf{IV}_{KD45}$ ), because  $V'$  is  $V$  (an RPS-interpretation valuation function). So, it satisfies  $\mathbf{I}_{RPS}$ ,  $\mathbf{II}_{RPS}$ ,  $\mathbf{III}_{RPS}$ , and  $\mathbf{IV}_{RPS}$ . Because  $W$  is  $W \times 2^W$ , we can substitute each  $w$  for a  $\langle w, U \rangle$ , and what results is identical with  $\mathbf{I}_{KD45}$ ,  $\mathbf{II}_{KD45}$ ,  $\mathbf{III}_{KD45}$ , and  $\mathbf{IV}_{KD45}$ .

$V'$  satisfies the fifth clause,  $\mathbf{V}_{KD45}$ , because it satisfies  $\mathbf{V}_{RPS}$ :  $\langle w, U \rangle \in V(!\phi)$  iff for all  $u$  in  $U$ ,  $\langle u, U \rangle \in V(\phi)$ , and we know that  $R\langle w, U \rangle\langle w_1, U_1 \rangle$  iff  $w_1 \in U$  and  $U = U_1$ , so  $R\langle w, U \rangle\langle u, U \rangle$  iff  $u \in U$ . That is, for all  $w$  and all  $u$  in  $U$ ,  $R\langle w, U \rangle\langle u, U \rangle$ , and so we know that for all  $u$  in  $U$ ,  $\langle u, U \rangle \in V(\phi)$  iff for all  $\langle w_1, U_1 \rangle$  s.t.  $R\langle w, U \rangle\langle w_1, U_1 \rangle$ ,  $\langle w_1, U_1 \rangle \in V(\phi)$ . Thus, we know that  $\langle w, U \rangle \in V(!\phi)$  iff for all  $\langle w_1, U_1 \rangle$  s.t.  $R\langle w, U \rangle\langle w_1, U_1 \rangle$ ,  $\langle w_1, U_1 \rangle \in V(\phi)$ .

$V'$  satisfies the sixth clause,  $\mathbf{VI}_{KD45}$ , because it satisfies  $\mathbf{VI}_{RPS}$ :  $\langle w, U \rangle \in V(\imath\phi)$  iff for some  $u$  in  $U$ ,  $\langle u, U \rangle \in V(\phi)$ , and we know that  $R\langle w, U \rangle\langle w_1, U_1 \rangle$  iff  $w_1 \in U$  and  $U = U_1$  and that  $U$  is non-empty, so there exists a  $\langle w_1, U_1 \rangle$  such that  $R\langle w, U \rangle\langle w_1, U_1 \rangle$ . So, we know that for some  $u$  in  $U$ ,  $\langle u, U \rangle \in V(\phi)$  iff for some  $\langle w_1, U_1 \rangle$  such that  $R\langle w, U \rangle\langle w_1, U_1 \rangle$ ,  $\langle w_1, U_1 \rangle \in V(\phi)$ . Thus, we know that  $\langle w, U \rangle \in V(\imath\phi)$  iff for some  $\langle w_1, U_1 \rangle$  such that  $R\langle w, U \rangle\langle w_1, U_1 \rangle$ ,  $\langle w_1, U_1 \rangle \in V(\phi)$ .

So,  $(W', R, V')$  satisfies the definition of a KD45-interpretation. Since  $V'$  is identical to  $V$ , all members of  $\Gamma$  are designated at  $\langle w, U \rangle$  and  $\theta$  is not designated at  $\langle w, U \rangle$  on  $(W', R, V')$ . So,  $\Gamma \not\#_{KD45} \theta$ . So, if  $\Gamma \not\#_{RPS} \theta$  then  $\Gamma \not\#_{KD45} \theta$ . Thus, if  $\Gamma \models_{KD45} \theta$  then  $\Gamma \models_{RPS} \theta$ .

So, we have one direction of the biconditional. Next, we must prove that the other direction also holds. That is:

if  $\Gamma \models_{RPS} \theta$  then  $\Gamma \models_{KD45} \theta$

Again, we do this by proving the contrapositive:

if  $\Gamma \not\#_{KD45} \theta$  then  $\Gamma \not\#_{RPS} \theta$

Suppose  $\Gamma \not\#_{KD45} \theta$ . Then, by the definition of KD45-consequence, there is a KD45-interpretation  $(W, R, V)$  and a  $w \in W$ , such that every member of  $\Gamma$  is designated and  $\theta$  is not designated at  $w$  in  $(W, R, V)$ .

We can construct an RPS-interpretation  $(W', V')$  on which, for some  $w \in W'$ , all members of  $\Gamma$  are designated at  $w$  but  $\theta$  is not designated at  $w$ , so that  $\Gamma \not\#_{RPS} \theta$ .

- Let  $W'$  be the set containing  $w$  and every member of  $U$ , where  $U$  is all  $w_1$  such that  $Rww_1$  ( $w$  may or may not be a member of  $U$ ).
- Let  $V'$  be a function defined as follows:  $\langle w, U \rangle \in V'(\phi)$  iff  $w \in V(\phi)$

It can be shown that  $V'$  satisfies the valuation clauses in the definition of an RPS-interpretation:

$V'$  satisfies  $\mathbf{I}_{RPS}$ , because  $\langle w, U \rangle \in V'(\neg\phi)$  iff  $w \in V(\neg\phi)$  (by definition of  $V'$ ), and  $w \in V(\neg\phi)$  iff  $w \notin V(\phi)$  (by definition of  $V$ , a KD45-interpretation valuation function), and  $w \notin V(\phi)$  iff  $\langle w, U \rangle \notin V(\phi)$  (again, by definition of  $V'$ ). So,  $\langle w, U \rangle \in V'(\neg\phi)$  iff  $\langle w, U \rangle \notin V(\phi)$ .

$V'$  satisfies  $\mathbf{II}_{RPS}$  because  $\langle w, U \rangle \in V'(\phi \wedge \psi)$  iff  $w \in V(\phi \wedge \psi)$ , and  $w \in V(\phi \wedge \psi)$  iff  $w \in V(\phi)$  and  $w \in V(\psi)$ , so  $\langle w, U \rangle \in V'(\phi \wedge \psi)$  iff  $\langle w, U \rangle \in V(\phi)$  and  $\langle w, U \rangle \in V(\psi)$ .

$V'$  satisfies **III**<sub>RPS</sub> because  $\langle w, U \rangle \in V'(\phi \vee \psi)$  iff  $w \in V(\phi \vee \psi)$ , and  $w \in V(\phi \vee \psi)$  iff  $w \in V(\phi)$  or  $w \in V(\psi)$ , so  $\langle w, U \rangle \in V'(\phi \vee \psi)$  iff  $\langle w, U \rangle \in V'(\phi)$  or  $\langle w, U \rangle \in V'(\psi)$ .

$V'$  satisfies **IV**<sub>RPS</sub> because  $\langle w, U \rangle \in V'(\phi \rightarrow \psi)$  iff  $w \in V(\phi \rightarrow \psi)$ , and  $w \in V(\phi \rightarrow \psi)$  iff  $w \notin V(\phi)$  or  $w \in V(\psi)$ , so  $\langle w, U \rangle \in V'(\phi \rightarrow \psi)$  iff  $\langle w, U \rangle \notin V'(\phi)$  or  $\langle w, U \rangle \in V'(\psi)$ .

$V'$  satisfies **V**<sub>RPS</sub> because  $\langle w, U \rangle \in V'(!\phi)$  iff  $w \in V(!\phi)$ , and  $w \in V(!\phi)$  iff for all  $w_1$  s.t.  $Rww_1$ ,  $w_1 \in V(\phi)$ .  $U$  is defined as all  $w_1$  s.t.  $Rww_1$ , so for all  $w_1$  s.t.  $Rww_1$ ,  $w_1 \in V(\phi)$  iff for all  $u$  in  $U$ ,  $u \in V(\phi)$ , and for all  $u$  in  $U$ ,  $u \in V(\phi)$  iff for all  $u$  in  $U$ ,  $\langle u, U \rangle \in V'(\phi)$ . Thus,  $\langle w, U \rangle \in V'(!\phi)$  iff for all  $u$  in  $U$ ,  $\langle u, U \rangle \in V'(\phi)$ .

$V'$  satisfies **VI**<sub>RPS</sub> because  $\langle w, U \rangle \in V'(j\phi)$  iff  $w \in V(j\phi)$  and  $w \in V(j\phi)$  iff for some  $w_1$  s.t.  $Rww_1$ ,  $w_1 \in V(\phi)$ .  $U$  is defined as all  $w_1$  s.t.  $Rww_1$ , and  $R$  is extensible, so  $U$  is non-empty. So, for some  $w_1$  s.t.  $Rww_1$ ,  $w_1 \in V(\phi)$  iff for some  $u$  in  $U$ ,  $\langle u, U \rangle \in V'(\phi)$ . Thus,  $\langle w, U \rangle \in V'(j\phi)$  iff for some  $u$  in  $U$ ,  $\langle u, U \rangle \in V'(\phi)$ .

Because any  $\phi$  is designated at  $\langle w, U \rangle$  in  $(W', V')$  just when  $\phi$  is designated at  $w$  in  $(W, R, V)$ , all members of  $\Gamma$  are designated at  $\langle w, U \rangle$  and  $\theta$  is not designated at  $\langle w, U \rangle$  in  $(W', V')$ . So,  $\Gamma \not\models_{RPS} \theta$ . So, if  $\Gamma \not\models_{KD45} \theta$  then  $\Gamma \not\models_{RPS} \theta$ . So, if  $\Gamma \models_{RPS} \theta$  then  $\Gamma \models_{KD45} \theta$ .

Finally, we have both directions of the biconditional:

$$\Gamma \models_{RPS} \theta \text{ iff } \Gamma \models_{KD45} \theta$$

So, the Relational Preprescription Semantics are equivalent to the accessibility semantics of KD45.

**Corollary 1** *The proof theory of KD45 is sound and complete with respect to Relational Preprescription Semantics.*

The proof of corollary 1 is immediate from the equivalence of RPS and the accessibility semantics of KD45, and the fact that KD45 is sound and complete.

## 12.4 Relationship with Deontic Logic

At this point, it may be worrying that my imperative logic is starting to look a lot like deontic logics that are based on modal systems. KD45, itself, has been proposed as a deontic logic (see Chellas (1980): 193), and even Standard Deontic Logic is the closely related system, KD (see Mally (1926), von Wright (1951), von Wright (1968), Goble (2001)). If RPS is equivalent to a deontic logic, then perhaps it will collapse into a form of cognitivism. If it

can be collapsed into a (cognitivist) deontic logic, then will it not encounter the problems outlined in chapter 6?

I have two points to make regarding this worry. First, RPS is *equivalent to* KD45, but this does not entail that it is *reducible to* it. There is an established precedent of translating one logic into another without any kind of philosophical commitment being required. For example, Kripke semantics can be defined that validate all and only the same arguments as intuitionist logic (see Moschovakis (2004)), but this does not mean that intuitionistic logic is not adding anything to the Kripke semantics. The two semantics are *extensionally* equivalent – the two languages have exactly the same validities. However, they are *intensionally* different, which is philosophically significant – they have different philosophical commitments. Exactly what is going on in cases like these is perhaps difficult to pin down, but it is not uniquely my problem – this is an instance of a wider phenomenon. In general, just because logic A can be translated into the language of logic B does not mean that logic A can be replaced with logic B. They may have very different philosophical explanations of the phenomena.

RPS and KD45 are logically equivalent, but they have very different philosophical implications for the semantics of imperative sentences. In particular, KD45 makes imperatives truth-apt, whereas RPS does not. They are, then, not interchangeable, even though they make all and only the same arguments valid. They are extensionally, but not intensionally, equivalent. In RPS, imperatives (and declaratives) express *preprescriptions*, which are two-dimensional – they have a “truth” dimension and a “compliance” dimension. Pure imperatives express a special kind of preprescription – those for which only the compliance dimension matters. They thus have compliance conditions, rather than truth conditions. These compliance conditions may exhibit logical behaviour *as if* they are truth conditions (in particular, about obligations), but this does not entail that they *really are* truth conditions. This difference between RPS and KD45 is important, for theoretical reasons, even though the two systems behave identically.

The second point I will make is that it is, I believe, a *feature* of RPS that it resembles deontic logic. After all, deontic logic is a logic of *obligations*, and imperatives are used to communicate some kind of obligation. It may not be *moral* obligation, but a sort of *imperative* obligation. Just as we can distinguish moral, legal, and practical obligations as similar in kind but distinct sets of obligations, I propose a kind of “imperative obligation,” which means “obligation due to having been issued an imperative.” Moral statements and imperatives are similar in some important ways – they both (in some sense) *tell you how to*

*act, or tell you to do something.* The kind of reason we have for doing that thing differs: a moral statement gives you a reason to do the thing because of some facts about morality (however that is fleshed out), while an imperative gives you a reason to do the thing because someone told you to do it (particularly if that person had the relevant kind of authority over you). When an imperative is *used* – that is, when a command, request or instruction is issued, by someone in the right sort of authority relationship with you, an obligation is created. It is not a moral obligation, a legal obligation, or a practical obligation. It is an imperative obligation. Because imperative sentences have this obligation-creating feature, it is hardly surprising that their logic turns out to be closely related to a logic of obligations. This might seem like a tension – on one hand, I claim that the logic of moral obligations and the logic of imperatives are very different things, and on the other hand I claim that they’re closely related. However, this is in no way inconsistent. My point is that, while imperatives are not the same as, a subset of, or a superset of moral obligations, they are nonetheless related. They are a different species of obligation.

Thus, it is a positive feature of RPS that it is equivalent to KD45 (a logic of obligations), because imperatives also express obligations, though they are imperative and not moral ones. RPS also adds something to KD45, although they are equivalent – it adds a philosophically satisfying interpretation of imperative sentences as not truth-apt but still expressing some kind of obligation.

## 12.5 RPS/KD45 on the Success Criteria

The theory of Relational Preprescription Semantics is general, conservative, adequate, and non-cognitivist about imperatives. It is general, because the contents of imperatives and declaratives alike are preprescriptions, so the definitions of entailment and validity apply in the same way to arguments made up entirely of declarative sentences, to those made up entirely of imperative sentences, and to those made up of a mixture of declarative and imperative sentences.

It is conservative, because in the special case where we have pure declarative arguments, all the premises and the conclusion will be preprescriptions in which the ideal worlds play no role whatsoever. That is, the content of a declarative sentence is a set of pairs, the first member of which is a possible world at which the sentence is true and the second member of which is irrelevant. So, it is equivalent to the set of possible worlds at which the sentence is true: to a proposition. Thus, when there are no imperative elements involved, we will get

all the expected results of classical propositional logic.

### 12.5.1 Non-Cognitivism about Imperatives

This theory of Relational Preprescription Semantics solves Chellas's theory's lack of non-cognitivism.

Chellas's theory is also based on KD45. However, his theory does not have a philosophically satisfying interpretation. He does not explain what it means for imperative sentences to "hold" or to "fail to hold." RPS, however, can fill this gap for Chellas. For an imperative sentence,  $!p$ , to hold at  $w$ , it must be complied with at all  $w_1$  such that  $Rww_1$ . Because the conditions under which imperatives hold are not truth-conditions, but *compliance*-conditions, holding does not amount to truth for imperative sentences. Rather, holding amounts to "being complied with at all the (imperatively) ideal worlds." So, RPS is non-cognitivist about imperatives. Holding for declarative sentences, on the other hand, amounts to being true. That is, a declarative sentence,  $p$ , holds at  $w$  whenever  $p$  is true at  $w$ . So, RPS is cognitivist about declaratives.

### 12.5.2 Adequacy

RPS is adequate, because it makes the intuitively valid arguments valid and the intuitively invalid arguments invalid. To demonstrate this, we will first consider again our original examples (from chapter 1):

- A1** Join me for a cup of wine, or leave me to drink alone again tonight!
- A2** Don't leave me to drink alone tonight!
- $\therefore$  **A3** Join me for a cup of wine!
  
- B1** If Bran wakes up, kill him immediately!
- B2** Bran wakes up.
- $\therefore$  **B3** Kill Bran immediately!

Arguments A and B are RPS-valid, because in each case, if you comply with both of the premises in every ideal world, then it follows that you comply with the conclusion in every ideal world. Getting these examples right, though, is not the difficult part. So, we will now consider several of the arguments that tripped up the other theories.

First, consider again the problem of unwanted consistencies that I raised against the

various forms of cognitivism. The problem was that those theories made the following two sentences consistent:

- (1) Attack!
- (2) Do not attack!

However, in RPS, (1) and (2) are indeed inconsistent. That is, they cannot both hold at once. For (1) to hold, it must be the case that you (the addressee) attack in every ideal world. But then, for (2) to hold, it must be the case that you do *not* attack in every world. Since there is always at least one ideal world ( $R$  is extensible, or, equivalently,  $U$  is non-empty in every  $\langle w, U \rangle$ ), and you cannot both attack and not attack, (1) and (2) can never hold at the same time – they are inconsistent.

Second, consider the familiar counterexample to Hare’s theory (see chapter 7):

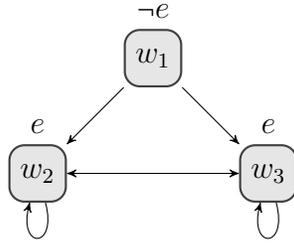
- C1** If you are a faithful subject, rise up Sir George!
- C2** Do not rise up, stay on your knees fellow!
- $\therefore$  **C3** George is not a faithful subject.

Hare’s theory made this argument invalid, when it is in fact valid. RPS correctly diagnoses this argument as valid. For C1 to hold, it must be that if he is a faithful subject (in this world), then George rises in all the ideal worlds. But then for C2 to hold, George does not rise in the ideal worlds. So, it follows that George is not a faithful subject.

Next, consider the counterexample to Hofstadter and McKinsey’s operator analysis (again, see chapter 7):

- D1** Do the dishes, Emily!
- $\therefore$  **D2** Emily will do the dishes.

Hofstadter and McKinsey’s theory (as well as the Predictions theory of imperative cognitivism) makes argument D valid, whereas RPS correctly says it is invalid. Here is a counterexample in KD45 (where  $e$  is “Emily does the dishes”):



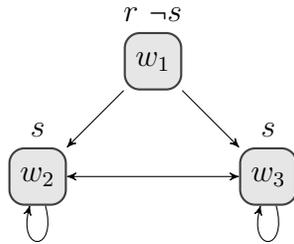
In this model, the premise holds, because Emily does the dishes at both the ideal worlds, but the conclusion does not hold because Emily does not (actually) do the dishes.

Next, consider again the following counterexample to the Smart/Pigden theory:

- E1** Whenever a woman enters the room, stand up!
- E2** A woman has entered the room.
- $\therefore$  **E3** Stand up!

- F1** Whenever a woman enters the room, stand up!
- F2** A woman has entered the room.
- $\therefore$  **F3** You will stand up.

Clearly, argument E is valid and argument F is invalid. The Smart/Pigden theory, recall, made argument E invalid and argument F valid when addressed to a paraplegic. Because RPS is not affected by who the addressee is in any scenario, this contextual fact will not make a difference. So, argument E is RPS-valid (it is of the same form as argument B). Argument F is RPS-invalid. Here is a counterexample in KD45 (where  $r$  is “a woman enters the room” and  $s$  is “you stand up”):

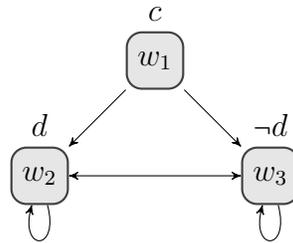


In this model, the premises hold at  $w_1$ , because a woman does enter the room (in  $w_1$ ), and also you stand up in all the ideal worlds, so the conditional “if a woman enters the room, stand up!” holds. However, the conclusion does not hold at  $w_1$ , because you do not, in  $w_1$ , stand up.

Next, consider the following counterexample to Vranas’s theory:

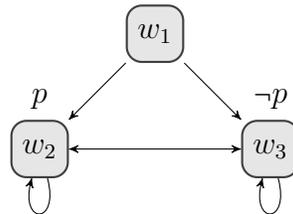
- G1** You have a conclusive reason to dance.  
 $\therefore$  **G2** Dance!

Argument G came out as valid on Vranas’s view. However, RPS correctly diagnoses this argument as invalid. Here is a counterexample in KD45 (where  $c$  is “you have a conclusive reason to dance” and  $d$  is “you dance”):



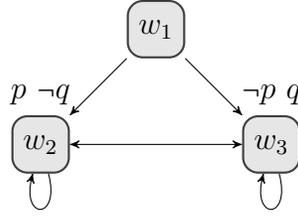
In this model, the premise is true at  $w_1$ , because you have a conclusive reason to dance at this world, but it is not the case that at all ideal worlds, you dance, so the conclusion does not hold.

Finally, RPS avoids the counterexamples to Parsons’ theory. RPS does *not* imply  $\neg! \phi \leftrightarrow !\neg \phi$ . One direction of the biconditional,  $!\neg \phi \rightarrow \neg! \phi$ , is an axiom, so that direction holds, but the other direction,  $\neg! \phi \rightarrow !\neg \phi$ , does not hold. Here is a counterexample in KD45:



At  $w_1$ ,  $\neg! p$  holds, since  $p$  is not complied with in every ideal world, so  $!p$  does not hold. However,  $!\neg p$  does not hold, since  $\neg p$  is not complied with in  $w_2$ . Because at  $w_1$  in this model, neither  $!p$  nor  $!\neg p$  hold, this also serves as a counterexample to the problem of imperatival exhaustion (that for all  $\phi$ , either  $!\phi$  or  $!\neg \phi$  holds).

Further, RPS does not imply  $!(\phi \vee \psi) \rightarrow !\phi \vee !\psi$ . Here is a counterexample in KD45:



At  $w_2$ ,  $p$  holds, so (by disjunction introduction),  $p \vee q$  holds. At  $w_3$ ,  $q$  holds, so (again by disjunction introduction),  $p \vee q$  holds. Thus, at  $w_1$ ,  $!(p \vee q)$  holds. However, at  $w_1$ ,  $!p$  does not hold because  $!p$  is only complied with at  $w_2$  and not at  $w_3$ , and  $!q$  does not hold because  $!q$  is only complied with at  $w_3$  and not at  $w_2$ , so  $!p \vee !q$  does not hold at  $w_1$ .

It may be noted, here, that  $!$  *sometimes* distributes over disjunction, but it does not *in general* do so. In particular,  $!(\phi \vee \psi) \rightarrow !\phi \vee !\psi$  holds when  $\phi$  and  $\psi$  are of the form  $!\phi$ . That is,  $!(\phi \vee \psi) \rightarrow !!\phi \vee !!\psi$  holds. However, this is not a problem, because it is just a special case of the general principle that any iteration of imperative operators is vacuous. The antecedent is equivalent to  $!\phi \vee !\psi$ , because each of the disjuncts is already an imperative, so adding an imperative operator to the whole thing is vacuous. The consequent is also equivalent to  $!\phi \vee !\psi$ , because each of the disjuncts has two imperative operators, so one can be removed from each of them. That is,  $!!\phi$  is equivalent to  $!\phi$ , and  $!!\psi$  is equivalent to  $!\psi$ , so  $!!\phi \vee !!\psi$  is equivalent to  $!\phi \vee !\psi$ . So, both the antecedent and the consequent are equivalent to  $!\phi \vee !\psi$ . It is important, though, that  $!$  does *not* in general distribute over disjunction.

This general principle that any iteration of imperative operators is vacuous is a desirable one, and was the motivation for making the accessibility relation *euclidean* (and for including the corresponding axiom,  $!_i\phi \rightarrow !_i\phi$ ). The intuitive idea is that once an imperative operator has changed a declarative sentence into an imperative one, it cannot be made any “more imperative.” We can (cautiously) appeal to natural language here. If we take  $p$  to be “you open the window,” then  $!p$  will be:

- (3) Let it be that you open the window! Or,
- (4) Open the window!

Adding a further  $!$  to this sentence would get us to  $!!p$ , which will be something like:

- (5) Let it be that let it be that you open the window! Or,
- (6) Let it be that open the window!

These sentences, (5) and (6), are both nonsensical. It is not just that they are ungrammatical, the problem with (5) and (6) is that there is nothing that they could be saying (over and above just “open the window!”) They cannot, for example, be saying that you are commanded *to command someone* to open the window, whether someone else or yourself, because  $!p$  is *not* “you command (someone) to open the window.” It is the (imperative) sentence “open the window!”

So, RPS meets all of the success criteria. It is general, conservative, adequate, and non-cognitivist about imperatives.

## 12.6 Imperative Permissions

I have already given three reasons to prefer Relational Preprescription Semantics (and KD45), over Parsons’ Functional Preprescription Semantics (and KDDc4). In KD45, we can make sense of the differences between  $!\neg\phi$  and  $\neg!\phi$ , and between  $!\phi\vee!\psi$  and  $!(\phi\vee\psi)$ , and we can have gaps in the imperatives that hold (that is, it can be that neither  $!\phi$  nor  $!\neg\phi$  hold). However, as a bonus, once we change from a functional to a relational accessibility relation, we also naturally get an account of permissions. When we move from functional to relational accessibility semantics, we get the dual of  $!$ :  $\text{j. } \text{j}\phi$  can be thought of as a permissive: something like “you are permitted to do  $\phi$ .” This explains what is going on when neither  $!\phi$  nor  $!\neg\phi$  is in force. When it is not obligatory to do  $\phi$ , it is permitted to do  $\neg\phi$ , and when it is not obligatory to do  $\neg\phi$ , it is permitted to do  $\phi$ . So, when neither “Kirsten, take your umbrella to work!” nor “Kirsten, don’t take your umbrella to work!” are in force, both “Kirsten, you are permitted to take your umbrella to work” and “Kirsten, you are permitted not to take your umbrella to work” are in force.

Perhaps it could be objected that, just as there is no commander to command Kirsten to take her umbrella, similarly there is no “permitter” to permit her to take it. However, the state of permission is very minimal; it is just the state of not-being-commanded. The point is *not* that nobody has commanded Kirsten to take her umbrella, (as similarly, nobody has permitted her to), the point is that it is genuinely *open* whether or not she takes it: neither option is *imperatively ideal*. Relational Preprescription Semantics can make sense of this, whereas Parsons’ Functional Preprescription Semantics cannot.

## 12.7 Logical Consequence

Now that we have an account of imperative consequence that is general, conservative, adequate, and non-cognitivist about imperatives, we can return to our original problem. Recall, the problem of imperative consequence is a trilemma:

**T1** Imperatives can be the relata of the consequence relation.

**T2** Imperatives are not truth-apt.

**T3** The relata of the consequence relation must be truth-apt.

My account of imperative consequence as RPS/KD45 is a version of solution 3, that is, I reject T3. I have given an account of consequence that applies to imperatives (that is, I claim that T1 is true), and it does not make imperatives truth-apt (that is, I also claim that T2 is true), so, the relata of the consequence relation do not have to be truth-apt. I ended chapter 1 with the account of the concept of logical consequence given by J.C. Beall and Greg Restall, the Generalised Tarski Thesis:

An argument is valid<sub>*x*</sub> if and only if in every case<sub>*x*</sub> in which the premises are true, so is the conclusion (Beall and Restall (2006): 29).

However, this conception defines logical consequence fundamentally in terms of (necessary) *truth*-preservation. So, how can we fix it? According to my proposal, consequence guarantees the preservation of *holdingness* rather than truth, where a declarative sentence holds at a world iff it is true there, and an imperative sentence holds iff it is complied with in all of that world's ideal worlds. Note, here, that this is *not* a merely disjunctive definition of holdingness. Pure declarative sentences and pure imperative sentences are just *special cases* of the general definition of holdingness. This definition is genuinely general, because it is really the contents of declaratives and imperatives that hold or fail to hold, and the contents of these different "types" of sentence are *all preproscriptions*. So, pure declarative sentences appear to hold in one way, and pure imperative sentences in another way (and, of course, mixed sentences do not hold in either of these simple ways, but in a more complex, two-dimensional way). Really, though, they appear to hold in different ways because they are special cases where we need only to assess them in one of the two dimensions.

So, with this more general conception of holdingness in mind, we can modify Beall and Restall's definition to:

### The More Generalised Tarski Thesis

An argument is valid<sub>*x*</sub> if and only if, in every case<sub>*x*</sub> in which the premises hold, the conclusion also holds.

Because “holds” reduces to “is true” in the case of declarative-only arguments, Beall and Restall’s conception, or any particular instance of their schema, is a special case of this More Generalised Tarski Thesis. However, the More Generalised Tarski Thesis *also* covers instances of consequence where some of the relata are imperatives.

## 12.8 The Frege-Geach Problem Revisited

Recall, from section 2.4, the problem of unasserted contexts for non-cognitivism about morality (the Frege-Geach problem). My logic of imperatives, Relational Preprescription Semantics, provides a solution to this problem for a form of prescriptivism. Recall that the problem arises when we consider moral arguments, such as:

- H1** Stealing is wrong.
- H2** If stealing is wrong, then so is committing copyright infringement.
- ∴ **H3** Committing copyright infringement is wrong.

RPS provides a solution for a version of prescriptivism. The prescriptivist can translate argument H into the following argument, which is ungrammatical in English, but (perhaps the prescriptivist could say) is meaningful nonetheless:

- I1** Don’t steal!
- I2** If don’t steal, then don’t commit copyright infringement!
- ∴ **I3** Don’t commit copyright infringement!

RPS is silent on whether I2 is the correct translation of H2, for the prescriptivist. RPS is not a prescriptivist theory of morality, because it is not a theory of morality at all (if the prescriptivist wants to translate H2 differently, they are welcome to, but in that case RPS will (probably) not help them solve the Frege-Geach problem). According to RPS, argument I is valid, because for I1 to hold, it must be that every compliance world has the addressee not stealing. If addressed to Ella, every compliance world has *Ella* not stealing, and if addressed to everyone in the world (as I take it the moral non-cognitivist would propose), every compliance world has nobody stealing. For I2 to hold, it must be that either there is some compliance world where the addressee steals, or all the compliance worlds have the

addressee not infringing copyright. Because I1 requires that there is no compliance world at which the addressee steals, the conclusion must also hold: there must be no compliance world where the addressee/s commit copyright infringement. My solution is similar to the norm-expressivism of Alan Gibbard (1990). Our theories each propose a possible worlds semantics on which the points of evaluation are pairs, where the first member of the pair is a possible world. The second member of the pair is what differs: on Gibbard's view, the second member is a "system of norms", while on my view, the second member is a set of possible worlds.

RPS cannot solve the Frege-Geach problem for every form of non-cognitivism about morality, because it is a two-part problem. The problem is: *how can the non-cognitivist plausibly translate moral sentences so that arguments with unasserted contexts (like argument H) are valid?* RPS gets the logic to work if the conditional moral sentence is translated in a particular way (and if we allow some ungrammatical sentences), but it is silent on the first (and, I think, more significant) part of the problem: determining how to translate these conditionals in a plausible way. This is a problem for the non-cognitivist about morality.

I have presented an account of imperative consequence in terms of Relational Preproscriptions. I proved that these semantics are equivalent to the accessibility semantics of the normal modal logic KD45, and thus the proof theory of KD45 is sound and complete with respect to my Relational Preproscription Semantics. This imperative logic is general, conservative, adequate, and non-cognitivist about imperatives. Finally, I presented a general conception of logical consequence as (necessarily) preserving holdingness, rather than truth.

# Conclusion

In this thesis, I presented the problem of imperative consequence as a modified version of Jørgensen's Dilemma. Specifically, I presented it as a trilemma:

**T1** Imperatives can be the relata of the consequence relation.

**T2** Imperatives are not truth-apt.

**T3** The relata of the consequence relation must be truth-apt.

I outlined and motivated each of these three claims, then outlined the three types of solution, which correspond, respectively, to the rejection of each of these three claims. First, I presented and reconstructed some arguments for eliminativism about imperative consequence, then I presented and discussed several versions of imperative cognitivism. Finally, I described and criticised several attempts at formal definitions of logical consequence.

In Part I, I discussed arguments for eliminativism about imperative consequence. I presented the argument from permissive presuppositions, proposed by Bernard Williams, which I claim fails for four reasons: because permissive presuppositions are analogous to conversational implicature, and so should be ignored when formulating imperative arguments; because there are other valid argument forms, which Williams does not consider, that do not suffer from different permissive presuppositions; because there are explanations for the change in permissive presuppositions in Williams' example other than a change of mind, so the different permissive presuppositions do not have to prevent accumulation; and because Williams accepts enough logical relations between imperatives (contradiction and a form of negation) for a definition of a valid imperative inference to follow naturally.

Next, I discussed a family of arguments for eliminativism about imperative consequence that all appeal to the rules of grammar. These were originally proposed by Gary Wedeking, Jörg Hansen, and Jonathan Harrison. I outlined the argument from premise- and conclusion-indicating words, the argument from motivating reasons, and the argument from analytic

conditionals. I pointed out why each of these arguments fail, and showed that the first and third of these arguments both appeal to instances of the same phenomenon: that an imperative cannot appear in a subordinate clause.

In Part II, I presented five versions of imperative cognitivism, and proposed four problems for them. All of the versions of cognitivism (the Reports theory, the Desires theory, the Deontic theory, the Predictions theory, and the Elliptic theory) fail some combination of the four problems: the problem of unwanted consistencies, the problem of unwanted validities, the problem of soft imperatives, and the problem of disjunctive threats. They are, then, inadequate as translations of imperatives.

In Parts III and IV, I discussed several proposals for formal definitions of logical consequence. These all outline formal logical systems that include imperatives. In Part III, I outlined and discussed several versions of imperative logic that are based on first-order logic. First, I looked at some early attempts, the neustic-phrastic analysis of R. M. Hare, and the operator analysis of Albert Hofstadter and J. C. C. McKinsey. I argued that the rules proposed by these theories have glaring and serious counterexamples, so they are unsatisfactory. Additionally, Hare's account is not fully developed, and Hofstadter and McKinsey's account collapses into a form of cognitivism.

Next, I outlined and discussed some formal definitions of consequence based on obedience-values, or termination-values as an analogue of truth-values. I discussed a version proposed by Charles Pigden (based on that of Jack Smart), and a version proposed by Nicholas Rescher. I pointed out some counterexamples to the definitions proposed, tried to fix them, and showed that fixing them just leads to different counterexamples. I then also argued that the two theories both suffer from two more fundamental, theoretical, problems. These two accounts require us to partition language into "imperative" and "indicative," without allowing any mixed sentences, but this appears to be impossible to do in a satisfying, principled way. Also, both of these theories have the result that consequence is a merely disjunctive concept: that there are really two distinct and *different* relations of "consequence."

Next, I outlined and discussed a version of imperative logic proposed by Peter Vranas, that we should take that analogue of truth, not as obedience, but as *reasons for obeying* the imperative. I argued that Vranas's definition relies on concepts that are not at all well defined, but insofar as they are defined, they produce several implausible results. Vranas' theory also inherits extensive problems, because it relies on truthmaker theory, as well as the problematic notion of reasons for acting. His theory also suffers from the theoretical problem that consequence is merely disjunctive: that there are two different things that are

both called “consequence.”

In Part IV, I turned to formal systems that are based on modal logics. First, I outlined and discussed the formal system of Brian Chellas, which introduces a relation of imperative alternativeness, and a concept of imperatively ideal worlds. I showed that his system is (a stronger version of) KD45: the normal modal logic K with the addition of axioms D, 4, and 5. However, I pointed out that this theory has no account of how it is that imperatives are not truth-apt. It assigns the same two values to imperatives and declaratives alike, with no explanation of how these values are anything more or different than truth-values. It is, thus, not non-cognitivist about imperatives.

Then, I outlined and discussed the theory of Functional Preposcription Semantics, proposed by Josh Parsons. This theory gives an account of the contents of sentences, not as propositions (sets of possible worlds), but as two-dimensional functional preproscriptions (sets of pairs of possible worlds). These semantics, it turns out, are equivalent to KDDc4: the normal modal logic K with the additional axioms D, Dc, and 4. However, this account, I claimed, is inadequate as it suffers from a series of important counterexamples.

Finally, I outlined my proposed formal logic of imperatives. I pointed out that the counterexamples to Parsons’ theory can all be fixed by removing the constraint of functionality, and instead making the preproscriptions *relational*. That is, the contents of imperative and declarative sentences are sets of pairs, where the first member is a world and the second member is a *set of worlds*. I proved that these Relational Preproscription Semantics are equivalent to the accessibility semantics of KD45 and consequently that the proof theory of KD45 is sound and complete with respect to my Relational Preproscription Semantics. Now, then, we have an imperative logic that is general, conservative, adequate, and also non-cognitivist about imperatives. Instead of propositions being true or false, we can now talk of preproscriptions *holding* and *failing to hold*. I explained how imperatives, as well as declaratives, have preproscriptions as their contents, so imperatives as well as declaratives can hold or fail to hold. I concluded that logical consequence should be defined, not in terms of truth, but in terms of *holding*.

# Bibliography

- A. R. Anderson and O. K. Moore. The formal analysis of normative concepts. *American Sociological Review*, 22(1):9–17, Feb 1957.
- Aristotle. Prior analytics. In Jonathan Barnes, editor, *The Complete Works of Aristotle, vol. 1, translated by A. J. Jenkinson*. Oxford University Press, 1984.
- J.C. Beall and Greg Restall. *Logical Pluralism*. Oxford University Press, Oxford, 2006.
- H. G. Bohnert. The semiotic status of commands. *Philosophy of Science*, 12(4):302–315, Oct 1945.
- Bernard Bolzano. *Wissenschaftslehre*. Sulzbach, 1837.
- Alexander Broadie. *Introduction to Medieval Logic*. Clarendon Press, Oxford, 1993.
- Rudolf Carnap. *Logical Syntax of Language*. Humanities, New York, 1937.
- Hector-Neri Castañeda. There are command sh-inferences. *Analysis*, 32(1):13–19, Oct 1971.
- Brian F. Chellas. Imperatives. *Theoria*, 37(2):114–129, 1971.
- Brian F. Chellas. *Modal Logic: An Introduction*. Cambridge University Press, 1980.
- Michael Dummett. *Elements of Intuitionism*. Oxford University Press, Oxford, 1977.
- John Etchemendy. *The Concept of Logical Consequence*. Harvard University Press, Cambridge, MA, 1990.
- Melvin Fitting and Richard L. Mendelsohn. *First-Order Modal Logic*, volume 277 of *Studies in Epistemology, Logic, Methodology, and Philosophy of Science*. Kluwer Academic Publishers, 1998.

- P. T. Geach. Imperative and deontic logic. *Analysis*, 18(3):49–56, Jan 1958.
- P. T. Geach. Ascriptivism. *Philosophical Review*, 69, 1960.
- P. T. Geach. Imperative inference II. *Analysis Supplement*, 23:37–42, 1963.
- P. T. Geach. Assertion. *Philosophical Review*, 74, 1965.
- Alan Gibbard. *Wise Choices, Apt Feelings*. Clarendon Press, Oxford, 1990.
- P. C. Gibbons. Imperatives and indicatives (I). *Australasian Journal of Philosophy*, 38(2): 107–119, 1960.
- Lou Goble. *The Blackwell Guide to Philosophical Logic*. Blackwell Publishers, Malden, 2001.
- H. P. Grice. *Logic and Conversation*. Harvard University Press, 1989.
- C. L. Hamblin. *Imperatives*. Blackwell Publishing Inc, Oxford, 1987.
- Jörg Hansen. *Imperatives and Deontic Logic: On the Semantic Foundations of Deontic Logic*. PhD thesis, 2008.
- R. M. Hare. *The Language of Morals*. Clarendon Press, Oxford, 1952.
- Jonathan Harrison. Deontic logic and imperative logic. In P. T. Geach, editor, *Logic and Ethics*. Kluwer Academic Publishers, The Netherlands, 1991.
- A. Hofstadter and J.C.C. McKinsey. On the logic of imperatives. *Philosophy of Science*, 6 (4):446–457, Oct 1939.
- Frank Jackson and Philip Pettit. A problem for expressivism. *Analysis*, 58:239–251, 1998.
- Jörgen Jörgensen. Imperatives and logic. *Erkenntnis*, 7(1):288–296, 1937.
- Immanuel Kant. *Fundamental Principles of the Metaphysic of Ethics, Translated by Thomas Kingsmill Abbott*. Longmans, London, 1962.
- H. S. Leonard. Interrogatives, imperatives, truth, falsity and lies. *Philosophy of Science*, 26 (3):172–186, Jul 1959.
- David Lewis. General semantics. In *Philosophical Papers*, volume 1. Oxford University Press, 1983.

- Ernst Mally. *Grundgesetze des Sollens: Elemente der Logik des Willens*. Leuschner und Lubensky, Universitäts-Buchhandlung, Graz, 1926.
- Joan Moschovakis. Intuitionistic logic. *Stanford Encyclopedia of Philosophy*, 2004.
- Josh Parsons. Negative truths from positive facts. *Australasian Journal of Philosophy*, 84(4):591–602, Dec 2006.
- Josh Parsons. Cognitivism about imperatives. *Analysis*, 72(1):49–54, 2012.
- Josh Parsons. Command and consequence. *Philosophical Studies*, 164(1):61–92, 2013.
- Josh Parsons. Preprescription semantics and KDDc4. Unpublished manuscript, 2014.
- Gillman Payette and Peter Schotch. Preserving what? In Peter Schotch, Bryson Brown, and Raymond Jennings, editors, *On Preserving*, chapter 6, pages 85–104. University of Toronto Press, 2009.
- Charles Pigden. Is George a traitor? Geach’s problem and imperative consequence. Unpublished manuscript.
- Charles Pigden. Logic and the autonomy of ethics. *Australasian Journal of Philosophy*, 67:127–151, 1989.
- Charles Pigden. *Hume on Motivation and Virtue*. Palgrave Macmillan, 2010.
- Henri Poincaré. *Dernières Pensées*. E. Flammarion, Paris, 1917.
- K. R. Popper. What can logic do for philosophy? *Proceedings of the Aristotelian Society, Supplementary Volumes*, 22:141–154, 1948.
- Nicholas Rescher. *The Logic of Commands*. Routledge and Kegan Paul, London, 1966.
- Gideon Rosen. Modal fictionalism. *Mind*, 99(395):327–354, 1990.
- Alf Ross. Imperatives and logic. *Philosophy of Science*, 11(1):30–46, Jan 1944.
- J. J. C. Smart. *Ethics, Persuasion and Truth*. Routledge and Kegan Paul, Boston, MA, 1984.
- R. Stalnaker. Assertion. In *Context and Content: Essays on Intentionality in Speech and Thought*, pages 78–95. Oxford University Press, Oxford, 1999.

- Alfred Tarski. *Logic, Semantics, Metamathematics (2nd ed.)*. Hackett Publishing Company, 1983.
- G. H. von Wright. Deontic logic. *Mind*, 60:1–15, 1951.
- G. H. von Wright. *An Essay in Deontic Logic and the General Theory of Action*. North Holland Publishing Company, Amsterdam, 1968.
- Peter B. M. Vranas. New foundations for imperative logic I: Logical connectives, consistency, and quantifiers. *Noûs*, 42(4):529–572, 2008.
- Peter B. M. Vranas. In defense of imperative inference. *Journal of Philosophical Logic*, 39: 59–71, 2010.
- Peter B. M. Vranas. New foundations for imperative logic II: Pure imperative inference. *Mind*, 120(478):369–446, 2011.
- Peter B. M. Vranas. New foundations for imperative logic III: A general definition of argument validity. Unpublished manuscript, 2012.
- Peter B. M. Vranas. New foundations for imperative logic IV: Natural deduction. Unpublished manuscript, 2013.
- Gary Wedeking. Are there command arguments? *Analysis*, 30(5):161–166, Apr 1970.
- Bernard Williams. Imperative inference. *Analysis Supplement*, 23:30–36, 1963.