Synthese Library 458
Studies in Epistemology, Logic, Methodology, and Philosophy of Science

Benjamin Alan Burgis

Logic Without Gaps or Gluts

How to Solve the Paradoxes Without Sacrificing Classical Logic

Springer

Synthese Library

Studies in Epistemology, Logic, Methodology, and Philosophy of Science

Volume 458

Editor-in-Chief

Otávio Bueno, Department of Philosophy, University of Miami, Coral Gables, USA

Editorial Board Members

Berit Brogaard, University of Miami, Coral Gables, USA
Anjan Chakravartty, Department of Philosophy, University of Miami, Coral Gables, USA
Steven French, University of Leeds, Leeds, UK
Catarina Dutilh Novaes, VU Amsterdam, Amsterdam, The Netherlands
Darrell P. Rowbottom, Department of Philosophy, Lingnan University, Tuen Mun, Hong Kong
Emma Ruttkamp, Department of Philosophy, University of South Africa, Pretoria, South Africa
Kristie Miller, Department of Philosophy, Centre for Time, University of Sydney, Sydney, Australia

The aim of *Synthese Library* is to provide a forum for the best current work in the methodology and philosophy of science and in epistemology, all broadly understood. A wide variety of different approaches have traditionally been represented in the Library, and every effort is made to maintain this variety, not for its own sake, but because we believe that there are many fruitful and illuminating approaches to the philosophy of science and related disciplines.

Special attention is paid to methodological studies which illustrate the interplay of empirical and philosophical viewpoints and to contributions to the formal (logical, set-theoretical, mathematical, information-theoretical, decision-theoretical, etc.) methodology of empirical sciences. Likewise, the applications of logical methods to epistemology as well as philosophically and methodologically relevant studies in logic are strongly encouraged. The emphasis on logic will be tempered by interest in the psychological, historical, and sociological aspects of science. In addition to monographs *Synthese Library* publishes thematically unified anthologies and edited volumes with a well-defined topical focus inside the aim and scope of the book series. The contributions in the volumes are expected to be focused and structurally organized in accordance with the central theme(s), and should be tied together by an extensive editorial introduction or set of introductions if the volume is divided into parts. An extensive bibliography and index are mandatory.

More information about this series at https://link.springer.com/bookseries/6607

Benjamin Alan Burgis

Logic Without Gaps or Gluts

How to Solve the Paradoxes Without Sacrificing Classical Logic

Springer

Benjamin Alan Burgis
Department of Philosophy
Rutgers University
Atlanta, GA, USA

ISSN 0166-6991 ISSN 2542-8292 (electronic)
Synthese Library
ISBN 978-3-030-94623-4 ISBN 978-3-030-94624-1 (eBook)
https://doi.org/10.1007/978-3-030-94624-1

© Springer Nature Switzerland AG 2022
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Contents

1 Logic and the Liar Paradox 1
- 1.1 The T-Schema and the Laws of Logic 1
- 1.2 Capture and Release 3
- 1.3 Paraconsistent and Paracomplete Options 4
- 1.4 Comparing Dialetheism to Paracompleteness 7
- 1.5 Varieties of Gaps and Gluts 9
- 1.6 Once More on the Comparative Plausibility of Restricting the T-Schema and Rejecting Classical Logic 12
- 1.7 Summary of the Chapters 15

2 Graham Priest and Dialetheism 17
- 2.1 Three Priests Walk Into A Bar: What's Funny About Dialetheism 17
- 2.2 The Paradoxes of Motion and Change 18
- 2.3 Russell's Paradox and the Inclosure Schema 20
- 2.4 A Bishop and a Priest Walk Into a Bar. . . (Berkeley's Master Argument) 25
- 2.5 Inconsistent Laws and Norms 27

3 How to Solve Priest's Paradoxes Without Sacrificing Classical Logic 31
- 3.1 Normative Dilemmas Without Normative Contradictions 31
- 3.2 How to Solve the Paradoxes of Naïve Set Theory 34
- 3.3 How to Solve the Paradoxes of Motion and Change 38
- 3.4 Where Berkeley and Priest (and Max Black!) Go Wrong 44
- 3.5 What Has Not Been Solved 48

4 Dialetheism and the Laws of Logic 49
- 4.1 Dialetheism and Validity 49
- 4.2 Dialetheism and the Status of *Reductio Ad Absurdum* and *Modus Tollens* 53

4.3 Alan Weir on Dialetheism and Falsity Preservation 57
4.4 Why Successful Classical Recapture Is Vital to the Dialetheic Project 59
4.5 The Failure of Classical Recapture, Part I (The Failure of Principle R) 61

5 Dialetheism, Rejection, and Curry's Paradox 65
5.1 The Broader Importance of the Failure of Principle R (I: Curry's Paradox) 65
5.2 Why Priest Rejected Principle R (and JC Beall Should Too) 73
5.3 Is the *Classical* Logician Entitled to Principle R? 83
5.4 Re-Examining the Relationship Between Dialetheism and (Even "Informal") *Reductio* Arguments in Light of the Failure of Principle R 88
5.5 The Broader Importance of the Failure of Principle R (II: Rejection-Talk in Beall and Field) 92
5.6 Francesco Berto's Critique of Dialetheist Rejection-Talk 99

6 Dialetheism, Rejection, and Probability Theory 103
6.1 Probability and Quasi-Validity 103
6.2 Dialetheism and the Consistency of the Empirical World 111
6.3 Priest's Case for the Consistency of the Empirical World, Part I (The Argument from Pierre and the Conjunction Principle) 112
6.4 Priest's Case for the Consistency of the Empirical World, Part II (The Argument from Perceptual Illusions) 115
6.5 Why LPm, Logical Pluralism, and Local Validity Cannot Save Us 118
6.6 The Failure of Classical Recapture, Part II (The Failure of the Probabilistic Strategy) 120
6.7 Francesco Berto on Negation, Exclusion, and NOT 121

7 Hartry Field and Paracompleteness 125
7.1 Paracompleteness and Truth Preservation 125
7.2 Paracompleteness and Ordinary Reasoning 129
7.3 Paracompleteness and Probability 131
7.4 Paracompleteness and Rejection 134
7.5 Paracompleteness and Revenge 135
7.6 Paracompleteness and Hierarchies: Field Vs. Tarski and Maudlin 137
7.7 Paracompleteness and the Exhaustive Characterization Project 141
7.8 Paracompleteness and Properties: Field Solution to Russell's Paradox and Mine 144

8 How to Solve the Liar Paradox Without Sacrificing Classical Logic 147
8.1 Some Preliminary Points About Propositions 147
8.2 More Preliminaries: Disquotationalism and Inferentialism 151
8.3 How To Cash Out the Disquotation Analogy and Solve the Liar Paradox in the Bargain: Content Inheritance Disquotationalism 156
8.4 Comparing CID With Other Disquotationalisms 159
8.5 Objections and Replies 160
8.6 Comparing the Solutions: The Principle of Uniform Solution 173
8.7 Conclusion: Vagueness, Quantum Weirdness, and Other Loose Ends 174

References 181

Index 185

Chapter 1
Logic and the Liar Paradox



Abstract The Liar and related paradoxes create a prima facie conflict between two equally important principles of classical logic—the Law of the Excluded Middle (LEM) and the Law of Non-Contradiction (LNC). Dialetheists such as Graham Priest and JC Beall solve the paradoxes by accepting the negations of some instances of the LNC and rejecting classical logic in favor of a "paraconsistent" logical framework in which contradictions do not entail every conclusion. Hartry Field advocates the mirror image of this solution, rejecting certain instances of the LEM and thus rejecting classical logic in favor of a "paracomplete" framework in which not every LEM instance is entailed by every premise. Some paradox-solvers, eager to avoid either of these amputations of classical logic, reject paradoxical instances of the T-Schema. This is an easy way out, but the search for an adequate solution is best conducted elsewhere. The rest of the book will be spent on the project of resisting the paraconsistent and paracomplete solutions to the paradoxes and finally proposing a positive solution that retains classical logic without sacrificing the T-Schema.

Keywords Paradoxes · Revenge · Dialetheism · Paraconsistent · Paracomplete · T-Schema · Bivalence

1.1 The T-Schema and the Laws of Logic

The Liar Paradox looks like a conflict between two equally fundamental principles of classical logic. Take the following sentence:

1. Sentence (1) is not true.

Let us go through the options.

(i) (1) is true
(ii) (1) is not true
(iii) (1) is both true and not true.

The Law of the Excluded Middle (LEM) tells us that either (i) or (ii) must be the case. According to familiar "Liar reasoning," each of them separately entail (iii). . . . which is ruled out by the Law of Non-Contradiction (LNC).

© Springer Nature Switzerland AG 2022

B. A. Burgis, *Logic Without Gaps or Gluts*, Synthese Library 458,
https://doi.org/10.1007/978-3-030-94624-1_1

To fully appreciate the problem all of this poses for classical orthodoxy, let us take a moment to unpack that "familiar reasoning." Tarski's Biconditional Truth Schema (henceforth, "the T-Schema") tells us that for every proposition α, $Tr<\alpha> \leftrightarrow (\alpha)$. Since (1) just *is* $\neg Tr<L>$, when we feed it into the T-Schema, we get $Tr<L> \leftrightarrow \neg Tr<L>$. There are multiple roads from here to contradiction—the inference is clear enough that classical logicians often talk loosely and say that a sentence of the form $\alpha \leftrightarrow \neg\alpha$ "is" a contradiction—but all of them involve the LEM.

The most obvious way to solve the paradox without sacrificing classical logic would be to reject (some instances of) the T-Schema, so that we could assert either (i) or (ii) without being stuck with (iii). Gilbert Harman takes this line in his 1986 classic *Change of View*.

> To see that this schema is indeed the culprit, notice that one instance of it is
>
> "(L) is not true" is true if and only if (L) is not true.
>
> Since (L) = "(L) is not true," this instance is equivalent to the self-contradictory
>
> (L) is true if and only if (L) is not true
>
> [. . .] In practice the best solution may be to retain the Biconditional Truth Schema and yet avoid contradiction by interpreting the Schema not as something that holds without exception but rather as something that holds "normally" or "other things being equal." It is then a "default assumption." One accepts any given instance of the Biconditional Truth Schema in the absence of a sufficiently strong reason not to accept it (Harman, 1986, p. 16).

How plausible one finds this proposal might depend on one's theory of truth. If the T-Schema is part of (or even follows from) the definition of "true," then there could no more be exceptions to it than there could be the occasional three-sided exception to the "default assumption" that squares have four sides.

On the other hand, there are views of truth according to which the T-Schema and the definition of "true" are logically independent. For example, it could be that when we call a proposition "true," we are ascribing to it some property such as correspondence with reality (where "correspondence" is not defined by use of the T-Schema) or perhaps coherence with a set of propositions epistemic agents would hold under some set of idealized circumstances. If so, just as the typical properties of bears ground our rational fear of encountering them in the wild, general facts about correspondence or coherence or whatever might ground the legitimacy of (most instances of) the T-Schema. Just as we know better than to apply the Maxim of the Dangerousness of Bears to old toothless bears, we should know not to apply the T-Schema to "defective" cases such as (1). Of course, to even begin to make a solution along these lines sound plausible, a great many details would need to be filled in, but so far it looks like such a story *could* be told.

Interestingly enough, in his book *Truth*, Paul Horwich advocates a version of Harman's view that *couldn't* be backed up by any such story. He argues for a version of deflationism about truth according to which the T-Schema is literally all there is to be said about the meaning of "true." His preferred formulation is that the "axioms" of his preferred theory of truth ("the minimal theory") are just an infinite list of instances of the T-Schema. Nevertheless, in the section of *Truth* in which he deals

with the paradoxes, he concedes that "certain instances of the equivalence schema are not to be included as axioms of the minimal theory. . ." (Horwich, 1998, p. 40) This is a bit like saying that the definition of a square consists in an infinite list of instantiations of the schema "[Name of geometrical object] is a square iff it is a plane figure with four equal sides and four right angles," while also saying that some three-sided figures also count as squares, so of course those schema instances are not part of the definition of "square"... and leaving it at that.

1.2 Capture and Release

The counter-intuitiveness of the Harman/Horwich family of solutions comes into sharper focus when we switch from talking about the T-Schema to talking about the Capture and Release rules.

Capture: α // Tr<α>
Release: Tr<α> // α

The two ways of talking are equivalent given Harman and Horwich's embrace of classical logic. (Remember, preserving classical orthodoxy is much of the point of such solutions.) Thus, embracing the Harman/Horwich solution means thinking that there are at least some counterexamples to Capture and/or Release—cases in which an ideal epistemic agent who knew all the facts of the case would accept the premise of a Capture or Release inference and reject the conclusion. When it comes to ordinary everyday values of α, we would take such a combination of acceptance and rejection to be deeply confused. Imagine the following two exchanges.

Snow is White, Take One:

Duncan: Snow is white.
Richie: I disagree.
Tessa: Really? You disagree? Have you ever seen snow?

Snow is White, Take Two:

Duncan: Snow is white.
Tessa: That's true.
Richie: I disagree.
Tessa: You disagree that snow is white?
Richie: No, I disagree with what you said, not with what Ducan said. Of course, I agree that snow is white. I just don't think it's true that "snow is white."

Presumably, Duncan and Tessa would be surprised at the end of both versions of the conversation, but they would be surprised in different ways. At the end of Take One, they would be surprised that Richie was not aware of an empirical fact with

which they would expect him to be familiar. In the Take Two version of the conversation, they would think he did not understand the word "true."[1]

One of the basic assumptions in everything that follows—an assumption that I tend to share with the theorists I will be arguing against—is that if someone claims to accept some statement α and reject Tr<α > or vice versa, this should be taken as sufficient evidence that the speaker is talking about something other than truth, just as, if Richie agreed that some object only had three sides but still insisted that it was a square, we would take him to be confused not about the object but about the meaning of "square." A "solution" that uses the word "truth" to talk about something else is not so much a solution as a change of subject.

But if we cannot get around the paradox by sacrificing instances of the T-Schema, where *should* we look for a solution? After a quick survey of the options, even some readers initially inclined to side with me on the T-Schema question might be willing to re-evaluate their initial intuition. After all, the inferential rules questioned by non-classical paradox-solvers might seem to be even more obviously valid than Capture and Release. We will return to this question later in the chapter. Meanwhile, let us take a look at some of the major non-classical solutions.

1.3 Paraconsistent and Paracomplete Options

In classical logic, contradictions entail everything. This follows trivially from the combination of the LNC with the traditional claim that *validity is truth-preservation*.[2] Dialetheists such as Graham Priest and JC Beall accept the LEM but reject the

[1] I have seen Graham Priest respond to a similar example by saying that they would conclude not that he did not know what the word meant but that he was a philosopher with non-standard views about truth. My only response would be that the fact that these would seem to be the only options says something significant about the clash between those non-standard views and the standard intuitions of competent speakers.

[2] When I say "the LNC," I mean the sense of "the LNC" in which "dialetheists reject the LNC" is true. See below. It should also be noted that Graham Priest suggests in sections 4.6 and 4.7 of *Doubt Truth to be a Liar* that the "explosion" of consequences derivable from contradictions in classical logic could be understood in a different way. He starts his discussion at the end of 4.6 by arguing that the Liar Paradox shows that "truth and falsity overlap" and thus that there are counterexamples to the claim that the inference of arbitrary consequences from contradictions is universally truth-preserving. Since "truth-preservation is at least a necessary condition for validity," he takes this to show that explosion is invalid (Priest, 2006a, p. 84). In 4.7, he suggests that a partisan of classical orthodoxy could defend the explosion principle by building it into the definition of negation. He responds that a logical constant defined in this way (such that ¬α was defined as α→⊥, "if A, then everything") "is not negation" on the grounds that the version of the LEM it yielded would be implausible (Priest, 2006a, p. 85). I have no interest in disputing this last point. I disagree with Priest about whether explosion is valid because (and only because) I disagree with him about whether it universally preserves truth. My position is that ¬α and α→⊥ are equivalent *given that explosion is valid*, but that α→⊥ should not be taken as any sort of analysis or definition of ¬α. Similarly, Priest suggests in section 1.13 of the same book that the account of negation built into "orthodox modern

LNC (or at least accept that it has true counterexamples). They solve the Liar and similar paradoxes by accepting that the resulting contradictions are actually true and rejecting classical logic in favor of an inconsistency-tolerant (or "paraconsistent") logical framework in which arbitrary conclusions cannot be derived from every contradiction. "Paracomplete" paradox-solvers such as Hartry Field make the mirror image of this move, holding onto the LNC by rejecting the LEM and all the laws of classical logic that implicitly rest on it.[3]

Before considering the more substantive issues, an important technical clarification is in order. I have said that dialetheists "reject the LNC," and I will continue to talk this way for the rest of the book, because it is convenient to do so, because dialetheists themselves generally talk that way, because it highlights the intuitive parallel between the dialetheist and paracomplete positions, and because it avoids the need for a lot of clumsy repetition. However, there is an important sense in which dialetheists *don't* reject the LNC.

If the LNC is just the universal generalization that for every α, $\neg(\alpha \wedge \neg\alpha)$, dialetheists are actually committed (by logical assumptions that lie in the overlap between classical logic and the various paraconsistent systems put forward by dialetheist thinkers such as Priest and Beall) to accepting the LNC. To see why, start with some contradiction $(\alpha \wedge \neg\alpha)$ that the dialetheist takes to be true:

1. $\alpha \wedge \neg\alpha$	Premise
2. α	1 Conjunction Elimination
3. $\alpha \vee \neg\alpha$	2 Disjunction Addition
4. $\neg\neg\alpha \vee \neg\alpha$	3 Double Negation
5. $\neg(\alpha \wedge \neg\alpha)$	4 DeMorgan's Laws[4]

Thus, at least assuming the validity of the four logical rules appealed to above—none of which are obviously invalidated by the mere assumption of dialtheism—

logics, such as 'classical' and intuitionistic logic" is such that for every pair of statements α and $\neg\alpha$, the content of $\neg\alpha$ is the complement of the content of α (Priest, 2006a, p. 31). In other words, $\neg\alpha$ simply has whatever content α does not. But he provides no evidence for the claim that classical logic rests on such an account of negation. He simply presents it as an unargued assertion. An obvious reason for skepticism is that, as an explanation of the explosion principle, the attribution of the complementation account of negation to classical logic is unparsimonious. If you start with the orthodox views that validity is truth-preservation and that contradictions are never true, you get explosion for free. Given this, there is no reason to suspect that it is *also* grounded in any sort of controversial semantic or metalinguistic assumption built into the logic.

[3] As Field often emphasizes, rejecting the LEM does not mean rejecting all of its instances. Field and other paracomplete theorists typical reject LEM instances related to paradoxical sentences and perhaps also borderline applications of vague predicates and accept nearly all of the instances we are most likely to encounter in scientific and "everyday" reasoning. Similar points apply mutatis mutandis to dialetheism and counterexamples to the LNC.

[4] Since Priest and perhaps most other dialetheists accept the LEM, the proof could have started on Line 3, but it is interesting to show that the proof goes through even without that assumption. Dialetheim could in principle be combined with LNC-rejection but to be a dialetheist just *is* to affirm the truth of some contradictions. Of course, given the dialetheist's assumptions, there is no conflict between saying that every instance of the LNC is true and that some of them are also false.

dialetheists are committed by their own lights to saying that false instances of the LNC are *also true*. Needless to say, this goes "all the way up." (In other words, nothing changes if α is itself an instance of the LNC the dialetheist takes to be false.) In a sense, then, the dialetheist *accepts* the LNC.

A form of the LNC some dialetheists might be tempted to reject is the modal version expressed in natural language by saying that "contradictions can't be true." Things are a bit more complicated here, because on the assumption that $\neg(\alpha \wedge \neg\alpha)$ is a theorem (i.e., it is *logically necessary* that it is true)—which Priest at least accepts for the reasons I have indicated above—we can get to the claim that contradictions are impossible in a couple of easy steps.

1. $\Box\neg(\alpha \wedge \neg\alpha)$ Premise
2. $\neg\Diamond\neg\neg(\alpha \wedge \neg\alpha)$ 1 Intersubstitutivity of $\Box$ and $\neg\Diamond\neg$
3. $\neg\Diamond(\alpha \wedge \neg\alpha)$ 2 Double Negation

At this point the question of whether a dialetheist can resist the conclusion that "contradictions can't be true" would seem to rest on the question of whether "$\neg\Diamond\alpha$" and "α can't be true" are equivalent. A dialetheist *could* deny that, just as a dialetheist *could* deny Double Negation or for that matter the intersubstitutivity rule for the modal operators, but, again, none of these denials would seem to flow in any sort of obvious way from dialetheism itself.

Even so, there is clearly a sense in which it makes perfect sense for dialetheists to say that they "reject the LNC." JC Beall clarifies this nicely in his book *Spandrels of Truth*.

> "The" so-called law of (non-) contradiction is not one but many. One version I reject is the following.
>
> >> Rationality LNC (RLNC): It is irrational (or rationally incoherent, or rationally absurd, or etc.) to knowingly accept that a sentence and its negation are ttrue[5] (Beall, 2009, p. 101).

Beall's RLNC plausibly captures at least one of the things that both dialetheists and non-dialetheists seem to mean when they talk about "dialetheists rejecting the LNC," and even when they talk about the role(s) that might be played by "the LNC" in ordinary reasoning. Importantly, no dialetheist that I am aware of asserts that the RLNC is true.

I do not claim that this is the *only* version of the LNC dialetheists have good reason to reject. Depending on a given dialetheist's position on some of the secondary issues flagged above—e.g., the relationship between "$\neg\Diamond\alpha$" and "α can't be true"—some dialetheists might even wish to reject one or more of the other versions of the LNC discussed even in this extremely brief and partial survey. I certainly would not want to suggest that I have ruled that out. There are also various other possibilities, such as "no formula has both the value 1 and the value 0" (which we

[5] In context, "ttrue" can be read just as "true." Beall uses the phrase "transparent truth" (ttruth) because the notion of "transparency" is central to his conception of truth. This is defined in terms of the unrestricted Capture and Release rules discussed above.

could call the Truth Value LNC), which might turn out to be dialetheically separable from the forms of the LNC that I have suggested that the dialetheist has a reason to accept. Whether it really is separable depends on whether the dialetheist's third truth value can be properly understood as a straightforward conjunction of the first two—an issue we will return to in different guises later in the book.

Happily, none of the criticisms of dialetheism in this book rely on fine-grained disambiguations of the sentence *Dialetheists reject the LNC*, so we can safely leave it coarse-grained. Anything I say about "the debate about the LNC" can be read as "the debate about those versions of 'the LNC', such as Beall's RLNC, which dialetheists typically reject."

1.4 Comparing Dialetheism to Paracompleteness

I said in Sect. 1.1 that the LEM and the LNC were equally fundamental principles of classical logic. This might be doubted. While no statistically meaningful research has been done on the subtle variations in analytic philosophers' attitudes toward various laws of logic—and it might be a bit much to hope for the funding for such a study to come around any time soon—my strong anecdotal impression is that quite a few philosophers who would never accept the notion of *true contradictions* are at least open to the idea that there might be exceptions to the LEM. (This open-mindedness tends to owe less to the Liar Paradox than to subjects such as borderline applications of vague predicates or even the mysteries of quantum mechanics. These topics will be briefly addressed in Sect. 8.7.) While the LEM and LNC are "equally fundamental" in the strict sense that instances of Excluded Middle are logically equivalent to negations of contradictions in classical logic (and any other logic that contains both DeMorgan and Double Negation laws), the idea of accepting a logical framework in which LEM sometimes fails seems to be less shocking to the average philosopher than dialetheism.

Why the intuitive asymmetry? There are at least two plausible (and complementary) explanations. The first has to do with the routine incompleteness of human knowledge. The second is that the LNC plays a more obvious role in our reasoning practices than the LEM.

Let us start with the second point and work our way back. When we derive a contradiction from a theory, we normally take that as decisive evidence that the theory is *false*. The incompleteness of a theory, on the other hand, is not normally taken to be any great strike against it. If all Russell had been able to show about naïve set theory was that it did not contain the resources with which to answer the question of whether the set of all sets that are not members of themselves was a member of itself or not, set theorists the world over would have shrugged. A few might have started finding a way to prove that it was (or was not) self-membered, but no one would have taken their inability to answer the question as evidence that there was no such set, and that the unrestricted comprehension axiom of naïve set theory must therefore be false. No one expects specialists in any area of human knowledge to be

omniscient even about their area of expertise. Of course, we also do not expect them to be entirely free of false beliefs any more than we expect them to have a God-like awareness of all truths, but we *do* expect them to correct falsehoods that have crept into their belief sets once they become aware of them. One of the easiest ways to prove that something is false is to show that it is inconsistent. That is why Russell's Paradox (and Burali-Forti and the rest) had such a profound impact on the history of set theory, and why Bishop Berkeley was so interested in deriving a contradiction from the notion of a mind-independent material world, and why prosecutors are so eager to find inconsistencies in defendants' alibis. The failure of a mathematical or metaphysical theory or a criminal alibi to include information one way or the other on some issue simply does not sink it in the same way.

As such, it is psychologically unsurprising that a paraconsistent solution to the paradoxes strikes many philosophers as a more radical departure from classical orthodoxy than a paracomplete one. None of this, however, adds up to any particularly weighty reason to actually prefer paracompleteness. To see why not, start by noticing that the LEM plays an important role in many pieces of reasoning we normally associate with the LNC. For example, to generate a contradiction about the set of all sets that are not members of themselves, we need the assumption that such a set must either be a member or a non-member of itself. Nor is the LNC always in the picture when the LEM plays this kind of structural role in setting up an argument against the theory. As often as not, we reject a theory because we think that all possible answers to some question it raises are counter-intuitive, rather than because any of them involve logical impossibilities. We reason as follows:

If α is true, is β true? Well, the combination of α and β would generate Counterintuitive Consequence #1, and the combination of α and $\neg\beta$ produces Counterintuitive Consequence #2, so α is probably false.

We reject claims *when* we discover that they are inconsistent, and nothing similar is true about incompleteness, but this is a chronological difference and not a philosophically interesting asymmetry. The LEM enters into the chain of reasoning we use to reject some theory at the beginning, as a premise—indeed, very often it enters the chain as such a trivially obvious premise we have no reason to explicitly state it. The LNC plays its role at the end of the chain, where we cannot possibly miss it.

Finally, while *incompleteness* may be an unavoidable fact of epistemic life, *paracompleteness* is not. To say that a theory is incomplete is merely to say that for at least one proposition α, the theory does not answer the question of *which* of the two possibilities (α, $\neg\alpha$) is right. A *paracomplete* theory asserts that Excluded Middle has broken down and both α and $\neg\alpha$ are, in some sense, *wrong*.

This last point needs to be put very carefully, for reasons explained by Hartry Field in his essay, "Is the Liar Sentence Both True and False?"

> There are many reasons why one might be tempted to reject certain instances of the law of excluded middle. And it is initially natural to take "reject" to mean "accept the negation of." But if we accept the negation of a disjunction, we certainly ought to accept the negation of each disjunct (since the disjunction is weaker than the disjuncts). So accepting $\neg(A \vee \neg A)$ should lead us to accept both $\neg A$ and $\neg\neg A$ (Field, 2005, p. 23).

It looks like, if the paracomplete theorist glosses rejection in terms of acceptance of negation, a paracomplete solution to the Liar Paradox ends up collapsing into the paraconsistent solution. Instead, Field recommends understanding rejection "[a]s a propositional attitude, on par with acceptance."

> To a first approximation anyway, accepting *A* is having a high degree of belief in it; say a degree of belief over a certain threshold T, which may depend on context but must be greater than 1/2. (Degrees of belief are assumed to be real numbers in the interval [0, 1].) To the same degree of approximation, rejecting A is having a low degree of belief in it: one lower than the co-threshold 1 - T. This has the desired result that rejection precludes acceptance. (And it allows, as of course we should, for sentences that we are uncertain enough about to neither accept nor reject.) (Field, 2005, p. 26).

But wait! Anyone who knows a little probability theory should already see the problem here. It is an axiom of the probability calculus that P(*A*) and P(¬*A*) must always add up to 1. If rejecting *A* means assigning it a probability of less than one half, then *rejecting A means accepting ¬A and* vice versa. Thus, paracompletishly *rejecting* both *A* and ¬*A* should mean *accepting* both *A* and ¬*A,* and it looks like (on the second definition of "reject") paracompleteness has once again collapsed into paraconsistency. Field gets around this problem by making the reasonable point that the axioms of the standard probability calculus grow out of classical logic. Rejecting classical logic means rejecting classical probability theory in the bargain. "A dialetheist ought to allow that P(*A*) + P(¬*A*) can be significantly greater than 1 (perhaps as high as 2)" and Field and his fellow paracompletists should allow it to sink to "substantially less than 1 (perhaps as low as 0)."

This last point nicely illustrates the distinction between *incompleteness* and *paracompleteness*. If a subject matter expert has no idea whether *A* or ¬*A* is true, she should set the probabilities of both *A* and ¬*A* at 0.5. If she "goes paracomplete" on the question of *A* vs. ¬*A,* she has to fundamentally revise the orthodox probability calculus—the one we use to such good effect for purposes ranging from gambling to meteorology—to even express her position. If we are not accustomed to rejecting theories because they force us to make this move (as we are accustomed to rejecting them because they are inconsistent), this only indicates that theorists do not accidentally stumble into paracompleteness as they stumble into inconsistency, *not* that paracompleteness is any less radical or counterintuitive than dialetheism.

1.5 Varieties of Gaps and Gluts

While Field focuses on the Law of the Excluded Middle, an older solution to the paradoxes focuses on Bivalence, the principle that every proposition is either true or false. This position, which we can think of as a form of "proto-paracompleteness," is usually described as the theory that there are "truth value gaps." Paradoxical sentences such as (2) were held to be "gappy," i.e., neither true nor false.

2. Sentence (2) is false.

One problem with this solution is that it is vulnerable to particularly obvious "revenge" problems. A revenge paradox is a reworked version of the Liar Paradox designed to frustrate an attempted solution. For example, think about sentence (3).

3. Sentence (3) is either false or gappy.

An only slightly extended version of the "Liar reasoning" we rehearsed in Sect. 1.1 delivers a contradiction. If (3) is true, it is either true and false (and hence true and not true) or true and gappy (and hence true and not true). If it is false, it is true and false (and hence true and not true). Finally, if it is gappy, it is both true and gappy—and hence true and not true!

Worse yet, we can generate revenge trouble for the proto-paracomplete view without using as "exotic" a Liar as (3). (1) will do perfectly well. I have found it convenient to start the chapter with (1) and use it as my "go-to" Liar thereafter, since doing so cuts to the heart of the contemporary debate, but this reverses the usual presentation of increasingly difficult forms of the paradox. (2) is a version of what is usually referred to as the Simple Liar, which is the sort of Liar that gap theorists had in mind. (1) is a version of the Strengthened Liar, the difference being that we do not need to assume Bivalence to derive a contradiction from the Strengthened form of the sentence.[6] (1) "says of itself" not that it is false but merely that it is not true. Well, if it is neither true nor false, it is *not true*.

In his discussion of truth value gaps in *In Contradiction*, Graham Priest compares Liar sentences to sentences such as (4), which is a version of "the Truth-Teller."

4. Sentence (4) is true.

Priest grants that, "Such a sentence is an obvious candidate for a Truth value gap." While he ultimately rejects the possibility of gaps, he thinks that *if* such things existed, it would be reasonable to suspect the Truth-Teller of being one. After all, "The semantic rules [. . .] governing the predicate 'is True' appear not to be sufficient to determine the Truth value of the sentence." Sentences such as (1)–(3) seem quite different to him.

> By contrast, in the case of [. . .a Liar sentence] the semantic conditions of the words involved seem to overdetermine its Truth value. [. . .It] would therefore seem a more plausible candidate for a Truth value "glut" than a Truth value gap, which is exactly, of course, what it is (Priest, 2006b, p. 15).

[6] Interestingly enough, Robert Martin has argued that the conventional view of the difference between Simple and Strengthened Liars I have sketched out here is wrong, and that the Simple Liar is "just as independent of Bivalence as its big brother" the Strengthened Liar (Martin, 1984, p. 2). However, JC Beall and Otávio Bueno have argued persuasively that Martin is mistaken. In a paper in *Analysis*, they prove that the two principles about truth and falsity Martin appeals to in order to derive a contradiction from the Simple Liar jointly entail Bivalence (Beall & Bueno, 2002 pp. 22–26). One of these principles is just the T-Schema. The other is the principle that $Tr\langle\neg\alpha\rangle\rightarrow F\langle\alpha\rangle$. Gap theorists would be well-advised to reject the latter.

Subsequent dialetheists have followed Priest in using the language of truth value "gluts." Here, however, the usual parallelism between dialetheism and paracompleteness breaks down. While even some advocates of versions of the paracomplete position have described it as a version of gap theory, Field regards this as a "serious mistake." (Field, 2008, p. 71)[7] If Field is right that talking this way is a mistake, JC Beall commits it too. In his 2005 paper "Transparent Disquotationalism," Beall advocates a hybrid view according to which there are both truth value gluts and "gaps."

> With Field I agree that there are gaps in the language, that some meaningful sentences are "indeterminate"—that neither language (its rules, etc.) nor the world determines that such sentences are dtrue or dfalse.[8] Semantic paradoxes themselves, I believe, give no good reason to think there are gaps. Rather, the appearance of gaps arises from reflection on vagueness, non-denoting terms, and other such familiar phenomena (Beall, 2005, p. 15).

In 2009's *Spandrels of Truth*, he switches to a pure glut view, arguing that "negation is exhaustive." (Beall, 2009, pp. 3–4)[9] In "Transparent Disquotationalism," Beall handles vagueness and the rest with gaps, but even at this stage he argues that at least some Liar sentences need to be treated as dialetheias (i.e., statements that are both true and false).

> The idea can be modeled using a four-valued language [. . .] Intuitively, 1 models sentences that are dtrue but not dfalse, 0 sentences that are dfalse but not dtrue, b sentences that are both dtrue and dfalse, and n sentences that are neither (Beall, 2005, p. 15).[10]

Field finds this way of parsing paracompleteness misguided for multiple reasons. First, since he is committed to both the T-Schema and the equivalent principle for falsehood—$F < \alpha > \leftrightarrow \neg\alpha$—any attempt to express his rejection of Bivalence by asserting that some sentences are *neither* true nor false would be equivalent to asserting the negation of some instances of the LEM. We saw where that led in Sect. 1.3.

Still, he *does* reject paradoxical instances of Bivalence, so it might seem reasonable to describe his views as a sophisticated version of the "truth value gap" theory. He rejects all attributions of truth values to paradoxical sentences! Why not describe the result as a truth value "gap"?

---

[7] He ascribes this mistake to Scott Soames.

[8] "Dtrue" and "dfalse" just mean "true" and "false." Like his usage elsewhere of "ttruth" to emphasize his "transparent" theory of truth, "dtruth" emphasizes his deflationism.

[9] In "Transparent Disquotationalism," he argues that the English word "not" and its equivalents in other natural languages are ambiguous between two different kinds of negations, only one of which is exhaustive. Liars constructed with one type of negation are gappy and Liars constructed with the other one are glutty. The key difference between the two versions of his view then is that in *Spandrels of Truth*, he treats negation as *unambiguously* exhaustive.

[10] When Beall says "dtrue" and "dfalse," the reader can simply parse this as "true" and "false." Similarly, in *Spandrels of Truth,* he uses the expressions "ttrue" and "tfalse." See n. 3 above for what he is getting at with this usage.

He has a response to this point, but it is not entirely persuasive. He claims that the paracomplete approach "asserts neither that there are truth value gaps nor that there aren't," by which he means that for defective sentences, both the relevant instances of Bivalence *and* their negations end up being rejected. (To put it differently, he rejects both the instances of Bivalence and the claim that these are "gaps" where "gaps" is understood in a proto-paracomplete way.) The problem with this response is that Field has no problem describing himself as an opponent of (certain instances of) the LEM, although he also rejects the negations of those LEM instances. Indeed, it seems natural to gloss this as an even deeper rejection of the LEM. For every LEM instance he rejects, he also must reject an infinite series of LEM instances—the disjunction of the original instance and its negation, the disjunction of *that* disjunction and its negation, and so on forever and ever. The same points naturally apply to Bivalence, since by both my classical and Field's paracomplete lights, Bivalence and the LEM are equivalent (at least given the Capture and Release rules for truth and their equivalents for falsehood and negation). As such, why not describe the view that rejects any application of a truth value to α (or to any statement directly or indirectly attributing a truth value to α) as a very deep rejection of Bivalence, and hence as a version of the theory that α is "gappy" with regard to its truth value?

Field sees things differently. For him, a "truth value gap" necessarily means an assertion that $\neg Tr\langle\alpha\rangle \wedge \neg F\langle\alpha\rangle$. As such, instead of describing paracompleteness as the view that there are truth value gaps and dialetheism as the view that there are truth value gluts, his preferred way of talking is that the paracomplete theorist has "acceptance gaps" and "rejection gluts" and the dialetheist has corresponding "acceptance gluts" and "rejection gaps."

I am perfectly happy to either talk this way or use the traditional language about truth values. The phrase "gaps or gluts" in the title of the book you are holding can be read either way, since—in a way that should come into sharper focus when I present my own view on the paradox—the distinction between "acceptance gaps" and "truth value gaps" makes no difference to me. It should be good enough for now to note that, given my orthodox commitments with regard to both logic and probability, rejecting a proposition α, asserting $\neg\alpha$, asserting $Tr\langle\neg\alpha\rangle$, and asserting $\neg Tr\langle\alpha\rangle$ all come to the same thing by my lights, so my project in what follows is to show that consideration of the paradoxes should not lead us to accepting gaps (or gluts) of *any* of these kinds.

1.6 Once More on the Comparative Plausibility of Restricting the T-Schema and Rejecting Classical Logic

So far, we have looked at three approaches to the Liar Paradox. (We will expand our focus to look at other paradoxes shortly, hence the plural "paradoxes" in the title of this book.) I have argued that all three are implausible.

I will not just be throwing stones here. My own solution comes in Chap. 8. Although to get there, we need to cover a lot of ground. Some of this is a matter of establishing useful principles (about negation, for example) that will play a direct role in diagnosing paradoxical sentences and formulating my preferred solution, but I will also be pushing back against other dialetheist and paracomplete lines of attack on classical orthodoxy. My larger project is to show Priest and Field and their ilk that, to slightly mangle the Irving Berlin lyric, anything they can do, we (classical logicians) can do better. Or, more modestly, in the trickiest cases (like Sorites), I want to show that sacrificing the LEM and/or LNC will not help anything. (On that point, I am reduced to, "Anything you can do, I can do *at least as well*.")

Out of the three ideas we started out with, restriction of the T-Schema is the only one I *won't* be returning to. As I said at the end of the first section, I join most of my antagonists in this debate in insisting on what is often referred to in this literature as a completely "transparent" truth operator—one where we can always "see through" Tr<α > to α and vice versa.[11]

This may seem perverse to some readers. After all, as eminent non-existent epistemologist Sherlock Holmes was fond of telling his assistant Watson, once you have eliminated the impossible, whatever remains, however improbable, must be the truth. If the LEM and LNC are *laws of logic*, and the T-Schema is not, should not the Holmes Principle lead us to take T-Schema restriction *more* seriously than dialetheism or paracompleteness?

I will make four points about this, going from the most philosophically interesting to the most mundane. The first is that it is not clear what (other than tradition) leads us to draw the borders of "the laws of logic" so as to exclude Capture and Release. Coming up with necessary and sufficient conditions with which to differentiate logic from non-logic is (such as the parallel demarcation problem in the philosophy of science) notoriously difficult. Capture and Release fit at least some plausible conditions, such as subject matter neutrality. (The truth operator applies in the same way to propositions about every subject.) And it certainly *looks* like they universally preserve truth. Of course, whether one thinks they actually *do* preserve it depends on what one thinks of the T-Schema-restricting solution to the Liar Paradox, but one could say much the same thing about the laws of classical logic called into question by the paracomplete and dialetheic solutions to the paradox.

Second, even *if* the best way of drawing the borders of "logic" leaves out the T-Schema, Capture, and Release, it would still make sense to make the preservation of all three a parameter of success for our project of "solving the paradoxes without sacrificing classical logic." To see why, start by noticing that "preserving classical logic" can mean several different things, not all of them equally philosophically interesting.

[11] Note that my policy in what follows is to refer to what is often called "the truth predicate" as "the truth operator." While the phrase "the truth predicate" is firmly entrenched in the literature, "the truth operator" is more accurate. Truth is something that competent speakers of natural languages only apply to whole statements.

At the bottom rung, we could simply endorse Graham Priest's solution to the paradoxes, and grant that sentences such as (1) are both true and untrue. Priest is a logical monist, and he believes that the uniquely correct logic is paraconsistent rather than classical, *but* (like many "deviant" logicians) he believes that we can "recapture" classical reasoning for most ordinary purposes. For example, as we will see in Chap. 4, Disjunctive Syllogism is paraconsistently invalid, but Priest thinks we can go ahead and use it in ordinary contexts where we are unlikely to encounter true contradictions—"John is either at home or at the library, and he's not at home, so he's at the library"—in much the same way that we can safely use Newtonian principles to solve word problems about baseball in high school physics tests, knowing as we do that the questions of force and velocity at stake there are unlikely to impacted by quantum or relativistic distortions. Priest's project of "classical recapture" preserves classical logic in the sense that it authorizes us to continue to reason classically about *most* subjects.

At a nearby rung, we could endorse open-ended logical pluralism, saying that a variety of different entailment systems stem from different genuine relationships of logical consequence, and that several of these systems (including classical logic) are appropriate for one reasoning task or another. In one way, this seems like at least a marginally more impressive sort of "preservation" of classical logic, since (unlike on Priest's picture) classical entailment comes out as a genuine relationship of logical consequence. Disjunctive Syllogism, for example, is *really* valid, not just something we can safely treat as if it were valid in certain contexts. On the other hand, it is (equally really) *invalid*, in that—for any pluralist picture that includes at least one paraconsistent logic within its family of logics—there are other equally genuine relationships of logical consequence relative to which Disjunctive Syllogism comes out invalid. Note too that this pluralist "endorsement" of classical logic is compatible with any number of different and mutually incompatible positions about what we can call the "ontology" of logic—i.e., with several different answers to questions such as, "Are there true contradictions?" Someone could pluralistically "preserve" classical logic in this way while still thinking that the vast majority of propositions that had ever been expressed in human language were either exceptions to the LEM or true counterexamples to the LNC. (Go ahead, though, and reason your heart out classically about those few domains where classical logic will not steer you wrong!)

One level up, we come to the position defended by Thomas Hofweber in his paper "Validity, Paradox, and the Ideal of Deductive Logic." Hofweber maintains even in the presence of the paradoxes that all the laws of classical logic are valid. He also affirms the principle that *inference rules are valid iff they preserve truth*. So far, so good. So good, in fact, that you might be forgiven for starting to wonder what remaining *raison d'être* there might be for the book in your hands. Read on. Hofweber says that the principle that inference rules are valid iff they preserve truth "has two readings." One is the obvious one. The other one is generic.

> The right hand side has a generic reading that is nicely brought up by using the plural. Inferences are truth preserving according to the generic reading just like bears are dangerous. That doesn't mean that each and every instance is truth preserving... (Hofweber, 2007, p. 150).

If this sort of "defense" of classical logic rings hollow, I would suggest that the reason is that the distinction between denying that all the laws of classical logic are *valid* in all contexts and affirming that they are always valid but denying that they are always truth-preserving feels like a distinction without a difference. Similarly, I would argue that the distinction between the Hofweber approach (whereby inference rules are always valid, but they do not always preserve truth) and the Harman/Horwich approach (whereby the sort of "truth" they preserve does not always *act* like truth) is, if not quite a distinction without a difference, then at least a distinction without a very interesting difference.

In what follows, I will try to do better.

1.7 Summary of the Chapters

Most philosophers on both sides of the debate about dialetheism and the LNC consider the Argument from the Liar Paradox to be the most powerful weapon in the dialetheist's arsenal, but Graham Priest has supplemented it with several secondary arguments. He offers dialetheic solutions to the "paradoxes of motion and change," to the antinomies of naïve set theory, to the paradox he takes to be implicit in Berkeley's "Master Argument" for idealism, and to the problem of how to make sense of inconsistent laws and norms. In Chap. 2, I will discuss each of these Priestly arguments. In Chap. 3, I will respond to them, showing how we can handle each issue at least as well without rejecting classical orthodoxy.

In Chap. 4, I will start to explore the logical consequences of dialetheism. Some of the logical rules that matter most for scientific, mathematical, and everyday reasoning are invalid in Priest's paraconsistent framework. In Chaps. 5 and 6, I will do a deep dive into the various ways in which Priest and his fellow dialetheist JC Beall have tried to solve this problem. I will argue that all such attempts fail.

In Chap. 7, I will turn from dialetheism to Hartry Field's paracomplete approach. I will delve into the consequences of this solution for probability theory, the notion of logical validity, and our ordinary reasoning practices. I will also argue that it faces revenge problems and that it fails to characterize paradoxical sentences in an informative way.

Finally, in Chap. 8, I will present my own solution to the paradoxes, arguing for a form of disquotationalism about truth from which it follows that Liars and other similar sentences do not express propositions. I will address several objections, show how this compares favorably with the non-classical alternatives surveyed in the book, and end with a consideration of possible "loose ends" for my defense of classical logic—quantum mechanics, "underdefined" (and "overdefined") predicates, and most troublingly the problem of borderline applications of vague predicates.

Chapter 2
Graham Priest and Dialetheism



Abstract Most dialetheists and most non-dialetheists agree that the argument from the semantic paradoxes to the existence of true contradictions is the strongest argument for dialetheism. Indeed, apart from Graham Priest, most dialetheists regard contradicts about the alethetic status of paradoxical sentences as the *only* true contradictions. Priest, however, has made independent arguments for true contradictions in several other domains. He postulates true mathematical contradictions based on Russell's Paradox and other antinomies of naïve set theory, some of which seem to be closely related to the semantic paradoxes, as well as true contradictions arising from inconsistent laws and norms, true contradictions at the limits of conceivability, and even true contradictions in the physical world derived from what Priest calls "the paradoxes of motion and change." In principle, Field could offer paracomplete versions of all of these arguments, but he has declined to do so. The "priestly" arguments for dialetheism are explored as well as reactions to them by Priest and Beall.

Keywords Priest · Field · Beall · Sets · Motion · Change · Dialetheias

2.1 Three Priests Walk Into A Bar: What's Funny About Dialetheism

Graham Priest has been the philosophical figure most associated with dialetheism since before the publication of the first edition of *In Contradiction* in 1987. This means that he has been the butt of many, many dialetheism jokes. A few have even made it into the literature. In a footnote to his *Mind* article "Priest's Paraconsistent Arithmetic," Nicholas Denyer thanks his subject in an ambiguously gracious manner.

> Thanks are due to an anonymous referee, and to Graham Priest, for commenting on an earlier version of this paper. Naturally, Priest disagrees with much of what I write—which is not to say that he does not also agree with it (Denyer, 1995, p. 574 n. 7).

Commenting on this in a footnote of his own, Priest says this is "a good joke—or at least it was the first few dozen times I heard its kind" (Priest, 1996, p. 658 n. 11).

© Springer Nature Switzerland AG 2022

B. A. Burgis, *Logic Without Gaps or Gluts*, Synthese Library 458,
https://doi.org/10.1007/978-3-030-94624-1_2

I myself have not heard as many Graham Priest jokes as all that, but I hope that Priest will forgive me if I include my favorite example of this genre here. I heard it from Gareth Young, who was at the time a doctoral student in Glasgow.

Three priests walk into a bar—an Irish priest, a Scottish priest, and Graham Priest. The Scottish priest says, "I'll have a whisky." The Irish priest says, "I won't." Graham Priest says, "I'll have what they're having."

No joke benefits from explanation, but jokes told by philosophy professors and graduate students can be a revealing window into their philosophical views. The punchlines of Graham Priest jokes tend to jarringly insert violations of the LNC (or attempted violations, as in the "three priests" joke) into everyday situations, thus playing off the disconnect between dialetheism and ordinary experience. Priest has diagnosed this disconnect as the source of a "mental cramp that often arises when people consider dialetheism."

> A commonly heard complaint is as follows (said with an air of puzzlement): 'I just cannot see what it would be like for a contradiction to be true, what it would be like, for example, for something to be a cup and not a cup [. . .]' (Priest, 2006b, p. 170).

One dialetheist strategy for assuaging this worry is to start by granting the central premise that it would indeed be absurd to take "the world" to match such inconsistent descriptions. On this view, "This object is a cup" cannot be true and false, because such a sentence would have to be both *made true* by the world being one way and *made false* by the world being otherwise, but paradoxical sentences such as (1) and (2) are an entirely different matter. Their truth values come not from "the world" but from what JC Beall calls "the overall rules (or logic) of the language" in which we formulate such sentences (Beall, 2009, p. 16).

This is not Priest's strategy. Far from restricting himself to postulating dialetheias whose truth values arise in a robustly "world"-free fashion, he has argued for dialetheic solutions to Russell's Paradox and other antinomies of naïve set theory, for a dialetheic analysis of inconsistent legal (and perhaps even moral) norms, for a "paraconsistent account of vagueness," and even for a dialetheic theory of the metaphysics of motion and change. We will start with that last item. Priest's strategy for addressing the "mental cramp" just described is a full frontal assault. He argues that we *think* we cannot understand what it would be for something to be a cup and not a cup, but in fact "something is both a cup and not a cup the instant it breaks into pieces" (Priest, 2006b, p. 170).

2.2 The Paradoxes of Motion and Change

This conclusion is startling. It is one thing to accept that sentences such as (1) and (2) are gluts. Such sentences are, as Priest's fellow dialetheist JC Beall puts it, "odd and rare beasts, tangled up in knots of truth and falsity on their surface [. . .]" (Beall, 2014b, p. 791). None of this is true about breaking cups. Why, then, does Priest posit gluts about *them*?

Consider four options. At the "instant it breaks into pieces," the cup in question is:

(i) Intact
(ii) Not intact
(iii) Neither intact nor not intact
(iv) Both intact and not intact

Priest's initial reason for rejecting the first two options, at least in many cases, is that there "is no good reason to say one rather than the other" (Priest, 2006b, p. 160). He rules out (iii) based on the general philosophical considerations he uses to rule out the possibility of truth value gaps. If there is no fact of the matter that makes something true, he thinks, this very fact suffices to make it false.[1] This is of course deeply philosophically controversial, but to the extent that his target here is the classical logician rather than a gappy logician whose favored logic blocks this inference, the point can be made on a purely formal level. Given Double Negation and DeMorgan's Laws, (iii) entails (iv).

Of course, all this assumes that paraconsistent and classical options are the only ones on the table. A paracompletist could reject (i) and (ii) without accepting (iii) or (iv), thus preserving the intuition that we should handle (i) and (ii) symmetrically without resorting to contradiction. As a matter of fact, the chief advocate of the paracomplete approach, Hartry Field, is generally wary of postulating LEM failures in contexts other than the "odd and rare beasts" discussed in Chap. 1, but that should not stop us from remembering when we examine each of Priest's arguments for glutty solutions to paradoxes other than the Liar that a paracomplete theorist *could* offer a parallel paracomplete solution in each case. If my argument in 1.4 goes through, such solutions are exactly as good or bad as their dialetheic equivalents. In each case, what Priest says for his paraconsistent solutions and what I say against them should apply equally well in the paracomplete case.

In his initial presentation of this argument, Priest grants that "different changes may be changes of different kinds" and thus that his argument does not establish that *all* instants of change should be treated dialetheically, but merely that in "a number of cases, there is no objective fact that can be appealed to to break the symmetry" between the two consistent options (Priest, 2006b, pp. 160–161). He seems to be particularly taken with this intuition in cases of *change of location*, hence the inclusion of *motion* in his phrase "the paradoxes of motion and change." Although perhaps there are at least some properties (other than location) such that at the instant in which an object changes from having to lacking the property or vice versa, it either (just) has the property or (just) lacks it. Is there any way to rule out this possibility?

Unfortunately for us classical logicians, there is a quite an easy way to do so. At any instant at which (i) is true, the cup has by definition *not yet broken*. At any instant

[1] Priest makes this argument at length in Priest, 2006b pp. 64-66. He refers the reader back to that discussion during his exploration of instants of change on pp. 160-162. His example in the latter discussion is a pen being touched to paper to write a sentence. "Since the motion is continuous, there must be an instant at which the pen leaves the paper. At that moment, is it on the paper or off?" (Priest, 2006b, p. 160).

at which (ii) is true, the cup has by definition *already broken*. More generally, at any instant at which an object O is going from having property P1 to property P2, O can neither (just) instantiate P1 or (just) instantiate P2, because at any moment at which either of these consistent scenarios hold, O has either not yet or already changed from instantiating P1 to P2.

At this point, things look very bad. We started in Chap. 1 with the ambition of saving classical orthodoxy from the clutches of revisionary logicians even with regard to such odd and rare beasts as (1) and (2). Now, it looks like our universe of moving and changing things is practically swimming in contradictions (or failures of the LEM). What has to be done?

Happily, there is actually a fairly straightforward way out for the classical logician. Although first, let us continue our tour of Graham Priest's garden of paradoxes.

2.3 Russell's Paradox and the Inclosure Schema

While the "paradoxes of motion and change" look quite different from the Liar Paradox, Russell's Paradox does not.

The unrestricted comprehension axiom of what is now called "naïve" set theory holds that for every property, there is a set of all and only the objects that possess this property. Thus, for the property of being a set that is not a member of itself, there would be a set of all sets that are not members of themselves. The set of all cats would be a member of this set, since it is a set rather than a cat and hence not a member of the set of all cats. The set of all sets, on the other hand, would *not* be a member of the set of all sets that are not members of themselves. The problem, as Bertrand Russell famously discovered, is that the set of all sets that are not members of themselves (henceforth, "the Russell Set") would be a member of itself if and only if it was *not* a member of itself. The structural parallel between the self-membership status of the Russell Set and the alethic status of the Liar is obvious on its face. Moreover, both paradoxes seem to have something to do with self-reference.

This connection was originally made by Russell himself, who thought that any adequate solution to his paradox (and related set-theoretic antinomies such as the Burali-Forti Paradox) should also apply to the Liar.

> There were older paradoxes, some of them known to the Greeks, which raised what seemed to me to be similar problems, though writers subsequent to me considered them to be of a different sort. The best known of these was the one about Epimenides, the Cretan, who said that all Cretans are liars (Russell, 1959, p. 77).

Russell took it to be a virtue of his "theory of types" that it could be used to tackle both paradoxes in the same way. Similarly, Priest today argues that all of the "paradoxes of self-reference" are of a kind, and that they should therefore be solved in the same way—i.e., dialetheically.

> The Principle of Uniform Solution (PUS) counsels: Same kind of paradox, same kind of solution. Since standard solutions to the set theoretic paradoxes do nothing to address the semantic paradoxes, and vice versa, this shows that each is inadequate (Priest, 2000, p. 123).

> The only satisfactory uniform approach to all these paradoxes is the dialetheic one, which takes the paradoxical contradictions to be exactly what they appear to be (Priest, 2002, p. 169).

While much of the interest in the Liar in subsequent analytic philosophy can probably be traced to Russell's linkage of his paradox about set theory to the ancient puzzle about lying Cretans, the dominant tendency among subsequent paradox-solvers has been to separate the set-theoretic and semantic paradoxes into separate categories that can be legitimately handled in quite different ways. Still, it is hard not to see why Russell and Priest find the similarity compelling.

We can reinforce that point by considering a paradox about properties that is sometimes described as a non-set-theoretic version of Russell's Paradox. Think about properties such as redness, tallness, and tastiness. Redness is not red, tallness is not tall, and tastiness is not tasty, so none of these properties instantiate themselves. The property of being a property, on the other hand, *does* instantiate itself. How about the property of being a property that does not instantiate itself?

The same paradox can be run in terms of predicates rather than properties. (Philosophically, this might or might not make a difference, depending on how we understand the relationship between predicates and properties.) Does the predicate of being a predicate that does not apply to itself apply to itself or not? Once we notice that an equally natural way to say "apply to itself" is "is true of itself," it looks like this paradox occupies a gray area between being a version of Russell's Paradox and a version of the Liar Paradox.

Worse, we can easily generate a paradox that does not straddle the line between the two paradoxes but rather neatly falls into both categories. Consider sentence (5).

(5) Sentence (5) is not a member of the set of all true sentences.

Given the assumption that there *is* such a set, (5) acts exactly like a "Strengthened Liar" sentence such as (1). If it is true, then by the T-Schema it follows that it is not a member of the set of all true sentences. . .but if it is not a member of the set of *all* true sentences, it is not true! On the other hand, if we say that it is untrue—whether we say that it is false or attribute to it some third, fourth or fifth truth value dreamed up by non-classical logicians—it is not a member of the set of all *true* sentences.

Patrick Grim has argued that the powerset axiom rules out the possibility that there could be such a "set of all truths," but in the original paper where he presents this argument, he conscientiously allows that such thing might be possible in an alternative, "many-valued" set theory (Grim, 1984, pp. 206–208). JC Beall has argued that, if we adopt the view that propositions are functions from worlds to truth values (what he calls "the S-view"), Grim's result can be avoided even within orthodox axiomatic set theories (Beall, 2000, pp. 38–41). Whatever may be true on either of those fronts, the important point for our purposes is that given the unrestricted comprehension axiom of *naïve* set theory, there *must* be such a set. As

such, the paradox about (5) joins Russell's Paradox and Burali-Forti as an antinomy of naïve set theory while *also* being a recognizable version of the Liar Paradox. If anything, (5) is much closer to "canonical" Liars such as (1) and (2) than many of the "revenge" Liars the reader will encounter in this book, or in just about any other book about paradoxes. The existence of paradoxes that clearly belong to both families does not by itself show that those theorists who sharply distinguish the two types of paradoxes and believe that it is legitimate to handle them differently are wrong. It does, however, put some pressure on the distinction.

Priest argues that the common thread that unites all of these paradoxes is that they satisfy the Inclosure Schema, whose elements are Existence, Closure, and Transcendence. There *is* some set or sentence (Existence) that belongs to some category such as "true propositions" or "self-membered sets" (Closure) and *by virtue of belonging to it*, it breaks out of it (Transcendence).[2] A theorist who solves Russell's Paradox by making the standard move of restricting the set-theoretic universe to the cumulative hierarchy of some standard system such as ZFC and solves the Liar Paradox by rejecting paradoxical instances of the T-Schema is in one case denying Existence and in the other accepting Existence but denying Closure and Transcendence. This is, from Priest's perspective, unacceptably disunified.

As with the hypothetical paracomplete solution to the "paradoxes of motion and change" discussed above, a theorist who was willing to embrace a paracomplete form of naïve set theory *could* solve Russell's Paradox and the Liar in a uniform way. Field's stated reason for declining to make this move is that the orthodox solution to the set-theoretic antinomies is a respectable one.

> Russell's paradox has a standard solution: there simply is no such set. [. . .] This solution to the Russell paradox is motivated by the hierarchical picture of sets. And that picture is not only natural, it is mathematically fruitful: it provides the framework for virtually all of contemporary mathematics. [. . .] The notion of set is primarily a mathematical notion, so it is hard to criticize this solution to the paradox (Field, 2008, pp. 2–3).

Still, we should not be too quick to accept this defense against the charge of lack of uniformity. There are at least three issues here. The first is that this reluctance to tinker with mathematically fruitful frameworks on philosophical grounds is hard to reconcile with Field's approach to the probability calculus, discussed above in Sect. 1.4. Nor is this just a cheap accusation of individual inconsistency. Field's willingness to embrace revisionary conclusions about probability is not some personal eccentricity of his. It is central to his way of prying apart negation from rejection to make sense of the whole paracomplete project.

Even more importantly, as Priest has pointed out, it is entirely possible to accept that "the hierarchy is an important and interesting set theoretic structure," well worth the attention that mathematicians lavish on it, without taking it to "exhaust the

[2] See Priest, 2002, Chapters 9–11, for details. The crucial point brought out by Priest's set-theoretic model of this structure is that in all Inclosure paradoxes, Closure and Transcendence are a "package deal." Just as the objects transcend the relevant categories *because* they belong to them, they belong to them *because* they transcend them.

universe of sets" (Priest, 2006b, p. 30). Priest's point here is reinforced when we remember that there is not really "a" cumulative hierarchy, but many such hierarchies delineated by the different axioms of multiple competing post-naïve set theories. Indeed, shortly after the passage just quoted, Field says that by "*the* standard solution" he means "the solution embodied in set theories *such as* Zermelo-Fraenkel" (my emphasis). The standard Platonist response to this proliferation is to grant that, so to speak, mathematical Heaven has many rooms. On what grounds could a paracompletist deny that some of these rooms are occupied by sets that are not organized into *any* of the hierarchies?

What it seems to get down to is Field's belief that because "the notion of set is primarily a mathematical notion," philosophers should be reluctant to criticize it. On the paraconsistent side, JC Beall makes just this move.

> *Objection: Russell's paradox and maths*. Russell's paradox is often thought to be the set-theoretic counterpart of the Liar. [. . .] The trouble for you is that standard—say, type-theoretic or otherwise— classical solutions to Russell's paradox are in tension with your dialetheic approach to Liars. In short, you can't have a classical mathematics without an ad hoc solution to Russell's paradox.
>
> Reply. What should be noted is that, in effect, Russell's paradox amounts to two paradoxes, one for set theory, and one for semantics. By my lights, mathematics is free simply to axiomatize away its version of Russell's paradox, as is standardly done. Sets were introduced within and for mathematics, a role that need not be constrained by our 'intuitions' about natural language (Beall, 2009, p. 112).

The problem is that all this deference to the expertise of the mathematicians only does any work in avoiding the need to solve the Liar and Russell's Paradox in a uniform manner if we start from the premise that the mathematicians have shown there to be no Russell Set. . .but (sadly, from my classical perspective) this is not the case. To see why not, start by asking to *which* mathematicians we should defer? The ones working within the universe of ZFC sets, or the ones who accept the basic ZF structure but reject the Axiom of Choice (the "C" in ZFC)? How about Peter Aczel, whose Anti-Foundation Axiom allows for "hypersets" not built up through any sort of cumulative hierarchy at all? There is a relatively mathematically uninteresting (but massively philosophically important) distinction between axiomatic systems such as ZFC or NBG or Morse-Kelly whose cumulative hierarchies start with the empty set and those that allow for sets that include non-set-theoretic objects as members. When philosophers with reductive views about time analyze *moments* as "sets of co-occurring events" or philosophers of biology offer set-theoretic analyses of concept of *species*, are they showing insufficient deference to mathematicians whose systems do not allow for sets that cannot be built up cumulatively from the empty set? Worse yet, are set theorists who assume the axiom of choice failing to give sufficient deference to those set theorists who assume its negation? Are those who live their mathematical lives within the austere confines of ZFC giving insufficient deference to Peter Aczel? As mentioned above, the standard maneuver is to postulate a very well-populated mathematical Heaven, but this move is incompatible with the premise that mathematicians have shown that there is no Russell Set. If the fact that sets of co-occurring events, or AFA hypersets, cannot be built up from the

axioms of ZFC does not settle the ontological status of *those* sets, the fact that the Russell Set cannot be built up from those same axioms does not settle *its* ontological status either.

One could argue that, while the result that some set cannot exist within *one* of these systems does not add up to a finding of its non-existence, the conjunction of all such facts does, but then what about Zach Weber, who has devoted himself to proving interesting results within a paraconsistent form of naïve set theory? Does he count as one of the mathematicians to which we should all defer? Standard-issue Platonists only reserve rooms in mathematical Heaven for *consistent* systems, but that is because they believe in the LNC. Whatever reason a dialetheist such as Beall might have for maintaining this restriction, it cannot be *that* one.

Let us grant Beall the distinction between "Russell's paradox [. . .] for set theory" and "Russell's paradox [. . .] for semantics." (The latter is the paradox about predicates and/or properties discussed above.) This should strengthen rather than weaken the case for a uniform solution. Instead of two somewhat similar paradoxes, we have three:

Par1: Russell's Paradox (for membership in the Russell Set)
Par2: Russell's Paradox (for the application of the "Russell Predicate")
Par3: The Liar Paradox (for the application of the predicate "not true")

However much one makes of the structural similarity between Par1 and Par3, Par1 is *even more* like Par2 than it is like Par3! One could satisfy demands for Par1-Par2 uniformity by "axiomatizing away" not only the Russell Set but the Russell Predicate, postulating a ZFC-like hierarchy of predicates. Alternately, there is always Priest's move of extending the dialetheic solution to the semantic paradoxes into set theory. Instead of doing either, Beall draws the line between paradoxes sufficiently unlike each other to be solved differently between Par1 and Par2, the most similar paradoxes on the list.

This might seem more defensible if we view the series Par1-Par3 as something like a Sorites series, with Par1 and Par3 as clear cases and Par2 in the vague penumbra between the two. . .but this is undermined by (5), which is both a clear case of a Liar sentence and the basis of a clear case of a paradox of naïve set theory. Of course, Beall could argue that while uniformity is one virtue to consider in comparing different packages of solutions, it can be overridden by other considerations. (He has suggested elsewhere that the rules of rationality generally involve such tradeoffs (Beall, 2013, pp. 210–223)) At the very least, though, such uniformity *is* a virtue, and one that his solutions conspicuously lack.

This matters for the kind of classical solution I will present and defend in Chap. 8. To follow through on my promissory note in Sect. 1.6 to show that we classical logicians can do everything that Priest and other deviant logicians do better (or at least just as well), I will have to present a classical solution to the Liar as unified with what I have to say about the rest of these paradoxes as Priest's dialetheic solution to the lot of them.

2.4 A Bishop and a Priest Walk Into a Bar... (Berkeley's Master Argument)

Priest has suggested that Inclosure Schema paradoxes such as the Liar and Russell's Paradox tend to arise when theorists explore "the limits of thought," where it is difficult to say everything we want to say without bumping into inconsistency.[3] To get a sense of Priest's picture here, think of the trouble that philosophers offering rational reconstructions of Wittgenstein's *Tractatus* have had with the passage in which the great man simultaneously gives and violates the recommendation that what "cannot be said" should be passed over in silence, or the difficulty that thoroughgoing epistemic skeptics have had in expressing skepticism about *everything* without undermining their own position. While Priest grants that many such philosophical finger traps are created by faulty philosophical premises, he thinks that the problem will arise even for ideally rational inquirers dealing with questions posed at a certain level of generality. We want to talk about sets of *anything at all*, and we get Russell's Paradox and Burali-Forti. We want a truth operator powerful enough to say of *any* sentence that it is true or untrue, but the price is the Liar. "Even though our conceptual/linguistic structure is, in a sense a human product," he tells us, "it does not follow that we have control over what we produce." In by far the best turn of phrase in *In Contradiction*, he says that this "is the moral of *Frakenstein*, and, in a much more horrific way, of *Capital*" (Priest, 2006b, p. 47).

If Priest's picture is right, we need to re-examine the history of philosophy with fresh eyes. Over the course of the last twenty-four centuries of "footnotes to Plato," there have been many instances in which one thinker has derived a contradiction from another's premises to refute them. Have there been at least some instances in which this dialectic has unwittingly uncovered *true* contradictions?

In his book *Beyond the Limits of Thought*, Priest attributes just such an unwitting discovery of a dialetheia at the limits of thought to Bishop Berkeley. In *Three Dialogues Between Hylas and Philonous*, Berkeley's mouthpiece Philonous argues for an extreme form of idealism, according to which objects such as tables, chairs, human bodies, distant galaxies, and the rest simply *are* all the ways that minds perceive them. When teaching the *Three Dialogues* in Early Modern Philosophy courses, I often hold up a cup of coffee, ask the students to look at it, and then ask them to humor me by closing their eyes for three seconds. I do the same, and when we all open our eyes back up, I ask if, according to Berkeley, the cup ceased to exist for the three seconds during which none of us were looking at it. The clue, I tell them, comes in fact that he was *Bishop* Berkeley.

This theistic idealism is established in the *Three Dialogues* by a combination of arguments against the possibility of truly mind-independent material objects on the one hand, and a firm rejection of skepticism on the other. If the skeptic doubts that we live in a rich and varied universe containing many wonders that no *human*

[3] This is the central theme of Priest, 2002.

perceives, Berkeley considers this to be just the sort of thing that can be refuted with an incredulous stare.

> How vivid and radiant is the shine of the fixed stars! How magnificent and rich the careless profusion with which they seem to be scattered throughout the whole vault of the sky! Yet the telescope brings into view a new host of stars that escape the naked eye. But neither ˙sense nor ˙imagination is big enough to take in the boundless extent ·of the universe· with all its glittering furniture. With all the hard work that we give to ˙those two faculties, exerting and straining each of them to its utmost reach, there's always a vast surplus left ungrasped. [. . .] Isn't the whole system immense, beautiful, more glorious than we can say or think? Then how should we treat those philosophers who want to deprive these noble and delightful scenes of all reality? How should we think of principles implying that all the visible beauty of the creation is a false imaginary glare? To put it bluntly, can you expect this scepticism [. . .] not to be thought extravagantly absurd by all reasonable people? (Berkeley, 1713, p. 29)

Berkeley's argument is essentially that the incoherence of belief in a mind-independent "material" reality leaves us with two choices. Either we can deny the existence of the "vast surplus left ungrasped" or we can make sense of how such a surplus could exist by saying that God grasps (and perceives) what we do not. Hence, by Disjunctive Syllogism, he arrives at idealism.

The crucial first premise is that nothing can exist independently of perceiving minds. Philonous strenuously argues for this claim throughout the first of the *Three Dialogues*, eventually trapping poor bumbling Hylas with what is often called Berkeley's "Master Argument" against mind-independence.

> **Phil:** [. . .] I am willing to let our whole debate be settled as follows: If you can conceive it to be possible for any mixture or combination of qualities, or any sensible object whatever, to exist outside the mind, then I will grant it actually to be so.
> **Hyl:** By that test, the point will soon be decided. What is easier than to conceive a tree or house existing by itself, independently of and unperceived by any mind whatsoever? I conceive them existing in that way right now.
> **Phil:** Tell me, Hylas, can you see a thing which is at the same time unseen?
> **Hyl:** No, that would be a contradiction.
> **Phil:** Is it not as great a contradiction to talk of conceiving a thing which is unconceived?
> **Hyl:** It is.
> **Phil:** The tree or house therefore which you think of is conceived by you.
> **Hyl:** How could it be otherwise?
> **Phil:** And what is conceived is surely in the mind.
> **Hyl:** Without question, what is conceived is in the mind.
> **Phil:** Then what led you to say that you conceived a house or tree existing independently and out of all minds whatsoever?
> **Hyl:** That was an oversight, I admit [. . .] (Berkeley, 1713, p. 21)

Priest takes this argument to successfully demonstrate "that a contradiction follows from the claim that something is not conceived (and that this itself is conceived)" (Priest, 2002, p. 65). He does not, however, think that it succeeds in establishing the truth of Berkeleyan idealism, or even of the disjunction of that view and skepticism about the vast surplus. Berkeley thought this argument was a reduction to absurdity of the claim that some things are unconceived. Priest takes

it to be a sound argument for the existence of a true contradiction about the limits of conceivability.

He finds Berkeley's reading of his own argument untenable for at least two reasons, both of which are extremely plausible. The first is that the Master Argument (read in Berkeley's way as a *reductio* of the claim that "any sensible object whatever" can "exist outside the mind") can easily be tweaked to establish not just that everything that exists "is conceived" but that everything "'is conceived of *now*', 'is conceived of *by me*', 'was conceived of by Berkeley on his 42nd birthday', etc. None of this would have been acceptable to Berkeley [. . .] who is quite clear that things may exist unconceived by me, since God conceives them: God himself must exist independently of my conception" (Priest, 2002, p. 65). Philonous could have asked poor Hylas at 2:20 PM, whether there were any objects unconceived by Hylas between 2:20 and 2:21 PM, and when Hylas answered in the affirmative (as he doubtless would have done with even more confidence than in the original dialogue), Philonous' *reductio* could have proceeded in exactly the same way, and the resulting argument would have been exactly as good or bad as the original Master Argument.

Priest's second, related reason for thinking that (whatever its other virtues) the Master Argument does not succeed in showing that everything is conceived is that, clearly, *not* everything is conceived. He gives the example of "things going on on planets in galaxies far away in space and time that no sentient being will ever think of" (Priest, 2002, p. 66). Of course, as we have seen, Berkeley has an answer to this question of a galactic surplus left un-grasped by *us.* At this point, one might think the difference gets down to Graham being such an adamantly atheistic Priest that he prefers dialetheism to theism, but *that* question can be put to one side. Remember that the target of the argument can be narrowed and widened at will, to derive a contradiction from, for example, the claim that some things were not conceived of by a particular medieval monk. Priest is willing to embrace this contradiction, but not to reject its negative conjunct.

2.5 Inconsistent Laws and Norms

Priest believes that there are true contradictions about psychological states such as "conceiving of that which is not conceived," about physical states of motion and change, about semantic objects such as statements that attribute untruth to themselves, and about mathematical objects such as the Russell Set. This leaves only the realm of *social* facts contradiction-free, and Priest takes care of this in his discussion of legal contradictions in Chap. 13 of *In Contradiction*.[4]

> [S]uppose there is a pair of statues, one of which requires a car-owner to change registration plates on January 1st, and the other of which forbids working on a Sunday. About every seven years the average law-abiding citizen is embarrassed (Priest, 2006b, p. 184).

[4] We will save Priest's dialetheic theory of borderline applications of vague predicates for 8.7.

The two halves of the inconsistency do not relate to each other in the right way for this to have anything to do with the Inclosure Schema, but there *is* a well-known Schema paradox in the vicinity—Charles Chihara's "SecLib" case.

> In the official rulebook of a club called 'Secretary Liberation' (or 'SecLib' for short), we find the statement:
>
> A person is eligible to join this club if, and only if, he (she) is secretary of one and only club and he (she) is not eligible to join that club (Chihara, 1984, p. 223).

In Chihara's example, SecLib eventually grows to the point where they have to hire a secretary of their own, "Ms. Finline." The question of whether Ms. Fineline is herself eligible to join SecLib plays out exactly like the question of whether the Russell Set is a member of itself.

Of course, Chihara's case is about a conflict within the rules of a club, not the laws of a nation. If this distinction matters, it is easy enough to make it go away. To convert Chihara's example into a legal case, just add a government that, concerned with corruption and discrimination in private clubs, passes a law requiring all formal associations to honor their written membership rules!

As well as inconsistent legal *obligations*, we can have inconsistent legal rights. Here too, Priest provides a plausible example.

> Suppose that there is a certain country which has a constitutional parliamentary system of government. And suppose that its constitution contains the following clauses:
>
> In a parliamentary election:
>
> (1) no person of the female sex shall have the right to vote;
>
> (2) all property holders shall have the right to vote.
>
> We may also suppose that it is part of common law that women may not legally possess property. As enlightenment creeps over the country, this part of common law is revised to allow women to hold property. We may suppose that a de facto right is eventually recognised as a de jure one. Inevitably, sooner or later, a woman, whom we will call 'Jan', turns up at a polling booth for a parliamentary election claiming the right to vote on the ground that she is a property holder. A test case ensues. Patently, the law is inconsistent. Jan, it would seem, both does and does not have the right to vote in this election (Priest, 2006b, pp. 184–185).

Of course, well-functioning legal systems have various mechanisms by which such inconsistencies can be resolved. Perhaps the Supreme Court of this country has the right of judicial review. None of this touches Priest's point, which is about the status of Jan's legal rights *before* the Court makes its ruling.

We have four options, just like the options about the breaking cup in Sect. 2.2.

(a) Jan has the right to vote.
(b) Jan lacks the right to vote.
(c) Jan neither has nor lacks the right to vote.
(d) Jan both has and lacks the right to vote.

A paracomplete legal theorist could reject both (a) and (b), but it looks like the classical logician is committed by the LEM to accepting one or the other of them—as

well as being committed by the LNC, since (c) DeMorgan's into (d) in any case.[5] How do we pick? Of course, one or the other could be true without our having epistemic access to the fact, but the resulting ontological picture is implausible. Does fealty to the principle of Excluded Middle really commit us to believing that a truthmaker for one or the other of these competing claims is sitting in some legal analog to mathematical Heaven, waiting for the Supreme Court to discover it when it makes its ruling?

In what follows, I will try to show how consistency can be purchased in this and in the other realms we have been discussing without resorting to such *ad hoc* assumptions.

[5] I actually think this is a case where the paraconsistentist has a slight advantage over her paracomplete rival. If Jan tries to vote, she will doubtless be arrested, as early suffragettes were arrested in the United States when they did the same, and if she is convicted and her lawyer appeals the verdict based on the clear statutory language giving *all* property-owners the right to vote, the case will doubtless be allowed to proceed. It feels far more natural to describe this situation as one of legal inconsistency than legal completeness.

Chapter 3
How to Solve Priest's Paradoxes Without Sacrificing Classical Logic



Abstract Objections are made to (almost) all of the "priestly" arguments for dialetheism discussed in Chap. 2. For example, we can accept normative *dilemmas* either in the law or in other domains without having any particularly compelling reason to postulate the existence of true normative contradictions. An argument is made against realism about naïve set theory, and hence against belief in the truth of contradictions that arise from naïve set theory. It is granted that we might have a good reason to be realists about naïve set theory if we had sufficiently compelling non-set-theoretic reasons to accept dialetheism, but it is argued that in the absence of such reasons the antinomies of naïve set theory give us no independent reason to accept the existence of true contradictions. The "paradoxes of motion and change" are undermined by means of a defense of Bertrand Russell's theory of change against Priest's critique. Bishop Berkeley's argument is shown to rest on a fundamental equivocation—one that also arises in a classic argument by Max Black. Finally, it is granted that the version of Russell's Paradox that concerns not sets but predicates or properties has not been solves, and that any adequate solution to the Liar Paradox must also apply to these closely related paradoxes.

Keywords Priest · Russell · Set-theoretic · Berkeley · Dilemmas · Contradictions

3.1 Normative Dilemmas Without Normative Contradictions

In the previous chapter, I tried to fairly and charitably review Priest's case for glutty accounts of the paradoxes of motion and change, of the antinomies of naïve set theory, of the tricky issues about the limits of conceivability revealed by Berkeley's Master Argument (or, in Priest's hands, Berkeley's Paradox), and by the possibility of inconsistent legal rights and obligations. In each case, I have noted that philosophers with Field-like views about logic could if they were so inclined supply paracomplete duals of Priest's paraconsistent accounts of these subjects, and that at least in most cases whatever can be said for (and much of what can be said against) the paraconsistent account can be said just as fairly of its paracomplete dual. Thus,

© Springer Nature Switzerland AG 2022

B. A. Burgis, *Logic Without Gaps or Gluts*, Synthese Library 458,
https://doi.org/10.1007/978-3-030-94624-1_3

Priest's arguments are most powerful when framed as objections to *classical* accounts of these subjects. These are, of course, precisely the accounts I aim to defend. Happily, in most of these cases a fairly straightforward defense is available.

To start with the last case, Priest is no doubt right that the positive law, like any other normative system designed by fallible human beings, can contain "glitches" that generate inconsistencies about what rights and obligations we have. Not all inconsistent rights or obligations even result from glitches. Plausibly, for example, a well-designed legal system *should* hold those who knowingly promise to do two incompatible things and sign contracts to that effect in breach of contract when they inevitably fail to honor at least one of them.

Moreover, while I have never heard even an interesting prima facie example of someone simultaneously having and lacking a *moral right*, there may well be such creatures as inconsistent *moral obligations*, resulting in tragic situations in which anything an agent does runs afoul of the moral law. A full exploration of the issues such a case would raise in terms of both meta-ethics and normative ethics—where any such suggestion is immediately met with the ghost of Kant, wailing that *ought implies can*—is well outside the scope of my present project. As such, I will confine myself here to the observation that, whatever one makes of such a possibility, it would at least leave the honor of the *LNC* intact.

The reason is that O¬α and ¬Oα are two very different beasts. (Oα ∧ O¬α) is the logical form of an inconsistent obligation. (Oα ∧ ¬Oα) is the form of an inconsistency *about* our obligations. Only a true instance of the second would be a dialetheia. If standard deontic logics allow ¬Oα to be derived from O¬α, the reason is that a key assumption built into such logics is that the normative systems they were designed to model were themselves internally consistent.

A superficially reasonable objection could be made here. If the goal is to revise deontic logic to allow for Oα ∧ O¬α to be true without revising the classical base logic to allow for outright contradictions (about obligations or anything else) to be true, more work has to be done than just rejecting the inference from O¬α to ¬Oα. We also have to reject at least one of two other inferences that might both seem to be more basic and intuitive than that one—the inference from Oα to Pα or the inference from Oα to ¬P¬α. With both in place, we can infer from (Oα ∧ O¬α), not (Oα ∧ ¬Oα), but *a* contradiction.

1. Oα ∧ O¬α	Premise
2. Oα	1 Conjunction Elimination
3. ¬P¬α	2 Inference from Obligation to the Impermissibility of the Negation
4. O¬α	1 Conjunction Elimination
5. P¬α	4 Inference from Obligation to Permission
6. P¬α ∧ ¬P¬α	3, 5 Conjunction Addition

What has to be done? On the one hand, it might seem intuitively obvious that if you have an obligation to do something, you are permitted to do it. On the other hand, to block the move from Oα to ¬P¬α is to reject the deontic instance of the interdefinition of modal operators beloved by modal logicians of all kinds. The

answer is that the intuition that obligations must entail permissions—that, in other words, if there is a rule requiring you to do something, surely you are not breaking any rules by doing that very thing—is simply a disguised form of the intuition that obligations must never conflict. Once we recognize this, we can see that any attempt to sneak contradictions-about-permissions through the backdoor of the proof we have just considered crashes against one or the other horn of a dilemma. If obligations never conflict, 1 above is false and the LNC is safe. If they *can* conflict, the inference from 4 to 5 fails.

All of this nicely takes care of contradictions arising from inconsistent obligations, whether legal or moral, but what about Priest's example about directly clashing legal *permissions*? One could parse the question of Jan's right to vote in terms of the obligation of judges and police to either allow her to vote or to stop her from voting, and stand by everything that has been said so far about how such inconsistent obligations would not entail any outright contradiction would hold, but perhaps this is not entirely satisfying. The issue, after all, is not just of the obligations imposed on others *by* Jan's right. It is a matter of the right itself, and there the logical picture is bleaker. It is not simply a matter of ($P\alpha \wedge P\neg\alpha$), but ($P\alpha \wedge O\neg\alpha$), which is equivalent given the interdefinition rules relating O and P (and the Law of Double Negation) to the outright contradiction ($P\alpha \wedge \neg P\alpha$).

1. $P\alpha \wedge O\neg\alpha$ Premise
2. $P\alpha \wedge \neg P\neg\neg\alpha$ 1 Modal interdefinition of $P\alpha$ and $O\alpha$
3. $P\alpha \wedge \neg P\alpha$ 2 Double Negation

At this point, JC Beall's response strikes me as decisive.

> I admit that this is a common phenomenon, but I think that, contrary to Priest, the phenomenon is not one that requires countenancing base-language gluts. What we have, in such cases, is exactly what Priest says: according to legislative body X, those in category A have the right to vote but those in B do not have the right to vote. But this, I suggest, is not a contradiction. The 'according to X' is essential to laws, and—importantly—the 'according to X' does not enjoy appropriate Release features (Beall, 2009, p. 126).

As he points out, if there was a valid legal version of the truth operator's Release rule, letting us go from "according to legislative body X, Y" to "Y," we would not *need* carefully crafted examples of inconsistent rights and obligations of the kind Priest gives us to get legal dialetheias.

> Just witness Australia and the USA. According to Australian law, it's illegal to drive on the right. According to USA law, it's legal to drive on the right. We don't infer a glut—namely, that it's legal and not legal to drive on the right—because we do not take the legal 'according to' operators to Release (Beall, 2009, pp. 126–127).

One might claim that such laws come with an implicit clause limiting their scope to the territory of the country making the law, but even this move would falter against real-life examples such as Taiwan, emphatically considered by the government of the People's Republic of China to be part of the territory of the PRC. Taiwanese and mainland Chinese drivers use the same side of the road, but the laws of Taiwan and the PRC differ on all sorts of points, such as what law-abiding citizens

should do in the event of a Chinese invasion of Taiwan. The history of the world is full of examples of successful wars of national liberation being followed by treason trials for citizens of the newly independent country for the crime of acting like very loyal citizens of the larger country against whom the war was fought.

Before leaving this case, let us take seriously the possibility that Beall is wrong and there are "operator-free" legal facts of some sort—facts not just about what rights or obligations some legislative body says we have, but facts about our "actual" legal rights and obligations. If there are such things, surely, they do not include a universal obligation to drive on both the left and the right side of the road, or even obligations by the residents of Taiwan to help both sides in the event of an invasion. This premise begs no question against the dialetheist. Quite independently of contradiction-avoidance, no one thinks that drivers in the USA have a legal obligation to drive on both the right and left side of the road because *no one thinks they have a legal obligation to drive on the left side.* Similarly, no one outside the PRC, and presumably even very few people inside that country, seriously believes that Taiwan's taxpayers are breaking the law by not sending some part of their earnings to Beijing. Presumably the reason that even those theorists who might be inclined to believe in "actual" legal facts do not believe that Taiwanese taxpayers have this obligation is that they think such facts have something to do with the "facts on the ground" created by effective enforcement. . . .and this is exactly where the possibility of an "operator-free" contradiction about Jan's right to vote falls by the wayside.

Enforcement in these cases is a logical impossibility. Just as any possible decision of SecLib's membership committee regarding Ms. Fineline's application to join the club would violate their written rules, any possible decision by officials at the polling station regarding whether to let Jan vote would violate one or the other of the conflicting statues. A real-life version of the Barber Paradox would create no empirical contradiction about the barber's shaving habits. The barber (who is a man, and who lives in the town) could post a sign in the window of his shop stating that it was his policy to shave men in the town iff they did not shave themselves, and every other man in town could play along, but when it came to his own case, nothing that he did with shaving cream and a razor would make it true that he shaved and failed to shave himself. Similarly, nothing the authorities could do when Jan tried to vote would amount to both enforcing her right to vote and enforcing her lack of such a right and thus creating a free-standing contradictory legal fact.

3.2 How to Solve the Paradoxes of Naïve Set Theory

Structurally similar points apply to the paradoxes of naïve set theory. Are there, in fact, such things as sets and other abstract mathematical objects? If not, claims about them are true only "according to" the axioms of ZFC or NBG or AFA or whatever, and paraconsistent set theory can simply be added to the list. Although if such objects do exist, why think that the Russell Set is one of them? I said above that Beall and Field, *given their logical commitments*, have no principled reason for denying

that naïve set theory occupies one of the many rooms of mathematical Heaven. Although given my logical commitments, I have an excellent reason. Sets with inconsistent self-membership conditions are logical impossibilities. A dialetheist who simply helped herself to the contrary assumption in arguing against the logically and mathematically orthodox view would be begging the question against her targets.[1]

As nominalists never tire of pointing out, we have no direct epistemic access to sets (or other mathematical objects) of any kind. Here is Jody Azzouni, for example, in his 1994 book *Metaphysical Myths, Mathematical Practice: The Ontology and Epistemology of the Exact Sciences.*

> It is not merely that mathematical objects play no causal role in the processes we use to learn about their properties; it is that they seem to play no role at all. (Imagine that mathematical objects ceased to exist sometime in 1968. Mathematical work went on as usual. Why wouldn't it?) (Azzouni, 1995, p. 56).

A quick look at the constellation of positions the philosophy of mathematics shows that Azzouni's suspicion about how this counterfactual (and counterpossible?) scenario would play out is a common one. Platonists tend to come in two flavors—the "many rooms" Platonists discussed above, and Quine-inspired Platonists grudgingly willing to pollute the desert landscapes of their preferred ontology with exactly as much abstract mathematical flora and fauna as they judge to be indispensable to our best scientific theories. (The man himself, with his occasional loose talk of "rounding out" our overall theory of the world with extra abstracta, was not entirely consistent in his grudgingness, but his inconsistencies do not seem to stem from anything relevant to the matter at hand.) The grudging Platonist implicitly admits Azzouni's point. If our only evidence for mathematical abstracta is a sort of sideways indirect deduction from physics and chemistry and the rest, *mathematicians* clearly do not have a direct window into their subject matter. In their own way, the position of "many rooms" Platonists starts from the same concession. The nominalist believes that we have no reason to believe in the existence of any particular mathematical object, so (according to various nominalists) we should either disbelieve in or remain agnostic about the lot of them. The "many rooms" Platonist, disagreeing with the nominalist about whether we are justified in believing in the *category* but agreeing that we do not have any way of evidentially discriminating between, for example, ZFC sets and AFA sets, simply assigns a parcel of land in the mathematical realm to every system whose axioms do not produce logical impossibilities.

[1] To be clear, the point of bringing up my own commitments here is not to use my own commitment to classical logic or the authority of classical mathematics as an argument against the view that inconsistent sets exist. As the chief proponent of that view frequently points out (in both set-theoretic and other contexts), such moves would beg the question against him. My point is a comparative one. Beall, Field and I all regard Priest's view that inconsistent sets exist as implausible. We all wish to reject. I claim that I have good principles by my lights to avoid it while their situation relative to their commitments is quite different.

On any of these three views, the argument for dialetheism from Russell's Paradox should leave us cold.[2] If the nominalist is right that we are not justified in believing in abstract mathematical objects of any kind, it follows trivially that we are not justified in believing in the Russell Set. If the "many rooms" Platonist is right, the LNC is equally safe. When it comes to non-physical objects whose properties we cannot track in any of the ways we can track the properties of even electrons and protons, much less the sort of "mid-sized dry goods" we can see and touch, the default assumption that they are not exceptions to the *logical* laws that hold of everything we *can* track is just about our only tool for ruling anything out. Perhaps, if we had sufficiently powerful reasons (independent of set theory) to change those logical views, a corresponding change in the logical standards for admission to mathematical Heaven would be in order, with *non-triviality* replacing consistency, but I will argue over the course of this book that this at least is not the case *yet*. My project here is to defend classical logic as our best current theory about which inferences preserve truth, not to argue that it should be immune from revision if a good reason to reject it *were* to emerge.

. . . .and this brings us to grudging Platonism. In *Doubt Truth to be a Liar*, Priest offers an interesting thought experiment about future science.

[2] While the two forms of Platonism I have mentioned overwhelmingly dominate the part of the current landscape occupied by Platonisms, perhaps a third form of Platonism—rarely articulated by current philosophers of mathematics, but perhaps implicit in the views of some philosophers and mathematicians in an early era—goes something like this: *Mathematical Heaven only has one room and it's occupied by whatever we get with the simplest and most intuitive set-theoretic axioms.* As an anonymous reviewer rightly points out, the Naïve Comprehension Axiom is certainly simpler and more intuitive than the alternative axioms of any post-naïve theory. Given this third version of Platonism, the intuitiveness of Naïve Comprehension could add up to a set-theory-specific argument that could be non-question-beggingly deployed against my view. I would reply in two ways. First, an appeal to the raw intuitiveness of naïve set theory would have to be weighed against the raw intuitiveness of the LNC. "Set" is a theoretical idea with a somewhat unclear relationship to the pre-theoretical notions that introductory-level explanations try to associate with it, like bunches of grapes, packs of dogs, and so on. Arguably the overwhelming majority of humans have no intuitions of any kind about sets. It is far more plausible that everyone *does* have the intuition that nothing can be both true and false in exactly the same way at exactly the same time. Of course, when most people are introduced to the concept of a set, naïve comprehension sounds very intuitive, but the idea that there are sets that are and are not members of themselves still sounds quite counterintuitive—as does the idea that we cannot legitimately use Disjunctive Syllogism to reason about sets! At the very least, it is not obvious that the intuitiveness of Naïve Comprehension is more powerful than the competing intuitions at stake here. Second, this third form of Platonism does not strike me as having much going for it, considering that set theory is hardly the only branch of mathematics where multiple sets of axioms deliver very different results, and thinking only one of them serves us as an accurate roadmap of Mathematical Heaven—if such a place exists—seems hopelessly arbitrary. Indeed, in at least one such case believing that the *only* room was the one described by the simplest and most intuitive axioms would be deeply strange. Euclidean axioms are more intuitive than the axioms of non-Euclidean geometries, but thanks to Einstein we now know that the geometry of the actual physical world is non-Euclidean. It would be very odd to combine realism about geometry with anti-realism about the particular geometry actually indispensable to our best science!

> Let us suppose that we come to predict a collision between an enormous star and a huge planet. Using a standard technique, we compute their masses as $x_1$and y_1, respectively. Since masses of this kind are, to within experimental error, the sum of the masses of the baryons (protons and neutrons) in them, it will be convenient to take a unit of measurement according to which a baryon has mass 1. In effect, therefore, these figures measure the numbers of baryons in the masses. After the collision, we measure the mass of the resulting (fused) body, and obtain the figure z, where z is much less than $x_1+ y_1$. Naturally, our results are subject to experimental error. But the difference is so large that it cannot possibly be explained by this. We check our instruments, suspecting a fault, but cannot find one; we check our computations for an error, but cannot find one. We have a puzzle. Some days later, we have the chance to record another collision. We record the masses before the collision. This time they are $x_2$and y_2 . Again, after the collision, the mass appears to be z (the same as before), less than $x_2+ y_2$. The first result was no aberration. We have an anomaly.
>
> We investigate various ways of solving the anomaly. We might revise the theories on which our measuring devices depend, but there is no obvious way of doing this. We could say that some baryons disappeared in the collision; alternatively, we could suppose that under certain conditions the mass of a baryon decreases. But either of these options seems to amount to a rejection of the law of conservation of mass(-energy), which would seem to be a rather unattractive course of action.
>
> Then someone, call them Einquine, fixes on the fact that the resultant masses of the two collisions were the same in both cases, z. This is odd. If mass has gone missing, why should this produce the same result in both cases? An idea occurs to Einquine. Maybe our arithmetic for counting baryons is wrong. Maybe the appropriate arithmetic is one where z is the least inconsistent number, and p (the period of the cycle) $= 1$. For in such an arithmetic $x_1+ y_1= x_2+ y_2= z$, and our observations are explained without having to assume that the mass of baryons has changed, or that any are lost in the collisions! Einquine hypothesizes that z is a fundamental constant of the universe, just like the speed of light, or Planck's constant (Priest, 2006a, pp. 162–163).

In the circumstances Priest is describing, grudging Platonists might have a good reason to allow inconsistent mathematical objects into their ontology. The mere fact that the unrestricted comprehension axiom of naïve set theory has some initial intuitive force, however, gives them no such reason.[3]

One final note is in order before our walk back through Priest's garden of paradoxes moves onto the next item. Azzouni's speculation about what would have happened if all the mathematical objects had disappeared in 1968, plausible as it is, has not commanded universal assent. In Azzouni's 2006 book *Tracking*

[3] Friends of reviving naïve set theory sometimes claim that unrestricted comprehension is "analytic," but it is hard to know how to parse this. It certainly is not analytic in the sense of being logically derivable from the meanings of any of the relevant terms, such as, "Bachelors are unmarried." Perhaps "analytic" is being used in a looser sense, so that it means something like "obvious to anyone familiar with the relevant concepts," but this sounds suspiciously like a simple appeal to intuition. Unrestricted comprehension *does* sound very intuitive the first time you hear it. Then again, so does the wrong answer to the Monty Hall problem. So does the claim that all transfinite sets have the same cardinality. Mathematicians have an irritating habit of showing that things that once sounded obvious are not even logically possible. The point here is not that appeals to intuition are illegitimate, but that the *logical* intuition that no object can have and lack the same property at the same time runs considerably deeper than the intuition that all the objects that fall under a certain description somehow form a further, mathematical object outside of space and time.

Reason, he notes that (in private correspondence) Adam Morton has offered a different view.

> [M]athematicians might lose their imaginative powers: suddenly be unable to speculate about mathematical objects—consequently, be unable to hypothesize they could then try to prove. Instead, those mathematicians stubbornly intent on continuing their art in the wake of the brutal abstracticide just contemplated would have to resort to constructing numerous detailed but boring proofs of utterly trivial results in the hopes that doing so day after day, year after year, might finally result in groping their way back to something interesting. [. . .Sadly, the abstracticide would] prevent them from distinguishing interesting from uninteresting results. What makes a result interesting is its illumination of the objects it's about [. . .] Attempts to read previous mathematics—even their own—would be similar to experiences of certain aphasics after a stroke; they wouldn't (now) even understand what such proofs *meant* (Azzouni, 2006, pp. 140–141).

I find Morton's position difficult to take seriously.[4] The epistemic faculty being attributed to mathematicians is hard to square with their status as biologically standard human beings. (Perhaps I am biased on this point. As the son of a math professor, I might not be emotionally prepared for the possibility that my mother belongs to a distinct species.) Although let us accept for the sake of argument that Morton is right. What follows?

In Sect. 2.3, I argued that Beall and Field are being unreasonable (given their logical commitments) in deferring to mathematicians on the question of the Russell Set's non-existence. After all, I claimed, the only reason mathematicians assume there cannot be inconsistent truths about mathematics is that they think that there cannot be inconsistent truths of any kind. Remove that commitment to the LNC, and what is left?

If Morton is right, my objection to Beall and Field fails. . .and Russell's Paradox gives us even *less* reason to abandon classical orthodoxy.

3.3 How to Solve the Paradoxes of Motion and Change

Bertrand Russell discovered the paradox that bears his name in 1901. The "paradoxes of motion and change" go back to pre-Socratic thinkers such as Zeno. These paradoxes were, however, quite adequately solved. . .decades before Graham Priest took them up and used them to argue for dialetheism. . .by Bertrand Russell.

Before turning to Russell's solution, let us review what Priest gets right. Object O is going from having to lacking Property P. At the moment of change, does it. . .

(a) (Just) have P?
(b) (Just) lack P?

[4] To be fair, it is not entirely clear in context that Morton is advocating this position, as opposed to simply offering it up as another possibility. I will refer to it, as Azzouni does, as "Morton's position," for the sake of simplicity.

(c) Neither have nor lack P?
(d) Both have and lack P?

Priest might be too quick to dismiss (c). A certain sort of gappy logician, of the kind referred to as "proto-paracomplete" in Sect. 1.5, might reject the classical inference rules allowing us to derive (d) from (c) for reasons that make as much sense given his starting points as the dialetheist's reasons for rejecting the classical inference rules that allow us to derive triviality from inconsistency. In contexts where the laws of logic are themselves in dispute, one should be careful about cavalierly dismissing such possibilities.

That said, it is hard to defend either of the two classically acceptable answers to this little multiple-choice quiz. If (a) is true, O has not yet changed from having P to lacking P. It might be the moment immediately before the change, just as a moment where (b) was true might be a moment immediately after the change, but how could either be the moment *of* change?

The simple answer here is that there are not any such things as "moments of change" in Priest's sense. As Bertrand Russell puts it in *The Principles of Mathematics*, "Change does not involve a state of change" (Russell, 1903, p. xxxiii).

All change is change *between* different states of an object or event at different times. To say that O changes just *is* to say that it has P at T1 and not-P at T2. David Lewis summarizes the Russellean account of change nicely in the language of temporal parts.

> Change is qualitative difference between different stages—different temporal parts—of some enduring thing, just as 'change' from east to west is a qualitative difference between the eastern and western spatial parts of the same landscape (Lewis, 1976, pp. 145–146).

In terms of motion (i.e., change of location), to say that an object is in motion over the course of some period of time is to say that it occupies different locations at different moments within that period. To say that it is at rest over a certain period is to say that at every moment in that period it has precisely the same position. There can no more be an instantaneous state of rest than there can be an instantaneous state of motion. More generally, there cannot be an instantaneous moment of change—except in the sense that the overall state of an object at a given instant might necessitate (or just increase the probability) that the object will be in a different state at a future time.

Priest has complained that this picture fails to accommodate our intuitions about change.

> [On this view] there is no change in the world at all, just a series of states patched together. The universe would appear to be more like a series of photographic stills, shown consecutively, than something in a genuine state of flux or change (Priest, 2006b, p. 162).

> A sequence of states, even a dense and continuous one, indistinguishable from corresponding rest-states, does not seem to be a state of motion. If God were to make different temporal slices of an object at rest in different places and string them together in a continuous fashion, he would not make them move (Priest, 2006b, p. 174).

But what *exactly* is missing from Russell's theory of change? Put differently, what is added to it by the addition of inconsistency? Unless Priest wants to suggest that Liars are always in the middle of changing between truth and falsity, there cannot be an innate link between change and contradiction. What non-question-begging reason could he give for thinking that such a link arises when contradictions are true of *concrete* objects?

Assume for the sake of argument that Priest is right to say that on the "cinematic" theory, true change is somehow absent. Any given "photographic still" of an object includes all the physical facts that will, when combined with the causal laws, entail that the object's future properties will be different from its current properties. None of this, we are assuming, is satisfactorily change-like. Well, then, why would the addition of some new "static" facts, even ones that contradicted the static facts already in the picture, constitute the addition of *change*?

Priest's otherwise puzzling intuition comes into clearer focus when he turns to Zeno's Arrow Paradox. If all there is to motion is the arrow occupying different locations at different instants, then any individual instant at which the arrow is in the air, a complete snapshot description of the arrow's properties would not include the property of *being in motion*. Over the course of an interval of time, however, it moves. "How," Priest asks, "can going somewhere be composed of an aggregate of going nowhere?" (Priest, 1985, p. 340).

As a matter of mathematics, this objection straightforwardly fails. (Compare the isomorphic oddity of a line segment of a certain length being composed of an aggregate of zero-dimensional geometric points.) Priest acknowledges this but remains unmoved by it.

> That one can prove a small mathematical theorem or two is one thing. However it does not ease the discomfort one finds (at least that I find) when one tries to understand what is going on physically, when one tries to understand how the arrow actually achieves its motion. At any point in its motion it advances not at all; yet in some apparently magical way, in a collection of these it advances. Now a sum of nothings (even in infinitely many nothings) is nothing. So how does he do it? (Priest, 1985, p. 341)

This is an odd response. Whatever one's views on the question explored in the previous section about the appropriate limits of philosopher deference to mathematicians, if it is appropriate for philosophers to defer to mathematical orthodoxy about *anything*, it is surely appropriate for us to do so on the matter of summing. It is one thing to question whether every mathematically coherent structure necessarily represents a possible way that the physical world could be—one might have the intuition, for example, that there is some sense in which it would be metaphysically impossible for there to be an imaginary quantity of Unobtanium beneath the surface of Pandora. It is quite another to claim based on sheer intuition that really existing quantities do not have the quantitative relationships they can be shown to have by means of proving a small mathematical theorem or two! One of the delights of mathematics is the way that it continually and elegantly frustrates intuition. As obvious as it might seem that someone confronted by the Monty Hall Problem has an equal chance of winning by sticking with their original guess or switching, it is not so. That all transfinite sets have the same cardinality is both intuitively obvious

and straightforwardly wrong. So it goes. Furthermore, as JC Beall has pointed out, the contemporary physical *science* of motion also seems to fit far more naturally with the Russellean theory of motion—and the anti-Russellean intuition appealed to by Priest seems to belong to a family of dubious intuitions that conflict with various empirical results.

> Of course, some scientific theories are unintuitive, but 'intuition', especially about the physical world, sometimes needs to be re-schooled. Motion and change may well be cases in which [. . .] re-schooling along Russellean lines [is] advisable. Indeed, I think the guiding 'intuition' has problems of its own. [. . .] Consider, for instance, the all-too-common 'intuition' that intelligence cannot be the 'sum' of unintelligence. 'How can intelligence be composed of an aggregate of unintelligence?', one might ask. More directly: a sum of dim bulbs is at most dim.
>
> The analogy, of course, is slightly exaggerated, but I think that it is on target. It is difficult to see what is special about the motion case that is lacking in cases like e.g. evolution or the mind in particular. More generally, we have learned that many intuitions about composition-like—or aggregation-like—principles are simply unfounded in reality in general (Beall, 2009, p. 128).

Little more needs to be said about Zeno's Arrow Paradox or the more general issue of the alleged failure of Russell's picture to capture the intuition that the world is in a "genuine" (rather than merely "cinematic") "state of flux or change." However, Priest has a final argument to offer against the "orthodox" Russellean account.

> Suppose that the universe was a Laplacean universe whose state at any time was determined by its state at any (prior) time. Then the orthodox account of change is impossible. For the instantaneous state of an object in the universe cannot even determine whether it is at rest or in motion, and hence whether it is in the same or a different point at a subsequent time. (Recall that the velocity—or momentum—of an object is not defined by its instantaneous state.) Now I am certainly not insisting that the universe is Laplacean. (It is not.) But it is a curious theory which rules it out *a priori* (Priest, 1985, p. 340).[5]

In reply to this, Chris Mortensen seems to concede that we would need Priest's "Hegelean" theory of motion to make sense of Laplaceanism. He just does not see this as a problem. (Mortensen is a friend to paraconsistency but, like Beall, a skeptic about Priest's arguments on the issue of motion and change.)

> In order to set aside this second argument, let us be clear about what Priest is denying. He is denying that velocity being non-zero is sufficient for motion; that is, $v \neq 0 \rightarrow$ motion. [. . .] Now, I suggest, one should not go so far as to say that Laplacean determinism is *impossible* or *incoherent*. It is *possible* for motion to be as Priest and Hegel say it is. On the other hand, Priest and Hegel must admit that the relational facts of the matter are present even in their story; that is, $v \neq 0$. Hence, their view is at least vulnerable to the charge of multiplying entities beyond necessity. The alternative, that nonzero velocity is sufficient for motion, is simpler and explanatorily clearer. Moreover, there is also the difficulty that Laplacean

[5] Tanaka (2007) seems to endorse this argument without adding anything to it. I say "seems" since the dialectical context—Tanaka's criticisms of Chris Mortensen's criticism of Priest—renders Tanaka's precise position on some of this unclear (at least to me).

> determinism appears not to be true, because of the probabilistic nature of quantum measurement (Mortensen, 2004, p. 20).

Mortensen's point about the Russellean account's advantage in explanatory simplicity is a good one, but what should we make of the rest of his reply?[6] Mortensen's confidence about causal indeterminism strikes me as misplaced. Plenty of non-Copenhagen interpretations of quantum mechanics remain as live options, and whether future physics supply the kind of hidden variables that would be necessary to show the universe to be deterministic down to its subatomic innards, Priest is surely right that philosophers should not be in the business of ruling out such possibilities a priori. So—does the Russellean account do this? Does Priest's "Hegelean" account do any better?

Before we dig into those questions, it is worth pointing out that Priest's worry should apply just as well to Copenhagen-derived indeterministic pictures of the physical world. Take one of Priest's examples from *In Contradiction.*

> As I write, my pen is touching the paper. As I come to the end of a word, I lift it off. At one time it is on; at another it is off (that is, not on). Since the motion is continuous, there must be a moment at which the pen leaves the paper. At that instant, is it on the paper or off? (Priest, 2006b, p. 60).

According to Russell, there is no such instant. At each instant, the pen is on the paper or not. As such, there is no way to tell from its instantaneous state whether it will be on or off at the next moment.

Now, suppose for the sake of argument that the world is not entirely deterministic, perhaps because the Copenhagen interpretation of quantum mechanics is true. On this view, given any given total instantaneous state of the universe as a whole, giving subatomic indeterminacy its due, the probability that, given the physical laws, the pen will be off the paper a moment in the future will not be 1. Giving quantum indeterminacy its due, it will be. . . .perhaps 0.9999999. (The fact that Priest's pen is being held by a conscious agent changes things hardly at all. Even indeterministic philosopher Robert Kane, who manages to squeeze a libertarian account of free will out of the assumption of indeterministic quantum events in the human brain, thinks that causally undetermined human decisions are extraordinarily rare!) If, as Priest suggests in the passage about "Laplaceanism," the instantaneous state of a physical object necessarily contains *no* information about its future location on the Russellean picture, this should be just as much of a problem for a (plausible) indeterministic account of the physical world as for a "Laplacean" one.

Is it true that this information would be lacking on the Russellean account? An implicit fallacy about aggregation of the type discussed by Beall might be at work here. If the instantaneous state of the universe does not contain facts "about motion," does it follow that it cannot contain anything from which Laplace's demon could

[6] How good a point it is might depend at least in part on what one makes of the analogy between multiplying entities and multiplying contradictions. I will return to this analogy in Sect. 6.1. For now, it is worth pointing out that it is at least a good dialectical point against Priest's overall package of views. . .since Priest endorses just such an analogy in other contexts!

derive such information? I am not so sure. Assume that Russelleanism is true. An arrow is, at a given instant, poised in mid-air. This fact, together with the relevant physical laws, surely entails that the arrow will not occupy precisely the same position two seconds hence. . .and it does not seem out of the question that less crudely obvious facts about the arrow's instantaneous state could entail finer-grained facts about its future location. At the very least, answering such a question seems like a job for the physicist, not the philosopher. Nevertheless, let us assume that Priest is right to think that no such conclusion can be drawn. Let us see what follows.

At the moment when Priest finishes writing and his pen leaves the paper, he holds that it is both on and off the paper. Does this supply Laplace's demon with the necessary information with which to derive whether the pen will be both on and off the paper at any given moment? Clearly not. First, as mentioned above, contradictoriness cannot *entail* transition from one state to another given Priest's overall package of views. Liar sentences are not mid-transition from truth to falsity. The Russell Set is not in the process of going from being a member of itself to being a non-member of itself. (Whether sentences can change truth value is philosophically controversial, but to the best of my knowledge no philosopher or mathematician denies that sets have their members essentially and thus cannot gain or lose them!) Perhaps Priest could claim that true contradictions *about physical objects* only arise when they are in the process of changing properties. Even this would come with a significant philosophical cost, given Priest's more recent friendliness to paraconsistent accounts of vagueness. (A man with few enough hairs on his head for it to be vague whether or not he is bald can stay that way for a long time!) Let us put that last point to the side and assume that contradictions concerning physical objects always indicate instants of change. It still would not be the case that Laplace's demon would be able to derive from a total state of the universe that included Priest's pen both being on and not-on the paper that it would be not on in the future! At best, from this fact and the relevant physical laws, the demon could derive the disjunctive conclusion that the pen was in the midst of changing from on to off *or* from off to on. The probability of either disjunct is 0.5, which is what it would be if the demon had no information whatever the instantaneous physical state of the world and had nothing to go on but the LEM.

To raise the probability-given-the-physical-laws of the pen being *off* a moment hence from 0.5 to a proper deterministic 1, or even a plausibly indeterministic 0.9999999, we need. . .information about at least one past instantaneous state, so the direction of change can be determined. Although if Laplace's demon is allowed *this* information, he can derive the future position of the pen or the arrow on *either* the Russellean picture or Priest's "Hegelean" alternative.

3.4 Where Berkeley and Priest (and Max Black!) Go Wrong

Before we consider Priest's use of a modified version of Berkeley's "Master Argument" to argue for a contradiction at the limits of conceivability, let us consider the original argument. Berkeley scholars disagree about precisely what the argument was meant to establish and how Berkeley meant to establish it, but—if read in the conventional way, as a *reductio ad absurdum* of the claim that anything can be "mind-independent" in the *sense* of "mind-independent" at issue between idealists and non-idealists—the argument rests on two important equivocations. First of all, whatever one makes of the relationship between conceivability and possibility, there is no way to get from the premise that we cannot conceive of anything that *we* are not conceiving of to the conclusion that there is anything conceptually or otherwise impossible about a universe that does not contain any minds, and whose various objects thus are not being conceived by anything in *that* universe. Second, there is a profound difference between "conceiving of" a thing in the sense that the thing falls under the scope of some extremely general and abstract thought and "conceiving of" it in any more interesting sense. For example, think of the mental relationship that Berkeley's God must have to my coffee cup in order for Berkeleyan idealism not to fly in the face of common-sensical ideas about object permanence. If the cup continues to exist while my students and I are not looking at it or thinking about it, does it do so as an indistinct blur? This would be hardly less counterintuitive than the coffee cup blinking out of existence entirely during those seconds. If God is going to fulfill His role in reconciling Berkeleyan idealism to our ordinary ideas about the stable existing of coffee cups, He must perceive every detail—the shape and size of the cup, the flavor and temperature of the liquid, and, well, *so on*—so that the coffee cup remains as constant in all its properties as it would be if a team of super-scientists were scrutinizing it without interruption.

Since Priest is not interested in arguing for theistic idealism, neither of these points might seem relevant to the use to which he puts Berkeley's argument. Although as we will see, Priest's version of the argument rests on a version of Berkeley's second equivocation.

Assume that some materialist philosopher has the thought, *Some things exist even though no one conceives of them*. Further assume that a race of invisible flying carnivorous lizards live on a planet in a galaxy so different that its existence is not known to any human astronomer. The lizards themselves are not self-aware so they do not (in any sense) conceive of themselves. No species capable of conceiving of anything interacts with the lizards. Nevertheless, Priest and Berkeley would argue that if the materialist philosopher is right, he *is* conceiving of the lizards (and all other unconceived objects). Berkeley would take this to show that the materialist philosopher is (just) wrong, Priest that the philosopher's thought is a true contradiction. Let us take a close look at how Priest backs up the premises of his altered, dialetheist version of Berkeley's argument.

> Now, one premise of the argument is: it is conceived that there is something that is unconceived. Or, on the modal interpretation: it is conceivable that there is something

> inconceivable. The first of these is, again, unproblematically true since I, for one, conceive such things. Even to suppose them as the first step of a *reductio* requires one to conceive them (they do not even have to be possible) (Priest, 2002, p. 66).

Is this true? As we have already seen, different conceptions of conceiving and conceivability have starkly different consequences. Which one is Priest using? As it happens, he is admirably precise about this at the beginning of his presentation of his version of the argument.

> Suppose, for example, that you conceive of George Elliot. George Elliot was Marian Evans. Have you conceived of Marian Evans? In a sense you haven not; in a sense you have, but you may not know it. [. . .W]e will do best to interpret conception in its extensional sense here. It is worth noting that even those who hold that conceiving is mainly (primarily, normally) extensional cannot deny that there is an intensional sense, namely conception-under-some-description-or-other. And just as I can (and we will) interpret 'conceive' extensionally, we will interpret 'conceivable' extensionally (conceivable-under-some-description-or-other) (Priest, 2002, p. 66).

So far, so good. This is just the extremely minimal sense of "conceive of" sketched out above. For an object to be conceivable by a given agent in this sense is for there to be a way for that agent to form a thought "about" the object in the way that the materialist philosopher's thought can be construed to be "about" the invisible flying carnivorous lizards on the distant planet despite the philosopher's not knowing that they exist.

> The other premise of the argument is: there is something that is unconceived; or, in the modal interpretation: there is something that is inconceivable (Priest, 2002, p. 66).

Given our exploration above of the consequences of Priest's definitions of conceiving and conceivability, it may seem impossible that Priest could establish either version of *this* premise. It is. Watch for his implicit substitution of a different and richer sense of "conceives."

> It is obviously difficult to give an example of something that is unconceived [. . .] but we know that there are (and were) things going on on planets in galaxies far away in space and time that no sentient being will ever think of. These are unconceived things. Moreover, in my life I have only a finite number of thoughts, and so will conceive of only a finite number of things, and so will everyone else. The number of things conceived (even by the human race) is therefore finite. Since there is an infinite number of things (numbers, points in space, etc.), most of them will not be conceived (Priest, 2002, p. 66).

This is just an equivocation. The things going on on planets in galaxies far away in space and time are unconceived on all sorts of (more interesting) senses of "conceived" but they *are* conceived in Priest's extensional sense. "Being a thing going on on some planet in some galaxy far away from ours in space and time" is a description. Priest himself, in thinking of that sentence, was conceiving of all such things under that description and hence conceiving of them under-some-description-or-other! Similarly, the infinite number of things populating his ontology (numbers, points-in-space, etc.) are unconceived (even by the whole human race!) only under a definition of conceiving sufficiently rich that something is only conceived if a *specific* thought is devoted to it and only it. Priest himself, however, personally

conceived the infinite lot of them in the conceived-under-some-description-or-another sense when he was writing the quoted passage, since the single description "being one of the infinite number of things that populates reality" clearly applies to all of them.

He seems to proceed more carefully when he turns to *conceivability*.

> The claim that there are some things that are inconceivable is less clearly true. It is pretty clear that a medieval monk could not conceive of a computer programmer, a Hilbert space, or a black hole, at least under those descriptions: he did not have the concepts. And presumably, in exactly the same way, there are things of which we can not now conceive and which future generations will be able to. But recall that conceivability is extensional here; and it does not follow from this that these things are not conceivable by us under any description (Priest, 2002, p. 66).

Unfortunately, things go back off the rails almost immediately.

> The premise is, none the less, true. If something is conceivable there must be some way of singling it out before the mind. Such singling out may involve direct perception (acquaintance) or may be purely by description. Now consider, for example, a point in space. Continua being dense, there is no way that any particular point can be picked out by our finite perception; and since there is an uncountable number of points but only a countable number of descriptions in our language, there is no way that every point can be picked out by a description (Priest, 2002, pp. 66–67).

There is no way that each can be picked out with a *unique* description, but there is a world of difference between defining conceivability as being "conceivable-under-some-description-or-other" as Priest does at the outset of his presentation of his version of the argument and defining it as being conceivable-under-some-*uniquely-referring*-description-or-other.

In first sense of "conceived," the one without the uniqueness constraint, everything is most definitely conceived by at least some people at some times, because people sometimes think thoughts whose extensions are fixed by extremely abstract and general descriptions. For example, anyone who thinks a lot about metaphysics is liable to have a thought like, *All objects have properties*. Every object falls under the description "being an object." Better yet, since anyone *can* form such thoughts, it looks like everything is *conceivable* by everyone. If anything were unconceived or inconceivable in *this* sense, Priest would be able to establish a contradiction about conceivability—"everything is conceived, but some things are not conceived"—with a much simpler argument than the one he uses in *Beyond the Limits of Thought*. However, the extreme ease of "conceiving of" something under-some-description-or-other makes it extraordinarily implausible (even in dialetheist terms) that anything could be unconceived, much less inconceivable, in this sense.

How about the second sense—the one formed by adding "uniquely referring" to Priest's definition? Priest is surely right that plenty of things are not conceived of by anyone in this sense—and surely *wrong* to claim that it is "unproblematically true" that it is conceivable that some things are unconceived or unconceivable, that he himself conceives of such things, and that "even to suppose them as the first step of a *reductio* requires one to conceive them (they do not even have to be possible)." The assumption that it is possible to "single out" non-existent objects is neither

unproblematic nor uncontroversial—think of Quine's difficulty counting the possible fat and/or bald men in his doorway—and the assumption that this is possible with regard to *impossible* objects may seem even stranger. To be fair, Priest defends both assumptions in Priest (2005). Let us assume for the sake of argument that both assumptions are correct. It would still be mysterious and awe-inspiring if Priest were able to single out with a unique description something that no one has (or could!) single out with such a unique description—and equally awe-inspiring if anyone other than Priest turned out to have such remarkable mental powers.

Things get no better when Priest turns to what he describes as a clearer case.

> The matter is even clearer with uncountable totalities of mathematical entities, like the ordinals, where perception is not even on the cards. There are not enough names to go around for every ordinal. But even if there were, there is no way of singling out most ordinals to pin names on them (Priest, 2002, p. 67).

If something can only be conceived of if someone can either perceive it or pin a unique name on it, then the claim that Priest (or anyone else) can conceive of things that no one is conceiving of is straightforwardly false. To conceive of some particular object failing to be perceived by anyone and failing to be named by anyone, Priest would have to either perceive or name it!

Max Black's celebrated but fallacious argument against the necessary identity of indiscernibles relies on an almost exactly similar confusion. One of Black's key assumptions is that the only way of pinning the name a on one of his two spheres and b on the other is to supply a description that picks out one but not the other—"let us call the sphere that _____ a and the other one b." Speaking of them as sphere a and sphere b without being able to supply such a description is, he says, "just pretending to use a name" (Black, 1952, p. 157). Assuming the LEM, Black is either right or wrong about this. Either way, his thought experiment fails to establish the possibility of non-identical indiscernibles. To see why not, consider the possibilities. If he is wrong and there is some other way of pinning names on the spheres, then we should be able to distinguish the spheres by pointing to various relational properties which sphere a has and sphere b lacks—for example, even if a's configuration relative to b is symmetrical in every way to b's configuration relative to a, we can still assert that a is co-located with a but not with b. On other hand, let us assume that he is right. Given the symmetry between the configuration of the spheres, Black is clearly right to say that any way of filling in the blank above would apply equally well to each of the two spheres. If it follows from all of this that it is impossible to express or make sense of the claim that they *differ in various relational properties*, should not it be equally impossible to express or make sense of the claim that they *fail to be identical* despite *having all the same properties*? To make precise his assertion that the world in which only these two objects exist constitutes a genuine counterexample to the necessary identity of indiscernibles, should not Black have to be able to say of his spheres that for every property p, a has p iff b has p, but that $a \neq b$? If not—if his argument is not weakened by being expressed at an intuitive, imprecise level that does not require pinning names on his spheres—why should we demand that *objections* to that argument be formulated at a higher level of precision?

3.5 What Has Not Been Solved

Sadly, even if one takes all of the paradoxes discussed in this chapter to have been decisively solved by the considerations I have brought to bear, this would leave what most observers take to be the core of the case for dialetheism untouched. Even as sympathetic a critic as Priest's occasional co-author Jay Garfield dismisses "Priest's alleged dialetheias of motion and change" out of hand as "spurious consequences of bad metaphysics" (Garfield, 2004, p. 237). Striking a more respectful tone, JC Beall says that Priest's arguments on these matters "are to be taken seriously," but Beall does "not take them to establish their goal" (Beall, 2009, p. 130). The only dialetheias Beall is willing to postulate are the ones arising from the Liar and related semantic paradoxes, and in this he seems more typical of other dialetheists. Most critics of dialetheism seem to share the view that the argument from those paradoxes is the most powerful weapon in the enemy's arsenal.

Later, I will make the case for my favored solution to the Liar and its ilk, and for the superiority of that solution (and of classical solutions in general) to solutions that require sacrificing the LNC and/or the LEM. But what if I am wrong about all of that? If there are some true counter-examples to the LNC, how bad a result would that be? I will spend the next three chapters exploring that question.

Chapter 4
Dialetheism and the Laws of Logic



Abstract Graham Priest accepts the traditional definition of validity as truth-preservation, but conceptual headaches arise when we try to dialetheically make sense of truth-preservation. Does truth-preservation mean that true premises always lead to true conclusions, or that they never lead to false ones? Problems appear either way. Alan Weir points out that we normally expect valid inferences to preserve falsity in the direction going from conclusions to premises as well as preserving truth in the direction going from premises to conclusions, but it is deeply unclear that dialetheism can satisfy this requirement. Contraposition and *Modus Tollens* might be invalid if dialetheism is true. When we put this together with more obvious losses such as *Disjunctive Syllogism* and *Reductio Ad Absurdum*, it begins to appear that the dialetheist cannot make sense of massive chunks of ordinary reasoning. Thus, "classical recapture," whereby it is shown that it would still be rational to use the full resources of classical logic to reason about every domain if dialetheism was true, is crucial to making dialetheism plausible. Unfortunately, such recapture is far more difficult than it might appear. An initially promising attempt embodied in Priest's "Principle R" is shown to fail.

Keywords Validity · Contraposition · Weir · Priest · Dialetheism · Reductio · Truth-Preservation

4.1 Dialetheism and Validity

A straightforward objection to the dialetheist solution to the paradoxes might go like this:

The dialetheist says that some contradictions are true, but that can't be right! If any contradictions were true, everything would be true. For example, since contradictions logically entail everything, if sentence (1) really were both true and untrue, (6) follows.

(5) *The moon is made of green cheese.*

If you aren't willing to accept that the moon is made of green cheese, you're rationally obliged to reject dialetheism.

© Springer Nature Switzerland AG 2022

B. A. Burgis, *Logic Without Gaps or Gluts*, Synthese Library 458,
https://doi.org/10.1007/978-3-030-94624-1_4

Priest has a standard response to this sort of criticism. In his book *In Contradiction*, where he lays out the most exhaustive version of his case for a dialetheic solution to the paradoxes, he frequently says things like this:

> Any appeal to this principle [that from a contradiction, anything follows] or to classical logic in general, can only, therefore, beg the question (Priest, 2006b, p. 6).

> [W]hat is at issue here is the very correctness of classical logic. Hence the issue cannot ([even] partially) be settled in this way without thoroughly begging the question (Priest, 2006b, p. 160).

As we will see in the discussion of Curry's Paradox in 5.1, the issue of what does and does not constitute "begging the question" in a debate such as this is not nearly as clear-cut as it might initially seem but let us put that to one side for the moment and acknowledge the core of Priest's point. Those of us with orthodox views about logic think that the inference from contradictions to triviality (the conclusion that everything is true) is *vacuously* truth-preserving. We are not making a counterpossible claim, speculating about what *would* be true in a universe with a very different logical structure than our own. Rather, we think that the inference from contradictory premises to arbitrary conclusions never takes us from truth to falsehood *because we think that contradictions are never true*.

Since the dialetheist denies this assumption, they lack this motivation to believe that contradictions entail triviality. Moreover, they have excellent reason to deny it, since they believe for example that (1) is true and untrue and do not believe the moon to be made of green cheese.

However, this denial comes at a price. Consider one of the easiest ways to get from the contradiction about (1) to the truth of (6).

1. (1) Premise
2. $\neg$(1) Premise
3. (1) $\vee$ (6) 1 Disjunction Addition
4. (6) 2, 3 Disjunctive Syllogism

While there are more exotic strategies available here to the designer of a system of paraconsistent logic, the two most obvious ways to block the inference from (1) to (6) are to say that either Disjunction Addition or Disjunctive Syllogism is invalid. Priest takes the second approach, and other dialetheists tend to follow his lead. Disjunctive Syllogism is invalid in Priest's preferred logic LP ("the Logic of Paradox"), on the grounds that if some proposition α is both true and untrue and some proposition β is (just) untrue, ($\alpha \vee \beta$) and $\neg\alpha$ will be both be true, but the inference from these two premises to β will have failed to preserve truth.

A structurally parallel dialetheist argument *could* be given against Disjunction Addition, *if* the dialetheist adopted a slightly different understanding of truth preservation. Note that the following principles are classically but not dialetheically equivalent:

(TP1) Inference rules are truth-preserving iff they always take us from true premises to true conclusions.

(TP2) Inference rules are truth-preserving iff they never take us from true premises to untrue conclusions.[1]

If the premise and conclusion of a single-premise argument are both dialetheias, then the inference from one to the other will be truth-preserving according to (TP1) but not (TP2). This will be precisely the situation when we infer from some dialetheia α to the disjunction (α ∨ β) (where β is *just* untrue), provided that both of the standard principles governing the truth conditions of disjunctions hold.

(TCD1) A disjunction is true iff at least one disjunct is true.
(TCD2) A disjunction is untrue iff all of the disjuncts are untrue.

These too are at least classically equivalent. It is not obvious why or how they would come apart on a dialetheic picture (or why similar pairs principles about other truth functions would come apart on such a picture).[2]

The problem is that if the dialetheist accepted (TP2), *no* inference would come out valid. Even the inference from α to α would not be universally truth-preserving. After all, on the line of the dialetheic truth table where α was both true and untrue, the inference would take us from a true premise to an untrue conclusion!

A dialetheist could hold onto (TP2) and try to resist this consequence by denying that truth value "gluts" are literally *both* true and false. Rather than the third "glutty" truth value in the dialetheist's truth tables literally being a conjunction of truth and falsity (which would then entail each of its conjuncts) it could be conceptualized as something distinct from both. Perhaps, in the same way that mixing red paint with green paint results in yellow paint rather than paint that is simultaneously red and green, when truth is somehow "mixed" with falsity, the result is a new truth value rather than a conjunction of standard ones.

[1] In the subsequent discussion, I take the question of whether some inference "preserves truth" (however exactly that notion is cashed out) to be equivalent to the question of whether it is valid, but this could be questioned. Hartry Field has argued (notably in Field, 2009) that valid inferences sometimes fail to preserve truth—a conclusion forced on him by his treatment of Curry's Paradox. Although he softens the blow of this by insisting that validity at least preserves truth "when it matters," since the only cases in which it fails to do so have to do with "certain degenerate instances" of valid inferences whose premises "involve 'ungrounded' occurrences of predicates such as 'True'." The reason he thinks these failures of truth preservation do not "matter" is that such premises "could never be established or be rationally believable" (Field, 2009, p. 266). Thus, even on Field's account of validity where validity is not analyzed in terms of truth preservation (or anything else) but regarded as a primitive concept that governs our reasoning practices, and where conditionals of the form "if the premises are true, so is the conclusion" are rejected when it comes to some valid inferences, it is still not possible for a valid inference to *actually take us* from true premises to false conclusions. Thus, even this account could not help a dialetheist pry apart validity from truth preservation when it comes to dialetheias—which are, after all, *true*.

[2] Perhaps the dialetheist could propose (TCD3), *A disjunction is untrue iff all of its disjuncts are (just) untrue*, but this seems arbitrary. It is one thing to say that a disjunction is *just* untrue iff all of its disjuncts have the same status, but why should *just* untruth (as opposed to any old kind of untruth) be a condition for the whole disjunction having any old kind of untruth?

The first thing to notice about this is that if the dialetheist made this move, it would no longer make sense for them to describe the third value as "glutty." The sentences in question would not have a "glut" of truth values. They would have exactly one, like all other sentences!

A second and much more important point is that the Liar Paradox manifestly cannot be solved by saying that sentences like (1) and (2) have a singular truth value that emerges from some sort of (non-conjunctive) "mixture" of standard truth values. If (1), for example, has the non-conjunctive "mixed" value, then it is true that it has a value other than truth. . .which makes it true, and (1) now has a conjunction of values rather than merely a "mixture" of them. Moreover, if a "mixture" is a distinct thing from a conjunction of values, it is no longer clear that the standard Liar reasoning rehearsed in 1.1 can deliver the result that paradoxical sentences have the "mixed" value. The link between Liar reasoning and the claim that such sentences have the "mixed" value could be preserved if we postulate that sentences like (1) are literally conjunctively true *and* false, and that having this conjunction of values entailed *also* having the "mixed value," but also that the entailment does not work the other way around—but this seems *ad hoc*. If sentences *can* be conjunctively true and false, what remaining motivation is there for postulating the "mixed" value?

Although they do not use the "mixing" metaphor I have employed to try to make sense of the dialetheist's third truth value failing to inferentially "detach" into truth and falsity, I owe the suggestion that there are "detachable" and "undetachable" readings of the "both" value to Arenhart and Melo (2017) (and the anonymous reviewer who directed me to their paper). Arenhart and Melo treat the "undetachable" reading as implicit in the formalism of Priest's LP, and the "detachable" reading as implicit in his informal/intuitive discussions of the paradoxes, but the point can be made more sharply by noticing that nothing about the formalism of LP *commits* Priest to undetachability. "Both" treated as a third value (or a third "relation" between sentences and truth in Priest's preferred language), but this by itself tells us nothing one way or the other about whether B<α> entails or fails to entail Tr<α> (or F<α>). Moreover, if it does not, in what sense are such sentences "both" true and false rather than simply a new thing *rather* than being true or false? Any dialetheist seeking to escape from the problems posed by combining dialetheism with (TP2) through embracing undetachability would seem to be sacrificing the intuitive point of a dialetheic solution to the paradoxes—the idea that we can simply embrace the surface impression that the combination of bivalence and "Liar reasoning" naturally lead us to say that such sentences have *both* values.

A far simpler solution to this problem would be for the dialetheist to simply reject (TP2) in favor of (TP1) as their preferred interpretation of truth preservation. Rejecting (TP2), however, will not suffice to get the dialetheist out of some thorny related problems. On the orthodox picture, (TP2) is inextricably bound up with the validity of inference rules such as *Reductio Ad Absurdum* (RAA).[3] Since valid

[3] I say rules *like* RAA, since possible dialetheic arguments against the unrestricted validity of RAA would presumably also apply to *Modus Tollens*. See below.

inferences cannot take us from true premises to untrue conclusions, when we use a rule we know to be valid to infer to some conclusion we know to be untrue, we can therefore know that at least one of the premises that got us there must be untrue as well. What becomes of RAA on the dialetheic picture?

4.2 Dialetheism and the Status of *Reductio Ad Absurdum* and *Modus Tollens*

Two related issues should be separated here. The first is whether RAA comes out as *deductively valid* in a dialetheic framework. The second is the one raised by Alan Musgrave in his paper "Critical rationalism and logical omniscience."

> If a contradiction may be true, then a *reductio ad absurdum* argument loses its force. And if philosophers cannot invoke *reductio ad absurdum* arguments, then they (and others) lose their chief critical weapon (Musgrave 1999a, b, p. 338).

To see how the two issues might come apart in a dialetheic framework, recall the distinction we explored in Chap. 1 while exploring Hartry Field's paracomplete theory of the paradoxes. On either of the two non-classical approaches under consideration here, *rejecting* a proposition comes apart from *asserting its negation*. The concepts come apart in one direction for the paracompletist, who sometimes sees fit to reject *both* a proposition and its negation, and in the other direction for the dialetheist, who will sometimes assert the negation of a proposition without rejecting it. Thus, a dialetheist might accept that RAA is valid without *rejecting* every proposition α that entails a contradiction ($\beta \wedge \neg\beta$). In one of his surprisingly few pronouncements on the *validity* of RAA, in *Doubt Truth To Be A Liar*, Priest seems to endorse exactly this approach.

> The purpose of a *reductio* argument is often to establish something of the form $\neg\alpha$ by deducing a contradiction from α. Dialetheism need not affect this enterprise. [...] In a polemical context, the point of a *reductio* argument is not normally to establish something, but to force an opponent to give up a view. In a dialetheic context, establishing the negation of a view is not *logically* sufficient for this (Priest, 2006a, p. 86).

The possibility that a dialetheist could use RAA to establish the negation of a proposition without rejecting the proposition itself would seem to trigger Musgrave's worry about RAA as a "critical weapon." Priest reassures us that this worry is misplaced.

> [W]hilst a contradiction may be *logically* possible, it does not at all follow that it may be rational to believe it. That I will turn into a fried egg tomorrow is logically possible, but a belief in this is grounds for certifiable insanity. (*A fortiori* that I both will and will not turn into a fried egg, since this entails it.) And an argument against an opponent who holds α to be true *is* rationally effective if it can be demonstrated that α entails something that ought, rationally, to be rejected, β. For then it follows that they ought to reject α (Priest, 2006a, p. 86).

Priest, then, seems to affirm both the *validity* of RAA and its continued usefulness as a "critical weapon," although he separates the issues. Variations on point about *usefulness* are sprinkled throughout his writings. Here it is, for example, in *In Contradiction*:

> A view can be criticized and made untenable if it can be shown to imply something that is rationally rejectable; for anything that implies something rationally rejectable is itself rationally rejectable. This is essentially how arguments by *reductio ad absurdum* work. Nor does the rationally rejectable consequence have to be a contradiction—anything that is rationally rejectable will do. For example, 'People turn into frogs when tapped on the shoulder' is as good as most contradictions, and better than some (Priest, 2006b, pp. 104–105).

. . .and in his paper "What is so bad about contradictions?"

> A person's views may be rationally criticized if they can be shown to entail something that is rationally unacceptable. This might be a contradiction, but it might be some non-contradiction. Some non-contradictions, e.g. that I am a fried egg, might be better than some contradictions, e.g. that the Liar sentence is both true and false (Priest, 2004, pp. 34–35).

He has far less to say about the validity issue. Even in the passage quoted above, he leaves himself some wiggle room, saying only that dialetheism "need not" affect the validity of RAA. His subsequent explanation is similarly qualified.

> If the deduction establishes that $\alpha \rightarrow (\beta \wedge \neg\beta)$, and the $\rightarrow$ in question contraposes, then we have $(\beta \vee \neg\beta) \rightarrow \neg\alpha$ by contraposition. [. . .] $\neg\alpha$ follows by LEM (Priest, 2006a, p. 87).

Priest's adherence to the LEM is one of his firmest philosophical commitments, but what about Contraposition? When Priest considers the thorny question of how various conditionals could be added to his logic LP in 1987's *In Contradiction*, he takes it for granted that even if conditionals fail to contrapose in *some* contexts, whatever conditional we use to define *entailment* must contrapose.

> For not only do we use the fact that something is entailed by true sentences to refute it, but we use the fact that something entails false ones to refute it. Thus, we require an entailment to preserve falsity from consequent to antecedent too. Of course, classically, truth preservation forwards and falsity preservation backward go together. However, once truth and falsity are seen to be partially independent, this is no longer the case. We therefore need to specify that both truth and falsity be preserved in the appropriate direction. [. . .] [T]he necessary preservation of truth and falsity in the appropriate directions is a necessary condition for entailment (Priest, 2006b, p. 84).

In 2006's *Doubt Truth To Be a Liar,* on the other hand, all Priest says about contraposition outside of the paragraph affirming that RAA "need not be" considered invalid by dialetheists is that some conditionals have this feature and others do not.

> Indeed, the claim that there are different kinds of conditionals (entailments, causal conditionals, indicative conditionals, subjunctive conditionals) is well-known. [. . .] A minimal condition for a conditional of any kind would seem to be that it preserve truth in an appropriate way from antecedent to consequent. Hence, we have *modus ponens*, α, $\alpha \rightarrow \beta \vdash \beta$. The most important question concerning a conditional in the present context is whether it preserves falsity in the reverse direction. For some conditionals, at least, this would seem

> to fail [...] Of course, there may well be conditionals that do preserve falsity in the appropriate way; in fact, one can always define one, $\Rightarrow$, in a simple fashion: $\alpha \Rightarrow \beta$ is just $(\alpha \rightarrow \beta) \wedge (\neg\beta \rightarrow \neg\alpha)$. For such a conditional contraposition will hold (Priest, 2006a, p. 82).

Once again, notice the non-committal language. There "may well be" contraposable conditionals. That is a far cry from his insistence in 1987 that entailment conditionals at least must contrapose. Nor does the point about $\Rightarrow$ do much to change his general posture of studied neutrality. Surely the substantive issue is about the ordinary uses of "if" and/or philosophically important cases such as Tarski's biconditional about truth. If the only conditionals that contrapose are ones stipulatively defined into existence to do so, it seems fair to gloss the situation as "contraposition is invalid." Compare: *Denying the Antecedent is invalid. . . . unless by* $\rightarrow$ *one means* $(\alpha \rightarrow \beta) \wedge (\neg\alpha \rightarrow \neg\beta)$. The italicized sentence is true enough, but what does it add to an unadorned claim that *Denying the Antecedent is invalid*?

Priest's fellow dialetheist JC Beall repeatedly indicates his own willingness to jettison contraposition in his book *Spandrels of Truth*.[4] Beall's sole concern about contraposition there is his desire to stop his dialetheic theory from slipping over into outright triviality, but there are reasons short of triviality-avoidance for dialetheists to reject contraposition.

To see why, start with the way Priest separates the issues of RAA's validity from the use of "*reductio* arguments" as a critical weapon. His idea is that when we—and that is a very broad "we," encompassing both classical and non-classical logicians, and philosophers who specialize in other areas, including ones who could not explain the difference between paracompleteness and paraconsistency to save their lives—use *Reductio* arguments as a "critical weapon," we do not require that the absurd consequences we extract from our opponents' positions be logically inconsistent. We simply need those consequences to be rationally unacceptable. When an ethicist rejects utilitarianism on the grounds that utilitarians have absurd commitments in various hypothetical examples about trolleys and transplants and the like, she does not regard the utilitarian as committed to *logically impossible* conclusions, but merely *unacceptable* ones.

Of course, dialetheists and non-dialetheists will differ about *which* results are rationally unacceptable. Non-dialetheists regard *all* contradictions as rationally unacceptable, while dialetheists think that some of them are rationally acceptable. For example, take sentence (7).

(7) It is not the case that the sentence written in blue marker on the whiteboard at the front of Room 214 of Scott Hall at the New Brunswick campus of Rutgers is true.

If (7) itself is the sentence that has been written in blue marker on the whiteboard at the front of Room 214 of Scott Hall at the New Brunswick campus of Rutgers, then Priest and other philosophers who embrace a dialetheic solution to the semantic

[4] See Beall 2009, p. 110, p. 121, p. 124, and p. 131. On p. 131, he comes frustratingly close to analyzing the underlying conceptual issues, and backs away from the precipice, remarking simply that, "Contraposition is contestable, and I shall not dwell on that issue."

paradoxes may consider a contradiction about the truth value of (7)—"it both is and is not the case that (7) is true"—to be not only rationally acceptable but *true*. On the other hand, we can identify at least three broad categories of contradictions that even dialetheists can be expected to find rationally unacceptable:

(i) Contradictions that entail triviality.
(ii) Contradictions with rationally unacceptable conjuncts.
(iii) Contradictions concerning domains which we can safely regard as contradiction-free.

Of the three, (i) is most obvious, since dialetheists typically see separating the premise that *some contradictions are true* from the classically derivable conclusion that *everything is true* as a core component of their project. Moving on to (ii), it is fairly straightforward that if "People turn into frogs when tapped on the shoulder" is rationally unacceptable, so is "People turn into frogs when tapped on the shoulder and it is not the case that people turn into frogs when tapped on the shoulder." That brings us to (iii).

As we have already seen, different dialetheists have very different ideas about which domains are dialetheic. Some, such as JC Beall, solve the Liar Paradox dialetheically but firmly reject the possibility of true contradictions in the "base language." (In other words, such dialetheists think that no sentences in which "true," "false," or similar words are not used can be dialetheias.) On the other end of the spectrum are dialetheists who embrace paraconsistent accounts of vagueness, thus postulating true contradictions in any domain where vagueness problems arise, which is to say pretty much every domain. Priest himself has occupied various positions in this spectrum at different points of his career.

Even so, we can produce at least one example that no dialetheic philosopher I am aware of at any point in this spectrum would regard as a rationally acceptable contradiction (and not because it fits into categories (i) or (ii)). Let us go back to the whiteboard in Room 214. Even if "it both is and is not the case that (7) is true" is rationally acceptable, the same surely cannot be said for "(7) both has and has not been written in blue marker on the whiteboard in Room 214."

But wait! Given a dialetheic analysis of the paradoxes, both conjuncts of the dialetheically unacceptable contradiction follow from the claim (7) has been written in blue marker on the whiteboard in Room 214. Thus, in a scenario where the first conjunct of the dialetheically *unacceptable* contradiction is true—and far from this being an implausible thought experiment, I can personally testify that I have written paradoxical sentences on that very whiteboard in markers of various colors!—the dialetheist should accept all three of the following claims.

1. If (7) has been written in blue marker on the whiteboard of Room 214, (7) is true.
2. If (7) has been written in blue marker on the whiteboard of Room 214, it is not the case that (7) is true.
3. (7) has been written in blue marker on the whiteboard of Room 214.

Given a logic with *Modus Ponens*, *Modus Tollens*, and Conjunction Addition, the dialetheically unacceptable contradiction follows in three easy steps. Of course,

Priest rejects *Modus Ponens* for the *material* conditional. (He has to, given his other logical commitments. If $\alpha \rightarrow \beta$ is materially equivalent to $\neg\alpha \vee \beta$, and Double Negation also holds, then the failure of $\neg\alpha \vee \beta$ and $\neg\neg\alpha$ to entail β by Disjunctive Syllogism entails the failure of $\alpha \rightarrow \beta$ and α to entail β by Modus Ponens.) All of this might give Priest an out. If he could argue that 1. is true if read as a material conditional but false if read as some other kind of conditional, he is off the hook. But is that plausible? On what grounds could it be argued that 1. is false if read as a conditional of some stronger variety? There is certainly no failure of relevance here, for example. If he wants to keep *Modus Ponens* for *any* kind of conditional, he has a problem here.

To see how this unsavory result generalizes, recall the traditional definition of "paradox" as an argument from premises which look like they *must* be true to a conclusion that *can't* be true. Generally speaking, a non-dialetheic solution will involve the counterintuitive claim that one of the premises—perhaps an implicit one—is *false*. A dialetheic solution avoids this particular cost, buying the ability to avoid having to say anything counterintuitive about any of the premises by saying something counterintuitive about the *conclusion*—that it is true. Adjudicating the comparative plausibility of dialetheic and non-dialethic solutions is a matter of deciding which of these costs is greater. . . .

. . .unless the dialetheist's entailment relationship contraposes, in which case the dialetheist is simply biting both bullets for the price of one.

4.3 Alan Weir on Dialetheism and Falsity Preservation

Alan Weir has argued, on slightly different grounds, that Priest's attempt to save contraposition in his discussion of entailment in *In Contradiction* is a bad fit with Priest's commitment to dialetheism. He regards this as a serious flaw of dialetheism. After all, falsity preservation backwards is as central to the role of logic in ordinary reasoning as truth preservation forward.

> One does not have to be a rampant Popperian to see that refutations do play an important role in science [. . .] If a theory entails a false prediction, the theory is false (though we might well not know which bit of it is false); more generally to the degree that we are sure the prediction is false we should be sure the theory is.

> Graham Priest seems to acknowledge this key role of falsity-preservation by building it into his account of an entailment conditional (1987:105). But if the $\rightarrow$ genuinely does represent entailment ($\models$) in the object language then this should hold:
> $\models P \rightarrow Q \leftrightarrow P \models Q$
>
> Yet where $\models$ is the logic LP, the one most utilized by Priest, this biconditional fails (Weir, 2004, p. 390).

At this point in our discussion, a charitable reader could be forgiven for thinking that Weir is being uncharitable to Priest. After all, did we not just see Priest saying

that entailment was not "the" conditional, but rather one of many conditionals, governed by all sorts of different logical rules?

We will have to wait for our discussion of Curry paradoxes in 5.1 to cash this particular check, but for the record: Weir's point here survives Priest's pluralism about conditionals. Given Priest's solution to Curry, he cannot accept that *any* conditional (or at least any conditional for which the T-Schema holds) honors both sides of Weir's biconditional.

In any case, Weir has a deeper reason for thinking that Priest is not going to be able to capture the role of falsity preservation in philosophical, scientific and everyday reasoning.

> LP, considered algebraically, is the three-valued strong Kleene logic in which the 'top' two values are designated. For Priest the values are True {T}, Both {T, F} and False {F}. Hence a premise that goes from a True premise to a 'glutty' conclusion that takes the value Both is not thereby shown to be invalid, since we have truth in the premise and truth in the conclusion; but an argument that goes from Both to false is invalid, since we have truth in the premise but not in the conclusion (Weir, 2004, p. 390).

If this result were just an artifact of the technical details of LP, one might see Priest's difficulties with falsity preservation as a flaw in his current formal model of dialetheism rather than as a flaw in dialetheism as a philosophical position. Unfortunately, *no* formalization of the dialetheic picture can capture the intuition to which Weir is appealing without disastrous consequences. As we saw in our discussion of (TP2) in Sect. 4.1, the combination of dialetheism and a definition of validity that renders every inference from a true premise to a false conclusion invalid renders the concept of validity empty. Even the inference from α to α comes out as invalid. The situation is at least as bad if we try to build falsity preservation directly into our definition. Here is Validity Defined Disjunctively:

(VDD) An inference is valid iff it always *either* takes us from (all) true premises to a true conclusion *or* from a false conclusion to (at least one) false premise.

Given the simple truth tables of classical logic, (VDD), (TP1), and (TP2) are three different ways of saying the same thing. The distinctions make a difference once we expand the picture to include gluts.

(VDD) captures Weir's intuition just as well as (TP2), but it avoids the unpleasant consequence that no argument forms can ever be valid. Even so, it has an equally disastrous consequence from a paraconsistent point of view. The inference from $(\alpha \wedge \neg\alpha)$ to $(\beta \wedge \neg\beta)$ comes out as valid! Given dialetheism, the options are that the premise and conclusion are both just false, that they are both dialetheias, that the premise is just false while the conclusion is a dialetheia, and that the premise is a dialetheia while the conclusion is just false. *All four possibilities satisfy the second disjunct of (VDD).* This simply sneaks the explosion principle in the backdoor. If all contradictions follow from contradictions, then by Conjunction Elimination *everything* follows from contradictions.

4.4 Why Successful Classical Recapture Is Vital to the Dialetheic Project

Dialetheist readers, or even "di-curious" ones who possess some sympathy for Priest's project but are not yet committed to dialetheism, might wonder if things are quite as bad as all that. After all, even if I am right that the dialetheist's only way out of the dilemma about validity and truth preservation is to cast *Modus Tollens* and *Reductio Ad Absurdum* into the fire right along with Disjunctive Syllogism and the other rules that break down once we allow for the possibility of true contradictions, can we not have them all back in some—indeed *most*—contexts?

The idea here is what is often referred to as "classical recapture." Non-classical theorists often seem to take it as a matter of course that whatever classical principles we give up to deal with problems that arise in one domain can still play their customary roles in guiding our reasoning in other domains. Does the distribution of conjunction over disjunction fail to hold for propositions about the behavior of subatomic particles? The quantum logician should still feel free to Distribute to her heart's content when reasoning about macroscopic objects. Do we need to reject instances of the LEM to deal with the Liar Paradox? No problem. The paracomplete logician can still exclude every middle in sight if he is reasoning about "grounded" sentences.

Graham Priest has proposed multiple strategies for "recapturing" classical logic. I will argue in what follows that none of these strategies succeed, and that classical recapture projects in general face insurmountable obstacles. Nevertheless, Priest deserves credit for being a non-classical theorist who *realizes that there's a problem* when it comes to how to justify the use of classical reasoning after jettisoning the assumption that such reasoning universally preserves truth.

That said, the problem would seem to go much deeper than even Priest has acknowledged. If the dialetheist says that paraconsistently invalid classical rules can be reintroduced in "consistent contexts" or "contexts in which no contradictions are true," this might sound intuitively reasonable—but it assumes that the dialetheist can even express these concepts. As we will see when we get to the dialetheist's difficulty in expressing the notion that some sentences are "just" false, this is not as safe as an assumption as it might seem to be. For now, notice that it certainly will not suffice for the dialetheist to say that in a given domain, $\neg(\alpha \wedge \neg\alpha)$ holds for every α, since $\neg(\alpha \wedge \neg\alpha)$ is dialetheically compatible with the truth of $(\alpha \wedge \neg\alpha)$![5]

Although even assuming for the sake of argument that the dialetheist has some way of expressing the idea of a "consistent domain," classical recapture will be extremely difficult. To start to see the scope of the problem, notice first that the rules

[5]Thanks to the anonymous reviewer for making this point! Along somewhat similar lines, see Woods, 2019. Woods argues that when Priest or his non-classical colleagues attempt to recapture classical logic in certain domains, they are necessarily doing it in a "classical metalanguage"—in other words, they are reasoning *about* which logic is appropriate for which domains in ways that only really make sense given classical assumptions.

the dialetheist must regard as invalid—Disjunctive Syllogism, *Modus Tollens*, RAA—are among the deductive inferences most often used in scientific, mathematical, philosophical and day-to-day reasoning. Ask a mathematician how we can be sure there is no largest even number, and she will tell you to assume for the sake of argument that there is such a number, E, and add two to it. Now we have E+2, an even number larger than the largest even number, which is impossible, so our original assumption was wrong. That is *Reductio Ad Absurdum*. The refutation of scientific theories by experimental evidence (and the clearing of criminal suspects through alibis) is *Modus Tollens.* As to Disjunctive Syllogism, those readers most familiar with these debates will have to forgive me since, with my miniature schnauzer Lucy curled up by my feet as I write this, I cannot resist the temptation to trot out the most well-worn example in the literature.

> At Cambridge in 1615, the claim that dogs use logic was defended by John Preston (1587–1628) of Queen's College. "He instanced in a hound who hath the major proposition in his mind, namely, *The hare is gone either this way or that way*; smells out the minor with his nose, namely, *She is not gone this way*; and followed the conclusion, *Ergo this way*, with open mouth." The inference which the dog is purported to have followed is *disjunctive syllogism*, which we might abbreviate as "*P* or *Q*, not-*P*, therefore *Q*." Preston was answered by Matthew Wren (1585–1667) of Pembroke College, for whom dogs were distinguished by the excellence of their noses, not their reasoning: surely the dog determined the quarry's path by scent. While the moderator, Simon Reade of Christ College, agreed with Wren, the King [James VI of Scotland and I of England] took Preston's side, gave an example of a reasoning dog from his own experience, and suggesting that Reade should "think better of his dogs or not so highly of himself." Reade successfully mollified his sovereign with the suggestion that dogs that hunt by royal prerogative must be exceptions to laws that govern common hounds, and the debate concluded in good spirits (Aberdein, 2008, p. 168).

Historically, many people have thought that dogs did just what Wren thought they did. (Various ancient Greek sources attribute this argument to the Stoic philosopher Chrysippus of Soli, whose own account is sadly lost.[6]) More recently, there is at least some experimental evidence for the use of just this sort of "exclusion procedure" by non-human animals, although a contemporary heir to Reade's argument might grant the use of such a procedure but question whether it can be meaningfully described as an instance of propositional reasoning.[7] While the finer points of canine cognitive science lie outside of the scope of this paper, the mere fact that it is so intuitively natural assume that dogs "use Disjunctive Syllogism" in this way is instructive. DS is so basic to our everyday reasoning that it is difficult for us to imagine anything even *resembling* human reasoning that does not involve some form of it.

JC Beall and Greg Restall seem to be on Wren's side of the 1615 argument. They discuss the example in the context of their discussion of relevance logics. (Note that

[6] See Aberdein pp. 170–171 for a summary of the Greek sources.

[7] Andy Coghlan describes a study involving grey parrots. First, treats were hidden under opaque cups. "Next the researcher hid the cups behind a screen, removed the treats, and showed the bird which one had been taken. Finally, the screen was removed to see if the parrot could work out which treat must remain, and under which cup it must be" (Coghlan, 2011). The cleverest parrot, a female named Awisa, was able to perform the deduction correctly 23 of 30 times.

they follow the Australian convention of calling what other English-speaking logicians call "relevance" logics "relevant" logics. As an American, I can only regard this as a quaint Australian eccentricity on par with vegemite-eating, crocodile-hunting, and Australia's vastly superior health care system.) While Priest's logic LP is not a form of relevance logic, all that matters for our present purposes is that both dialetheists and relevance logicians regard DS as invalid.[8]

> It might be thought that this is not *really* a worry for proponents of relevant logics because, even though disjunctive syllogism isn't relevantly valid, disjunctive syllogism is not all we have at play here. After all, the DOG does not merely know that A ∨ B and know that ~A. The DOG also knows that it's not the case that A and ~A. This might be the case. However, it's not at all obvious how this is going to help, because the argument
>
> $$\frac{A \vee B \sim A \sim(A \wedge \sim A)}{B}$$
>
> is no more relevantly valid than is disjunctive syllogism. (After all, the third premise ~(A ∧ ~A) is relevantly entailed by the second premise ~A [. . .] so adding it really does not do much for the argument. [. . .] On the other hand, one could add a second conclusion to the argument, giving this:
>
> $$\frac{A \vee B \qquad\qquad \sim A}{B \qquad\quad \sim(A \vee \sim A)}$$
>
> which is relevantly valid. Any situation in which A ∨ B and ~A are true is one in which either B or A ∧ ~A are true (Beall & Restall, 2006, pp. 95–96).

But wait! To get from "B or A ∧ ~A" to our desired conclusion B, do we not need another instance of Disjunctive Syllogism? Have we not just reinstated the original problem after adding a layer of unnecessary complexity? Well. . .

4.5 The Failure of Classical Recapture, Part I (The Failure of Principle R)

This brings us to the first of the classical recapture strategies Priest proposes in *In Contradiction*. This particular strategy is worth examining in some detail both for its own sake and because both Field and Beall seem to be confident that something along the same lines will work. (More on that later.) Priest proposes in effect that, while a

[8] Of course, the same philosopher can be both a philosopher and a relevance logician, and that bit of overlapping logical space has multiple real-world representatives, but the relevance logician *qua* relevance logician has a slightly different reason for wanting their logic to be paraconsistent than the dialetheist *qua* dialetheist. To see the difference, notice that while both types of logicians want to block the inference from violations of the LNC to arbitrary conclusions, relevance logicians also disallow the (dialetheically innocent) inference from arbitrary premises to instances of the LEM. Without engaging the nuances of the philosophical motivations for relevance logic, it is not too much of an oversimplification to parse the distinction by saying that dialetheists reject the "explosion" of inferences classically derivable from contradictions because they think most such inferences *fail to preserve truth*, while relevance logicians reject explosion on the grounds that *truth-preservation is not enough* for one proposition to truly "follow" from another.

dog (or a cognitively superior creature in the same epistemic position) may not be justified in using DS, they *would* be justified in using what he calls Principle R.

Principle R

If a disjunction is rationally acceptable and one of the disjuncts is rationally rejectable, then the other is rationally acceptable.

Suppose that someone accepts a disjunction, $\alpha \vee \beta$. Then nothing, as yet, forces them to accept either disjunct: they may be agnostic about both. Let us suppose that they now come to reject one of the disjuncts, say α, while continuing to accept the disjunction. Then the person is rationally committed to accepting the other, β. The rationale for this is precisely the truth condition for disjunction: $\alpha \vee \beta$ is true iff α is true or β is true. So if one limb of the disjunction is not on, the other must be.

It is important to see that this is not an application of DS. The argument is not: $\alpha \vee \beta$ is true but α is not true; therefore β is true. The justification is not a formal one but a pragmatic one: $\alpha \vee \beta$ is rationally acceptable; β is rationally rejectable; hence β is rationally acceptable. If this distinction is not clear, observe it in another context. If someone accepts that $\alpha \rightarrow \beta$ and accepts $\neg\beta$, then they ought to accept $\neg\alpha$. The rationale for this is precisely the formal validity of *modus tollens*: $\{a \rightarrow \beta, \neg\beta, \vdash \neg\alpha\}$. If, on the other hand, she accepts $\alpha \rightarrow \beta$ and rejects β, she had clearly better reject α. But this is not *modus tollens.* As we observed [...earlier in the book,] there is a crucial difference between rejecting α and accepting $\neg\alpha$ (Priest, 2006b, pp. 113–114).

This last point may not be entirely persuasive. After all, while accepting $\neg\alpha$ and rejecting α may indeed be different psychological states, that is not all that Priest needs here. He needs the *circumstances under which it is rational to accept* $\neg\alpha$ to come apart from *the circumstances in which it is rational to reject* α. For a partisan of either gaps or gluts, there is indeed such a distinction. For the defender of classical orthodoxy, there is not. To the extent that the purpose of Priest's argument for the possibility of classical recapture is to persuade those who are not yet convinced of dialetheism (perhaps because, although they see the advantages of a dialetheic solution to the paradoxes, they worry about the counterintuitive consequences of declaring DS to be invalid), it may not be legitimate for Priest to help himself to that distinction.

Let us put aside that worry for now and examine the work that Principle R does for Priest in this passage. At the very least, Principle R does not get the dialetheist into trouble in the obvious way that DS does. If α is true and false and β is (just) untrue, Principle R will not be triggered, for the simple reason that, from the dialetheist's perspective, if we *know* that α is both true and false, the mere fact that it is false will not suffice for α to be rationally rejectable.

Returning to Beall and Restall, this is precisely how they (tentatively) suggest that a relevance logician could get from the relevantly valid multiple conclusion argument given at the end of the previous section to the intuitively clear single conclusion B.

If we have, therefore, a warrant for the DOG to ***reject*** $A \wedge {\sim}A$ (as opposed to a warrant for the DOG to ***accept*** ${\sim}(A \wedge {\sim}A)$), then perhaps we have a relevantly valid argument that goes some way to representing the DOG's reasoning (Beall & Restall, 2006, p. 96).

They go on to raise worries about whether this is the best way to reconstruct the dog's reasoning. For one thing, even Wren and his dog-loving King might worry that

this stretches the limits of the reasoning powers we can safely attribute to non-human animals. To see the point in sharper relief, let us move from dogs to humans. Recall one of the famous instances of DS in popular culture. In Stanley Kubrick's classic film *Full Metal Jacket*, an abusive Marine drill sergeant, Gunnery Sergeant Hartman, taunts Texan recruit Leonard Larence ("Private Pyle") by telling him that "only steers and queers come from Texas [. . .] and you don't like a steer to me, so that kind of narrows it down." We have no reason to think that Hartman, Lawrence, or any of the other recruits silently witnessing the homophobic taunt have studied formal logic, but no one has any trouble telling what conclusion Sergeant Hartman has left hanging in the air. There is nothing unrealistic about that. Few bits of reasoning are as rudimentary as rattling off a list of options, eliminating all but one and embracing what is left. Although it is a bit more of a stretch to imagine that (even on some rudimentary, not fully conscious level) Hartman or any of his listeners are engaged in multiple-conclusion reasoning, first considering and rejecting the possibility of $(A \wedge \sim A)$ before settling on B.[9]

[9] Kevin Sharp raises a structurally similar point about another recapture strategy Beall has proposed ("shrieking") in his excellent and excellently titled paper, "Shrieking in the face of vengeance" (Sharp, 2018). Beall has suggested that the dialetheist's difficulties with saying that some sentences are "just true" or "just false" can be solved by adding "shriek" rules to the dialetheist's preferred paraconsistent logic. This amounts to saying that some sentences are "just" true where "just true" means "the sentence is true and the claim that it's false entails triviality." This essentially turns sentences of the form "This sentence is just false" into instances of Curry's Paradox, about which I will have more to say about in Chap. 5. Beall points out that it is a bit much to think that anyone who says "that's true" about an ordinary statement is unconsciously stipulating that the statement not being true would entail triviality. (Presumably, most ordinary speakers would respond to the suggestion that this is what they meant by asking, "What's triviality?") When we take a step back from the truth predicate and look at the bigger picture, this problem gets a lot worse. Beall plausibly claims that, in the vast majority of contexts, people mean every use of every predicate in the way that "just" is supposed to capture here. If Bob says "the cat is on the mat," he presumably means to exclude the possibility of the cat not being on the mat. But applying the "shrieking" strategy not just to "true" but to every predicate meant "consistently" gets us some bizarre results. He uses the example, "Bill is jumping rope slowly in the garage." Given dialetheism, this is compatible with "Bill is not jumping rope slowly in the garage," "Bill is jumping rope not slowly in the garage," and "Bill is jumping rope slowly not in the garage," but any ordinary speaker in any ordinary context presumably means to rule out *all* of these possibilities. If we apply Beall's shrieking strategy, we end up getting the result that what any ordinary speaker uttering this sentence would *really* mean is "Bill is *just* jumping rope *just* slowly *just* in the garage" (Sharp, 2018, pp. 458–459). This is loading a *lot* of baggage into every normal English sentence that it is not plausible that ordinary speakers are consciously or even unconsciously putting there. The anonymous reviewer who directed me to Sharp's paper puts the point very sharply: "*Classical recapture strategies seem to saddle us with the most complicated unconscious reasoning in order to ensure consistency. This is strange, given the dialetheists' commitment with the idea that natural language and reasoning processes should be held intact when they lead to the Liar, but not when we discuss about the weather or football. . .*" Moreover, the point arises with particular force with regard to Beall's work, since Beall has in another context said that the best reason to view some paradoxical sentences as gluts is "aesthetic"—"such target liars simply look like gluts" (Beall, 2015, p. 583). Not to put too fine a point on it, "Bill is jumping rope slowly in the garage" does not look like and is not most naturally treated as, "Bill is *just* jumping rope *just* slowly *just* in the garage."

Of course, if Principle R is there to get the DOG (or the homophobic drill sergeant) from the acceptance of a disjunction and the rejection of one disjunct to the acceptance of the remaining disjunct, a step can be skipped in Beall and Restall's presentation of the argument. The apparent user of Disjunctive Syllogism could be understood to be reasoning straight from the fact that "one limb of the disjunction is not on," where being "not on" is understood in terms of rejection rather than negation, to the conclusion that "the other must be." This reconstruction strategy, however, faces a different problem. Attributing a disjunction of intermediary conclusions and then an extra step to the apparent DS-user might or might not be implausibly complex, but it at least preserves the key facts of the case—that very often people start from the premises $(\alpha \vee \beta)$ and $\neg\alpha$ and end up at the conclusion that β. Tessa says that Duncan is either at home or at Charlie's dojo, and that he is *not* at home, so she concludes that he is practicing his martial arts skills at the dojo. If we parse Tessa's argument as consisting in nothing but a single application of Principle R, we are undermining the plausibility of the explanation by rejecting one of the apparent facts of the case. We are saying that Tessa *thinks* that one of the premises she is reasoning from $\neg\alpha$, but she is really reasoning from the non-equivalent *rejectability* of α.

So Beall and Restall's worry here seems like a good one, even when we shift the focus from our four-legged friends to adult humans fully capable of complex propositional reasoning. Both the paraconsistent reconstructions of "apparent" uses of DS that Principle R can be used to generate come with significant intuitive costs.

A third option might be to forgo the reconstruction project entirely and grant that Tessa is unjustified in believing her conclusion (even if she is justified in believing both of her premises), but point out that she *would* be justified if she had used Principle R instead. The real question, a dialetheist pushing this last option might argue, is not whether accepting dialetheism will lead us to revise our reasoning practices, but whether the revised practices will be sturdy enough for our day-to-day epistemic needs. Note, however, the sheer size of the bullet the dialetheist who responded in this way would be biting. Massive chunks of apparently correct deductive reasoning are being written off as not just *invalid*, which is bad enough, but as failing to count as good reasoning of any kind.

Still, if Principle R can at least deliver the result that a dialetheist can make sense of what she is doing when, for example, she narrows down the list of murder suspects to two, clears one of them based on a convincing alibi and concludes that the other one committed the crime, that is not nothing. Unfortunately, as we are about to see, Principle R simply does not do this work.

Chapter 5
Dialetheism, Rejection, and Curry's Paradox



Abstract The dialetheist proposes to block the inference from paradox to triviality not by blocking the inference from paradox to contradiction but by blocking the inference from contradiction to triviality. This strategy is undermined by Curry's Paradox, in which the inference goes from paradox to triviality without stopping at the waystation of contradiction. Priest's attempted solution is critiqued for several reasons, including its failure to honor Priest's own "Principle of Uniform Solution," according to which paradoxes of the same type should be solved in the same way. Beall's solution to Curry fares better on that score, but it turns out that it relies on Principle R, and it has already been shown that this principle cannot be made sense of on dialetheic assumptions. A much larger problem lurks in all of this for dialetheist (and paracomplete) attempts to use rejection-talk to do work that cannot be done by negation once classical orthodoxy is abandoned. This strategy is undermined by the failure of Principle R as well as further problems pointed out from within the dialetheic camp by Francesco Berto.

Keywords Priest · Curry · Berto · Rejection · Uniformity · Negation

5.1 The Broader Importance of the Failure of Principle R (I: Curry's Paradox)

Priest himself abandoned Principle R when the second edition of *In Contradiction* came out in 2006. We will look at his reasons in Sect. 5.2. But first, a review of the broader significance of this issue is in order. Without some sense of the stakes, a reader might be forgiven for wondering why so much of her time is about to be spent on a strategy proposed by Priest in 1987. . .and abandoned by Priest in 2006.

Some of those reasons will have to wait for a fuller discussion until later in the chapter, after we have gone over quite a bit of groundwork, but one of them can be given here. The most otherwise plausible dialetheist solution to Curry's Paradox implicitly rests on Principle R. If Principle R fails, as I argue that it does, so does the solution.

Curry's Paradox concerns sentences like (8).

© Springer Nature Switzerland AG 2022

B. A. Burgis, *Logic Without Gaps or Gluts*, Synthese Library 458,
https://doi.org/10.1007/978-3-030-94624-1_5

(8) If sentence (8) is true, *everything* is true.

When we feed (8) into the T-Schema, we get the following, using the standard "triviality" symbol ⊥ for "everything is true."

Tr<8> ↔ (Tr<8> → ⊥.)

In a few easy (and classically valid) steps, we can get from there to *every* conclusion!

1. Tr<8> ↔ (Tr<8> → ⊥) Premise
2. Tr<8> → (Tr<8> → ⊥) 1 Df. ↔
3. Tr<8> → ⊥ 2 Contraction
4. (Tr<8> → ⊥) → Tr<8> 1 Df. ↔
5. Tr<8> 3, 4 Modus Ponens
6. ⊥ 3, 5 Modus Ponens

This paradox is a source of embarrassment for dialetheists, since their standard move for blocking the inference from the truth conditions of paradoxical sentences to triviality seems to fall flat in the face of sentences like (8). Dialetheists accept the inference from such truth conditions to *contradiction*, and use the resulting allegedly true contradictions to block the inference from inconsistency to arbitrary triviality, arguing for example that the truth and untruth of sentence (1) and the fact that it is *just* that the moon is made of green cheese constitutes a counter-example to the classical principle that *ex contradiction quodlibet* ("from a contradiction, anything follows"). The problem is that the inference from the truth conditions of (8) to triviality is an express route. It does not make a stop at Contradiction.

In the actual world, Haskell Curry died five years before *In Contradiction* was published (and discovered his paradox decades before that), but we can be fairly certain that in quite a few of the closest possible worlds where Curry was not the discoverer of the paradox in question, it was formulated at some point after 1987 *as a revenge paradox for the dialetheist solution to the Liar Paradox*.

(8) certainly *looks* like a canonical Liar sentence like (1) or (2). Granted, appearances can be deceiving, but it is striking that all three of these sentences share all sorts of common paradoxical features. They are all self-referential, all "saying something" about their own truth value that, once fed into the T-Schema, classically entails contradiction. (There are variations of Curry that at least superficially avoid self-referentiality, but the same is true of the Liar! See the discussion of Yablo's Paradox in 8.5. In both cases, the variants that are not obviously self-referential are at the very least "ungrounded.") Worse yet, we can easily concoct a sentence that is *both* a Liar and a Curry sentence, like (9).

(9) If sentence (9) is true, sentence (9) is false.

Graham Priest and other dialetheists have solved Curry's Paradox by rejecting Contraction for the conditionals in their favored paraconsistent logics, thereby

blocking step 3 in the proof above.[1] One obvious problem with this move is that it does not seem to have anything to do with paraconsistency per se. Contraction plays no role in the inference from *contradictions* to triviality—only in the "express route" inference from the truth conditions of sentences like (8) to triviality. As such, this solution seems *ad hoc*.

In Beall, 2015, JC Beall raises the worry that it is not only *ad hoc* but severely undermines one of the chief motivations for dialetheic solutions to the paradoxes. Priest in particular has relentlessly emphasized the idea that dialetheism is preferable to other solutions to the paradoxes because it can promise a "semantically complete" theory of English. A "theory" in the relevant sense, derived from Tarski, is a set of sentences closed under entailment, and a "semantically complete" theory is one that (a) could express all the truths *about* English but also (b) be expressed *in* English. In other words, the dialetheist does not have to limit either which sentences can be expressed or what is entailed by those sentences.

At least some non-dialetheic solutions to the Liar Paradox sacrifice semantic completeness by insisting either that an unrestricted truth predicate be excluded or that some entailments of a theory that includes paradoxical sentences be arbitrarily excluded from the theory. Beall points out that given a Validity predicate, such that V(α, β) means that β can be validly inferred from α--crucially, a series of valid inference steps connecting α to β can be "captured" into V(α, β) and V(α, β) releases into the conclusion that there is some such series of steps—a Curry sentence be constructed using this predicate.

The dialetheist's choices seem to be to either make such a predicate inexpressible or to arbitrarily block certain inferences from such Curry sentences to block triviality. This is precisely parallel to the choices Priest takes to be faced by a non-dialetheist paradox-solver who can either exclude certain sentences from being formed with the *truth* predicate or arbitrarily limit the entailments of such sentences to avoid contradiction. As such, Beall concludes that the promise of having semantic completeness "at the mere price of gluts is a false promise" (Beall, 2015, p. 580). The price of semantic completeness would be triviality, which is unacceptable. Thus, Beall thinks, the dialetheist must join he classical logician in giving up on the dream of semantic completeness, and the only remaining argument for handling paradoxical sentences dialetheically is a basically "aesthetic" one—Liars "look" like gluts. He acknowledges that this does not rationally "force" a glutty solution to the Liar Paradox, but he does think it "motivates" one.[2]

[1] To cash a check written in 4.3, this is why Priest cannot accept both sides of the biconditional P → Q ↔P |= Q, which Alan Weir thinks is crucial for capturing the intuition that valid inferences should both preserve truth in one direction and falsity in the other, for any conditional for which the T-Schema holds.

[2] Priest's response to this paper in Priest 2016 is to argue that (a) the problems Beall brings up are already well-known and do not pose a new challenge to dialetheism, and (b) they can be solved through adopting a "substructural" logic for reasoning about Curry sentences. A key move is that some collections of sentences are best understood not as conjunctions but as "fusions," and that some inference rules that hold for conjunctions do not hold for mere "fusions." He does not give a

All this would be bad enough for the dialetheist solution, but the point can be sharpened by looking at it from another angle. If, as Priest thinks, the Liar and the set-theoretic paradoxes can be seen as *prima facie* sound arguments for dialetheism—such that the challenge for non-dialetheists is to find a mistaken premise or a mistake in the reasoning connecting those premises to the resulting contradiction—Curry's Paradox can be seen as a *prima facie* sound argument for *trivialism,* the claim that *absolutely everything* is true.

Seen in this light, Priest's approach to Curry—rejecting an otherwise plausible logical rule to avoid the paradoxical conclusion—seems question-begging in precisely the sense that objections to Priest's dialetheic solution based on pure contradiction-avoidance are (according to Priest) question-begging. What is the difference?

One could argue that the difference is just that it is impossible to beg a question against the trivialist, because the trivialist accepts everything, necessarily including the wrongness of trivialism. You cannot illegitimately assume what is in dispute because nothing is really "disputed" by the trivialist.

This would, however, be entirely too hasty. The epistemic reasons that begging the question is barred from rational debate do not cease to apply in this case. Priest himself has some very sensible things to say about precisely this subject in Chap. 3 of *Doubt Truth to Be a Liar* and when looking at how he fares on these grounds, it will be sufficient to compare his actual performance with the standards he himself lays out in that discussion.

What he asks us to do in that book is to imagine, when arguing against the trivialist, that we are not trying to convince the trivialist (who, after all, is already convinced of everything!), but rather that we are trying to convince a neutral arbitrator who has not yet made up her mind about the issue of trivialism, and who will (using normal standards of evidence) accept or reject arguments brought to her by both sides. Convincing the trivialist of the wrongness of trivialism is a trivial matter. Convincing the imaginary arbitrator is not. It is the latter, harder standard to which we should set ourselves.

> Convincing the arbitrator is [. . .] not as easy as it might first appear. [. . .] For example, an obvious argument to use is that the trivialist's position is inconsistent, and so ought not to be believed. (Naturally, a dialetheist is not going to make this objection; but a classical logician might.) It is clear that this argument fails, though, since it begs the question. The trivialist affirms [. . .] all contradictions. Simply to claim that these cannot be true is to take for granted part of what is at issue (Priest, 2006a, p. 66).

Things are not so hopeless as they may appear, however. In fact, there is excellent reason to be confident that a nontrivialist with normal beliefs about the world could make a lot of good, convincing anti-trivialist arguments of the kind that the hypothetical neutral arbitrator *would* allow.

---

philosophical rationale for this change of logic, and in the absence of a sufficiently convincing one, Beall's charge that Priest is simply arbitrarily blocking certain inferences in ways that break the promise of semantic completeness.

> The trivialist believes many strange things. They believe, for example, that you are a scrambled egg. This is objectionable since there is just no evidence to justify it. The rational person should not believe anything for which there is no reason. The arbitrator must agree. And if the trivialist argues that there is a reason, namely that it follows from trivialism, then it is they who now beg the question. The trivialist will also claim that there is independent evidence. (They will claim everything.) But claiming does not of course make it so. The evidence needs to be produced for the arbitrator. [. . .] (Priest, 2006a, p. 66).

This should be enough to get across Priest's general idea about how to argue against trivialism.[3] We agree that this is a reasonable standard for evaluating anti-trivialist arguments.

Although what will happen when the trivialist presents the neutral arbitrator with Curry's Paradox, which seems to show that trivialism is true in just the way that (according to both Priest and the trivialist) the Liar Paradox shows that dialetheism is true? Oddly, Priest nowhere mentions Curry in his discussion of arguing with trivialists in *Doubt Truth to Be a Liar* (Priest, 2006b, Chapter 3). Elsewhere in the book (2006b, pp. 99–101), however, he mentions the paradox to justify the rejection of Contraction in his favored logic LP (Logic of Paradox). A footnote directs the reader to a longer discussion of Contraction in *In Contradiction* (Priest, 2006a, Chapter 6). There, he explicitly frames the entire discussion in terms of "Curry-proofing" his conditionals.

If this were Priest's defense against the trivialist's argument-from-Curry, the neutral arbitrator could hardly be impressed.

Finally, Priest's solution raises awkward questions about the Principle of Uniform Solution (PUS)—same kind of paradox, same kind of solution. Given Modus Ponens, the price tag of applying the bullet-biting dialetheist response to the other paradoxes of self-reference to Curry and declaring (8) and (9) to be *true* would be abandoning the *raison d'être* of paraconsistency—the avoidance of triviality! As we saw in 2.3, Priest cashes out the sense in which all the "paradoxes of self-reference" count as the "same kind of paradox" in terms of the Inclosure Schema. As Otávio Bueno and I have argued elsewhere (Bueno & Burgis, 2019), this results in a dilemma for Priest and any dialetheist paradox-solver who follows his lead in this regard. Either Curry *is* an Inclosure Schema paradox, in which case, by Priest's lights, it must be solved in the same way as the Liar, or it is not, in which case so much the worse for the IS as a tool for demarcating paradoxes of the same type.

In other contexts, Priest has argued that, even if solutions to various paradoxes are all *individually* satisfying, they should be rejected if they fail the test of the Principle of Uniform Solution (PUS).

> The Liar Paradox comes in many varieties. In some of them, self-reference is achieved by means of a demonstrative ("this sentence," etc.); in others, it is attained by a description ("the

[3] Once this framework is established, Priest goes on to introduce various complexities to the scrambled egg argument, and to make other arguments against trivialism. Whether or not all of his arguments go through is beside the point for our current purposes. The important part of the discussion from *Doubt Truth to Be a Liar* for our purposes here is not Priest's anti-trivialist argumentation, but his conceptual framework for how anti-trivialist argumentation should proceed.

> first sentence on such-and-such a page," etc.); in still others, it is achieved by some sort of diagonal argument employing Gödel numbers as names. At some level of abstraction, these paradoxes are, presumably, of different kinds. The first depends on context-dependent features of utterance; the second presupposes some account of the way that descriptions work. The third employs numbers. Suppose that I propose three corresponding kinds of solution. The first kind of paradox is solved by some theory of speech acts; the second is solved by some theory of the semantics of description; the third is solved by an appeal to mathematical nominalism. Even if all the solutions were to work for their own kind, it appears to me undeniable that, collectively, there is something wrong with them. [. . .] The solutions may avoid the paradoxes, but they do so by attacking peripherals, not essentials (Priest, 2000, p. 124).

So is Curry an Inclosure Schema paradox or not? In the past, Priest has moved very delicately on this front.

> [. . .] whether or not Curry paradoxes fit the Schema depends on how the conditional involved is interpreted. If it is interpreted as a material conditional, Curry paradoxes fit the Schema. If not, they do not. Such paradoxes are, therefore, of a different kind (Priest, 1998, p. 387; see also Priest, 2002, pp. 168–169).

This is very weak sauce. After all, a classical logician who solved the Liar Paradox by arguing that all sentences implicitly "say of themselves" that they are true (so that sentences like (1) and (2) came out as mere contradictions, not paradoxes), but who solved the paradoxes of naïve set theory by denying the existence of sets outside of the cumulative hierarchy of ZFC, could avail themselves of exactly the same defense. *Whether or not the Liar Paradox fits the Schema depends on whether one takes all sentences to implicitly say of themselves that they are true. If not, the Liar does fit the Schema. If so, it does not. . .*

Even the hypothetical paradox-solver Priest used to illustrate PUS violation above could offer up a precisely similar explanation for how she solved all *real* Inclosure Schema paradoxes in a uniform manner. *If such-and-such theory of speech acts were right, the version of the Liar that started with the words "this sentence" would be an Inclosure Schema paradox. If not, not. If a certain picture of how descriptions secured reference were right, the version of the Liar that secured self-reference with a definite description would be a Schema paradox. If not, not. . .*

More recently, Priest and a phalanx of co-authors—Zach Weber, David Ripley, Dominic Hyde, and Mark Colyvan—have offered a more sophisticated line of argument on the question of Curry's relationship to the IS and hence the PUS. They grant that Curry *does* fall under Inclosure, *if* "Curry's paradox" means the paradox about sentences with consequents such as the consequent of (8) above—that *everything* is true. *Everything* includes (8)'s negation. As JC Beall (2014a) argues, this gives us both Closure and Transcendence in pretty much the same way we get those elements in the formal structure of the Liar Paradox. Weber, Ripley, Priest, Hyde, and Colyvan—henceforth, following Beall's usage in the exchange, we will call them WRPHC—do not dispute Beall's technical point.

> So Curry's paradox fits the schema? No; this is not Curry's paradox! The problem posed by Curry-style reasoning is that it allows us to establish an *arbitrary* sentence, not just $\bot$. If A is any sentence and s is the sentence $\mathrm{T}s \rightarrow A$, the Curry reasoning allows us to establish A. This is problematic whether A is true or false (Weber et al., 2014, p. 823).

To see their point, let us consider a few other Curry sentences:

(10) If sentence (10) is true, the moon is made of green cheese.
(11) If sentence (11) is true, Hillary Clinton is the 45th President of the United States.
(12) If sentence (12) is true, Donald Trump is the 45th President of the United States.
(13) If sentence (13) is true, Canberra is the capital of Australia.

Assuming the validity of Contraction, the T-Schema truth conditions of (10) entail that the moon is made of green cheese. The truth conditions of (11) entail that Hillary Clinton is the President of the United States. Since astronomers tell us that the moon is not made of green cheese, and anyone who reads a newspaper knows the American Presidential election went the other way, it is obvious why these are paradoxical results... But what about (12) and (13)? "Donald Trump is the 45th President of the United States" and "Canberra is the capital of Australia" are not paradoxical results. They are not even false. Nevertheless, there does seem to be something paradoxical about the idea that they can be derived from a mere statement of the *truth conditions* of (12) and (13). As WRPHC put it, focusing on (13), there is "something crazy about thinking we can establish that Canberra is the capital of Australia in *this way* [...]" (Weber et al., 2014, p. 823, our emphasis).

Beall's formal analysis applies to (10) and (11). Since Hillary Clinton is not the President and the moon is not made of green cheese, in each case we can use Contraposition and *Modus Ponens* to derive the negation of the antecedent and hence the negation of the whole conditional. But since the consequents of (12) and (13) are *true*, there is no route through Contraposition to fitting these sentences into the formal structure of the IS.

> Beall is attempting to appeal to the principle of uniform solution: same kind of paradox, same kind of solution (Priest, 2002, Sects 11.5, 17.6). What is at issue, then, is the question of whether the Curry paradox is an inclosure paradox. All the instances of the Curry paradox—whatever the A is—are clearly of the same kind. And it is not an inclosure paradox, since instances of it do not fit the schema. By the uniform solution principle, the Curry paradoxes all require the same treatment, but not (necessarily) the same as that of the Liar [...] (Weber et al., 2014, p. 823).

Beall's reply to WRPHC on this point takes the form of a dilemma. What, he asks, is the IS supposed to apply to—entire sentential schemes, like $(s = (\mathrm{T}s \rightarrow A))$, or individual instances of that scheme?

> Here, we [...] branch into two cases.
>
> Super-paradoxical case: a sentential scheme is an inclosure paradox iff *all* instances are (inclosure) paradoxical.
>
> Sub-paradoxical case: a sentential scheme is an inclosure paradox iff *some* instances are (inclosure paradoxical) (Beall, 2014b, p. 843).

Consider (14), below, which is recognizably a version of the Liar Paradox—in fact, it is just the sort of sentence that might be brought up as a revenge Liar to frustrate certain non-dialetheist solutions to the paradox:

(14) Either this sentence is false or Hillary Clinton is the 45th President of the United States.

Given that the second disjunct is contingently false, (7) is true iff the first disjunct (and hence the whole sentence) is false. In other words, it is true iff it is false like any other Liar sentence, and hence it fits the IS. So far, so good. Now consider (15):

(15) Either this sentence is false or Donald Trump is the 45th President of the United States.

Given the contingent truth of the second disjunct, (15) cannot be made true by being false. Hence, the IS does not apply. But (14) and (15) are both instances of the sentential scheme $(s = (Fs \vee A))$! Thus, Beall concludes, if a sentential scheme must be super-paradoxical to be a schema paradox, even the Liar Paradox is not a schema paradox. If, on the other hand, being sub-paradoxical is enough, Curry is a schema paradox too. Either both paradoxes or neither of them fit IS.

So, who is right about all of this—Beall or WRPHC? The answer is that it does not matter. If Beall is right, so much the worse for Priest's claim to be offering a unified solution to all the IS paradoxes. If WRPHC are right and Curry is *not* an IS paradox, so much the worse for the Inclosure Schema as a plausible dividing line between types of paradoxes.

Of course, all of this assumes that there is no way to extrapolate a rejection of Contraction *from* dialetheism. This might be questioned. After all, Contraction is simply a short cut for the inferential work that can be done more slowly with the combination of Conditionalization and *Modus Ponens*. (If we have $\alpha \rightarrow (\alpha \rightarrow \beta)$ and we assume α for Conditionalization purposes, we can get to β with two applications of *Modus Ponens*.) There is no obvious conceptual connection between accepting dialetheism and rejecting Conditionalization, but as paraconsistent logicians have known since the early days of their project, there *is* a straightforward connection between accepting dialetheism and rejecting *Modus Ponens*. In classical logic, an indicative conditional $\alpha \rightarrow \beta$ has the same truth table as the disjunction $\neg\alpha \vee \beta$. Thus, given Double Negation, instances of *Modus Ponens* ($\alpha \rightarrow \beta$, α // β) are equivalent to instances of Disjunctive Syllogism ($\neg\alpha \vee \beta$, $\neg\neg\alpha$ // β).

Instead of embracing this consequence as the basis of a uniform solution to the Liar and Curry's Paradox, Priest (and most other dialetheists) have rejected the classical equivalence between $\alpha \rightarrow \beta$ and $\neg\alpha \vee \beta$ and gotten around Contraction by rejecting Conditionalization. Graham Priest has been particularly adamant about the unacceptability of giving up on Modus Ponens.

> Any conditional worth its salt, $\rightarrow$, should satisfy the modus ponens principle [. . .] This is, indeed, analytically part of what implication is. (Priest, 2006b, p. 83)

JC Beall, however, has recently argued that regarding *Modus Ponens* as invalid is not as counterintuitive a move as it may initially seem. We still have a way of reasoning away from, for example, the premises "Lucy is a miniature schnauzer" and "If Lucy is a miniature schnauzer, then Lucy is a dog" to the conclusion "Lucy is a dog." We just need an extra step.

> Many have thought that when we 'do modus ponens' or 'follow modus ponens', we are simply inferring (say, accepting something new) *in accord with logic*, that is, *in accord with what our theory logically implies*. Since $\{A, A \supset B\}$ doesn't imply B, we can't be 'in accord

> with logic' when we infer B from $\{A, A \supset B\}$, when we accept B into our expanded theory. We are close (Beall, 2013, p. 10).

The "close" inference that Beall still takes to be valid is the inference from the set of premises $\{A, A \supset B\}$ to a choice between the two elements of the set of possible conclusions $\{B, A \wedge \neg A\}$. Since Beall is committed to the principle that it is rational to reject contradictions whenever we can plausibly do so, in any case where A and B are ordinary, non-paradoxical sentences, it is rational to reject $A \wedge \neg A$ and hence to accept B.

The formal theory he uses to make sense of all this is LP+ a multi-conclusion version of Priest's Logic of Paradox. Crucially, though, he accepts—granting a point to which he credits an anonymous referee—that, "given background (so-called structural) features of LP+ together with the clauses for disjunctions," his strategy for explaining how we can rationally "do" what looks like *Modus Ponens* without *Modus Ponens* being valid could just as easily be run with the *single*-conclusion inference from $\{A, B\}$ to $(B \vee (A \wedge \neg A))$. He favors the multi-conclusion presentation for two reasons. The first is that the single-conclusion version "need not always hold for non-classical logics" and may not hold even for more complicated variations of his favored framework. The second is that "the 'choices' idea [. . .] seems to be more intuitive when presented in the more general multi-conclusion guise [. . .]" (Beall, 2013, p. 11).

Crucially, though, Beall seems to accept that the distinction between rationally rejecting one element of a set of possible conclusions and thus rationally accepting the other and rationally rejecting one disjunct of a disjunction and thus rationally accepting the other does not make a difference. In principle, however the choice is formally represented, if we have a choice between accepting two claims and a rational reason to reject one, we have a rational reason to accept the other.

That is just Principle R.

5.2 Why Priest Rejected Principle R (and JC Beall Should Too)

In fact, Principle R should be every bit as unacceptable to a dialetheist as Disjunctive Syllogism. To see why, start by noting an important distinction.

It may be psychologically impossible for the same individual to simultaneously accept and reject α, but it hardly follows that the *circumstances under which it is rational to accept* α never overlap with the *circumstances under which it is rational to reject* α. Given the LNC (in the sense of "the LNC" that dialetheists reject) and (quite controversially in this context) the claim that rejection should be understood in terms of negation, we might conclude that it is *logically impossible* for such an overlap to exist. Although failing that, what remaining reason do we have for regarding such overlap as metaphysically or epistemically impossible?

One might be tempted to summon the ghost of Kant here and assert a non-moral version of the "ought implies can" principle. How plausible this move seems is likely to depend on (a) how plausible one finds "ought implies can" in its home context of *moral* obligations, and (b) how relevantly similar one takes them to be to the norms of epistemic rationality. On (a), even in the first edition of *In Contradiction,* in which he endorses Principle R, Priest takes it for granted that it is possible for it to be the case that we morally *ought* to do two things which we jointly *can't*. This is an extended aside in his discussion of contradictory legal norms.

> Consider moral norms. It is a commonplace that moral norms frequently conflict *prima facie*. But it is not so easy to establish that this conflict is not just *prima facie*, and that inconsistent moral obligations may indeed arise [. . .] Still many writers have recently suggested that not all *prima facie* moral conflicts can be resolved, and I think that the truth lies with them.
>
> [. . .] Of course, granted that there are inconsistent moral obligations, when these rise a *practical* problem is posed with which dialetheism gives no help. Assuming that the obligations involved are not such as to be jointly realizable, one needs to decide which of them to fulfill. All that theory can say is that whichever one fulfills some obligation will be flunked. This is cold comfort (Priest, 2006b, p. 197–198).

According to Priest, it no more follows from the assumption that inconstant moral obligations exist (or even from the further claim that *moral dialetheias* exist—for more on the distinction, see Sect. 3.1) that I can both perform and not perform some Action P than it would follow from rejecting mainstream Einsteinian physics in favor of some neo-Lorentzian physical theory whereby time really was a universal invariant and it should be possible in principle for an object to accelerate to faster than the speed of light (without traveling back in time) that I will personally be able to get my bicycle to go that fast if I just pedal hard enough. To sum up his view on (a)—with which I am inclined to agree—in colloquial terms, if our moral obligations outrun our metaphysical capacities, that is *just too bad*.

Perhaps the most extreme defense of (a) in the history of western philosophy comes from Peter Geach, who contends that the creator of the universe "does not require of a faithful servant the desperate choice between sin and sin."[4] Geach's idea is that, while it may be possible for us to engage in armchair speculations about moral dilemmas, constructing internally consistent thought experiments in which genuine moral rules come into conflict, God in His benevolence would never allow such a scenario to come to pass. To see how someone who shared Geach's theological premises could extend them to issues of rational choice, think about Buridan's Ass, equally hungry and thirsty and exactly as far from a stack of hay and a pail of water. Perhaps a loving God would not allow a real donkey to die of hunger and thirst for lack of a rational reason to prefer one course of action to the other, and so no such rational dilemmas ever arise in our divinely ordered universe.

The problem is that *even if* we go along with Geach on (a), there are at least two reasons to think the argument would falter at step (b). The first has to do with the

[4] See Statman 1995, p. 44

content of plausible norms about the rational acceptability or rejectability of propositions. In *Doubt Truth To Be A Liar*, Priest endorses a principle he calls **Accept**.

> **Accept**. One ought to accept something if there is good evidence for its truth.

This much seems uncontroversial (or about as uncontroversial any such principle is likely to be). As he goes on to note, a "natural dual" to **Accept** is:

> **Reject(F)**. One ought to reject something if there is good evidence for its falsity (Priest, 2006a, p. 109).

> There is nothing in the existence of truth value gaps that threatens this principle, but it is unlikely to recommend itself to a dialetheist. Consider something taken to be both true and false, such as, 'this sentence is false.' Call this λ. Half of the paradoxical argument provides good evidence for ¬λ, but one would not on this count want to reject λ. For the other half of the paradoxical argument provides evidence for λ. **Accept** therefore counsels its acceptance.
>
> The lack of symmetry between **Accept** and **Reject(F)** might be felt to be uncomfortable. Why the one, but not the other? But one should recall that falsity is nothing other than truth of negation. So the correct principle concerning falsity is that delivered by **Accept**:
>
> One ought to accept ¬α if there is good evidence for the falsity of α (Priest, 2006a, pp. 109–110).

That is all very well, but the issue is not when to *accept* ¬α, but when to *reject* α. One might start to worry that Priest has reasoned himself into a corner here, and that he will be left without a plausible principle about rejection. This would be a major embarrassment problem for his account—*especially* if he wanted rejection to do work independent of acceptance-of-negation.

Priest is, however, willing to accept a rejection principle formulated in terms of *untruth*:

> **Reject(U)**. One ought to reject something if there is good evidence for its untruth (Priest, 2006a, p. 110).

But wait! Should not "untruth"(¬Tr<α>) be equivalent to "truth of negation" (Tr<¬α>) in any framework where the intersubstitutivity rules discussed in 1.2 operate? It is easy enough to derive Tr<¬α> from ¬Tr<α>:

1. ¬Tr<α> Premise
2. ¬α 1 Release
3. Tr<¬α> 2 Capture

. . . .and vice versa:

1. Tr<¬α> Premise
2. ¬α Release
3. Tr<¬α> 2 Capture

The short answer is that Priest does not endorse unrestricted Capture and Release. Part of the story here has to do with the technical details of Priest's various attempts to add conditional operators to LP—note that Priest *does* accept the unrestricted T-Schema!—and part of it has to do with his desire to limit the "spread" of contradictions to which he is committed. The last piece of the puzzle has to do

with his theory of truth. Priest's fellow dialetheist JC Beall not only accepts the universal intersubstitutivity of α and Tr<α> but subscribes to a version of deflationism about truth according to which the substitutivity rules (Capture and Release) literally constitute the meaning of the truth operator. For him, the chief attraction of non-classical solutions to the paradoxes is that they allow non-classical paradox-solvers to hold onto this theory of truth. Priest, on the other hand, subscribes to a teleological theory of truth, whereby truth is defined by its role as the *telos* or goal of assertion. This, for him, makes truth more fundamental than falsehood. Whether a proposition is true is in some sense a more "basic" question for Priest than whether it is false, so F<α> should not suffice to settle the question of whether Tr<α> or ¬Tr<α>.

On this last point, it seems to me that Alan Weir's response is decisive.

> As against this, one could assert that falsity is the telos of denial and deny that assertion is more fundamental than denial. More importantly one could, and in my view should, deny that linguistic behavior has any overriding aim or goal at all: with Wittgenstein and Chomsky one should stress the highly variegated purposes to which linguistic acts, including assertion, are put. Even if the statistical norm closed linked to observation is that they be true, there is no reason to suppose that it is true in general; nor is it so clear what it could mean to say that the norm, in some non-statistical sense, of assertion is truth (Weir, 2004, p. 391).

It should be noted that no other dialetheist accepts Priest's claim that falsity is distinct from untruth.[5] It is not hard to see why. Most dialetheists, like most critics of dialetheism, think that the argument from the Liar Paradox is the best argument for dialetheism. Well, as we have already seen, the entailment from falsity to untruth can only be blocked by rejecting the universal validity of the intersubstitutivity rules. . . . but if they sometimes fail, why not just use their failure in the case of paradoxical sentences to block standard Liar reasoning? Why give up on Disjunctive Syllogism and Contraposition and the rest if denial of the intersubstitutivity rules is acceptable in principle? If we really are faced with the choice between the intuitive costs of rejecting certain instances of those rules and accepting counterexamples to the LNC, reasonable people might disagree about which cost is greater, but why bite *both* bullets?

Regardless of what one thinks about the issue of falsity entailing untruth, there is a straightforward reason to think that the distinction will not do the work Priest wants it to do in this context. His point about Simple Liar and **Reject(F)** should apply equally well to the Strengthened Liar and **Reject(U)**. Priest acknowledges this point and considers the possibility of rejecting even **Reject(U)**. (He talks about rejecting certain instances of it, seeing it as "generic," but in this context, that amounts to

[5] Gareth Young says that Priest's stance on this issue "is not the standard one for glut theorists. I know of no other proponents of the view who reject the entailment from falsity to untruth [. . .]" (Young, 2015, p. 269). Professor Young wrote his doctoral thesis on revenge problems for dialetheism, and he has an encyclopedic knowledge of this literature. If dialetheists other than Priest who denied this entailment existed, Young would know about them. Since I take Contraposition to be universally truth-preserving, I am comfortable concluding that there are no such dialetheists.

much the same thing. Saying that he "accepts" **Reject(U)** in this generic sense would be like saying that he accepts **Reject(F)** or for that matter that he accepts the principle that contradictions should always be rejected.) Although he seems to find this unsatisfying, as with his comments on the validity of RAA (which we looked at in 4.2), his language is non-committal:

> The failure of symmetry in this case is might be thought to be more troubling, however. Falsity is not the opposite of truth [. . .] Untruth, by definition, is. So does the acceptance of **Reject(U)** damage dialetheism. The answer is 'no.' It is quite open to the dialetheist to accept both **Accept** and **Reject(U)**. Suppose that we have some α for which there is good evidence that it is both true and untrue. Then, by these two principles, one ought to accept α and one ought to reject α. We have a dilemma, since we cannot do both (Priest, 2006a, p. 110).

This, then, is the first of Priest's two reasons for rejecting Principle R. Just as, even in the first edition of *In Contradiction*, he accepts that there could be moral dilemmas (and even moral dialetheias), he is now open to the possibility of rational dilemmas. Once this possibility is taken seriously, an unrestricted acceptance of Principle R threatens to reduce dialetheism to triviality.

> Arguably, the situation is exactly the same as that concerning the disjunctive syllogism (DS). Given $\alpha \vee \beta$, $\neg\alpha$ delivers β only if one does not also have α. Similarly, if one accepts $\alpha \vee \beta$, rejecting α delivers β only if one does not also accept α (Priest, 2006b, p. 276).

Recalling that Principle R is about rational *acceptability* and *rejectability*, not actual acceptance and rejection, the point can be rephrased: If α is rationally acceptable and rationally rejectable, and β is *just* rationally rejectable, Principle R will counsel accepting β. Given **Reject(U)**, this exact situation arises with disastrous consequences when we consider sentences like (16) or worse yet (17).

(16) Either (1) is true or people turn into frogs when tapped on the shoulder.
(17) Either (1) is true or everything is true.

In the first edition, he tries to head this off in two ways.

> Principle R is clear enough. Still, it might be objected *ad hominem* that it is not correct, since someone might rationally accept *and* reject α. There are two points to be made in reply to this. The first is that joint rational acceptance and rejection are not possible [. . .] (Priest, 2006b, p. 114)

Priest's phrasing here is ambiguous in an important way. Is the idea that (i) it is psychologically impossible for the same person to both accept and reject α at the same time (and hence impossible for them to do so in a rational manner), or that (ii) it is impossible to be rationally obliged to do so?[6] On the first reading, the reply is

[6]Priest ends the sentence by saying that the impossibility of joint rational acceptance and rejection was established in Sect. 7.4. Sadly, flipping back to 7.4 simply deepens the confusion. In the relevant passage, he does not clearly distinguish between the two issues. I quote it in full so readers will not worry that they are missing some nuance that would clarify his position. "To see that they are incompatible, note that what is rationally acceptable and rejectable (in a certain context of inquiry) is what the ideal rational agent (a fictitious but useful creature) accepts and rejects. But, as we saw in Sect. 7.3, acceptance and rejection are mutually incompatible. In other words, if one is

simply irrelevant, since Principle R concerns what we are obliged to do, not whether we are successful in fulfilling those obligations. As with the moral case, if we are obliged to do what we cannot, that is just too bad.

His second reply is that even if joint rational acceptance and rejection were possible, "Principle R would not be undercut."

> The argument that it would goes as follows. Suppose that a disjunction and one of its disjuncts is rationally acceptable. Then it does not follow that the other is rationally acceptable. Hence from the facts that a disjunction is rationally acceptable and that one of its disjuncts is rationally both acceptable and rejectable, it does not follow that the other is rationally acceptable. The argument is of the form: $\neg(\gamma \rightarrow \delta)$; hence $\neg(\gamma \wedge \eta \rightarrow \delta)$. Enough said (Priest, 2006b, p. 114).

To see what is wrong with his reasoning in this passage, note that it could be adapted point-for-point to accuse anyone who thought DS failed in a dialetheic framework of the same fallacy! *True contradictions are impossible, and even if they were possible, DS would not be undercut. The argument that it would goes as follows: Suppose that a disjunction and one of its disjuncts is true. Then it does not follow that the other one is true. Hence from the fact that a disjunction is true and that one of its disjuncts is true and false, it does not follow that the other is true. The argument is of the form* $\neg(\gamma \rightarrow \delta)$*; hence* $\neg(\gamma \wedge \eta \rightarrow \delta)$*. Enough said.* In both cases, this is a decisive refutation of one very silly argument that someone somewhere perhaps *could* make for the dialetheic unacceptability of the rule/principle in question, but the argument being refuted is hardly "the" or even the most obvious argument for that conclusion.

---

presented with a claim that it is *prima facie* both acceptable and rejectable, this conflict must be resolved in favor of one or the other party, or of agnosticism. (This might be done by the kind of weighing procedure I shall discuss in the next section.) Since it is impossible jointly to accept and reject the same thing, it is, *a fortiori*, impossible to do this rationally. Hence rational acceptance and rejection are incompatible" (Priest, 2006b, p. 103). It looks to me like Priest is talking about (i) until the last two sentences, in which he slides over to (ii). It is also unclear why he takes all *prima facie* conflicts in *this* domain to be resolvable, given the comfort he expresses in the previously quoted passage for saying that ideal *moral* agents would have to perform impossible tasks. In fact, there is an obvious symmetry between his description of rational "weighing" in 7.5 and his description of moral weighing in 13.5. He argues that *prima facie* inconsistent obligations arise because a variety of different considerations are relevant to moral rightness and wrongness, and that while they might be merely *prima facie* if some transcendental overriding consideration can be appealed to in order to clean things up, he expresses skepticism about the existence of the right sort of consideration. Most strikingly (in terms of the comparison with his line on moral norms), he declares that, "What the examples of *prima facie* conflicts do is place the burden of argument on those who would claim that apparent conflicts can *always* be resolved to show just cause" (Priest, 2006b, p. 198). Oddly enough in light of all of this, when he discusses rational "weighing" in 7.5, he does not try to discharge this burden by offering any sort of argument for the impossibility of rational dilemmas. Nor does he appeal to a single overriding factor to clean up the appearance of dilemmas. Indeed, he revels in the messiness of calculating rational norms. "If this is all disconcertingly non-algorithmic, that is just an unfortunate fact of life" (Priest, 2006b, p. 106). Perhaps there is some interesting disanalogy between the two realms he could have brought up at this point. If so, he never produces it. In any case, his arguments in the second edition (and in *Doubt Truth to Be a Liar*) settle the issue. The possibility of genuine rational dilemmas is just part of the price tag of dialetheism.

Priest goes on to give a positive argument for the claim that Principle R would continue to operate even in the presence of overlap between rational acceptance and rational rejection:

> Some confusion may arise from the thought that something's being rationally acceptable 'cancels out' its rational rejectability. This is just a confusion. If something is rationally acceptable and rejectable, it is still rationally rejectable. Any consequences that this fact has, therefore, still stand (Priest, 2006b, p. 114).

The problem, of course, is that if the consequences that allegedly "still stand" include triggering Principle R, the core rationale for paraconsistent logic—blocking the inference from inconsistency to triviality—has been sacrificed. The combination of Principle R and rational dilemmas takes us to triviality as fast as the combination of DS and dialetheias. To his credit, Priest flatly disowns this line of argument in the second edition.

> In reply to an objection to principle *R* in Sect. 8.3, I argued that, if one accepts a disjunction and rejects one of the disjuncts, one is still committed to accepting the other, even if one also accepts the disjunct. This, it now seems to me, is just wrong (Priest, 2006b, p. 276).

Of course, even given the failure of the second reply, Principle R could still be saved, if the first reply—on reading (ii)—was correct. Sentences (13) and (14) are only counter-examples to the principle given Priest's acceptance of **Reject(U)**. How bad would it be for a dialetheist to reject it?

To start to answer that question, note that **Reject(U)** contains relatively little information about when propositions should be rejected. We should reject propositions when there is "good evidence" that they are untrue. How much evidence, and of what kind, is sufficient to count as "good evidence"? Priest presumably leaves the fine-grained questions to the epistemologists. All **Reject(U)** tells us is that when the standard for rationally concluding $\neg Tr\langle\alpha\rangle$ has been met, *whatever that standard might be*, α should be rejected. As such, a dialetheist who rejected even **Reject(U)** would not be left with much of anything that they could (universally) assert about rational rejectability. Indeed, it is hard to see how they could do better than the annoyingly uninformative **Reject(R)**. . . .

Reject(R). One ought to reject something if there is good evidence that it should be rejected.

. . .which is saved from being a *completely* uninformative tautology only by the unhelpful gesture in the direction of the not-very-informative notion that rational oughts have something or other to do with evidence. A dialetheist might be tempted to attempt to salvage the idea of a definitional connection between a proposition's truth value and its rational rejectability by asserting **Reject(JF)**.

Reject(JF). One ought to reject something if there is good evidence that it is (just) false.

This will not work. We will explore the difficulties faced by dialetheists when they try to make sense of the notion of "(just) false" in due course, but it is enough for now to notice that **Reject(JF)** produces rational dilemmas in the faces of sentences like (18) below in *just the same way* that **Reject(F)** produces them in

the face of sentences like (2) and **Reject(U)** produces them in the face of sentences like (1).

(18) Sentence (18) is (just) false.

By a straightforward application of the same "Liar reasoning" the dialetheist uses to conclude that sentences like (1) and (2) are both true and false, (18) is both true and (just) false! If (18) is true, it is (just) false. If it is false, it is either (just) false (in which case it is true) or both true and false. If it is both true and false, it is true, and hence (just) false. Similarly, it will not help anything to say that it is (just) true. If it is (just) true, it is true, and hence also (just) false!

Dialetheists such as Priest and Beall have tied themselves into knots trying to say something plausible about "(just) false." Some of the attempts to get around these problems involve trying to explicate "(just) false" with rejection-talk. Even if such a strategy were successful, it would render **Reject(JF)** as uninformative as **Reject(R)**. Moreover, there are good reasons to think that the strategy fails. We will get to all of that later. What we have said so far should suffice to show that **Reject(JF)** does not represent a lifeline for the dialetheist trying to save Principle R. Given (9) and a dialetheic solution to the paradoxes, there are sentences that are both true and (just) false. Given **Accept**, such sentences should be accepted. Given **Reject(JF)**, they should be rejected. Given sentences that are both rationally acceptable and rationally rejectable, Principle R reduces paraconsistent dialetheism to outright trivialism.

So far, we have only produced such dilemmas by introducing principles such as **Reject(F)**, **Reject(U)**, and **Reject(JF)**. However, even if the dialetheist rejected all such principles (even the annoyingly uninformative **Reject(R)**!), and remained entirely silent on the whole subject of when propositions should be rejected, this would not be enough to save Principle R. Priest's second and more decisive reason for rejecting it has to do with the Irrationalist's Paradox. Consider sentence (19), below.

(19) Sentence (19) is either rationally rejectable or falls into the category of sentences about which we should be agnostic.[7]

This, it seems to me, is equivalent to the sentence provided by Priest in his discussion of the paradox, sentence p, where p is, "It is irrational to believe p" (Priest, 2006a, p. 111). I prefer to talk in terms of rational acceptability and rejectability to avoid any ambiguity, given that Priest uses such talk (rather than the slightly more ambiguous "irrationality" talk he uses to formulate sentence p) in Principle R. . . .and, as we will see in 3.7, he uses rational acceptability and

[7]Priest's raising of this paradox, in which "rationally rejectable" is treated as operating on sentences, might seem to sit uncomfortably with his insistence elsewhere that acceptance and rejection are not connectives and cannot operate on sentences as if they are, but this conflict can be resolved when we remember that the issue here is not *rejection* but *rational rejectability*. By analogy, consider, "Bob gets angry when he reads α" (or more to the point "it would be irrational for anyone to get angry when reading α.") Anger does not operate on sentences but "angering Bob" (or "being a sentence it would be irrational for anyone to get angry about") does.

rejectability talk when defending the right of dialetheist's to use *reductio* arguments, an idea he holds onto even after abandoning Principle R! Also, the wording of (19) has the advantage of making clear a dilemma or some sort results *even if* the rationally correct thing to do is to reserve judgment.

If (19) is rationally rejectable or neither rationally acceptable nor rationally rejectable, then either way it is true and hence by **Accept** it is rationally acceptable. It could be objected that **Accept** does not say that all true propositions should be accepted, but only that all propositions for the truth of which we have sufficient evidence should be accepted. Although in this context, distinction does not make a difference. If it is rationally rejectable, then—even assuming that rejection is not governed by anything more informative than **Reject(R)**—the state of the evidence must be such that we are warranted in believing it to be rejectable. Well, if we are warranted, in believing that we are warranted in believing the disjunction of that and the claim that it is neither acceptable nor rejectable, which just *is* (19). Feed that into the T-Schema, and we are warranted in believing that (19) is true. The same argument applies *mutatis mutandis* to the possibility that (19) is a sentence about which we should remain agnostic. If (19) is such a sentence, parallel steps will fulfill the conditions for **Accept**. Thus, given reasoning closely parallel to that used by the dialetheist to establish contradictions about sentences like (1) and (2), rational dilemmas must be accepted.

A dialetheist who wanted to hold onto Principle R in the teeth of this problem might try to hold the line here, pointing out that so far the argument from this version of the Irrationalist's Paradox has not established that there is at least one proposition α that is both *rationally acceptable* and *rationally rejectable*, but merely that there is at least one proposition α for which the following disjunction holds, where N(α) signifies that α should neither be accepted nor rejected (presumably because, given the current evidence, we should reserve judgment):

$(A(\alpha) \wedge N(\alpha)) \vee (A(\alpha) \wedge R(\alpha))$

The second disjunct would blow up Principle R, but the first disjunct would not. Perhaps, then, all that the paradox shows is that there is at least one proposition which we should both accept and be agnostic about, neither accepting it nor rejecting it. This will of course entail a contradiction $(A(\alpha) \wedge \neg A(\alpha))$, but the dialetheist has *already* accepted that some of those are true. The important point is that there are no propositions about which $A(\alpha) \wedge R(\alpha)$ are both true.

To test the possibility of holding the line here, let us consider what should superficially be an easier case.

(20) Sentence (20) is rationally rejectable.

On first glance, the best policy for dealing with (20) is to hold that this is a sentence about which we should rationally remain agnostic. (Sadly, no term like "rationally acceptable" or "rationally rejectable" suggests itself. "Rationally agnosticable" does not exactly roll off the tongue.) After all, if (20) is rationally

rejectable, then by the same reasoning we used in the case of (20) (but with one less step) we get the result that it is both acceptable and rejectable, and if it is rationally acceptable, we get the same result. (Presumably, if we should accept the claim that some claim should be accepted, we should accept it!) By putting it into that third category of sentences that should neither be accepted nor rejected, we can avoid both logical and rational paradoxes.

Although what would it mean to put it in this category? On Priest's assumptions, residence in the third category is a matter of an obligation to *remain agnostic*.

> One can reject something without accepting its negation. The statistician who rejects a certain hypothesis does not, thereby, reject its negation. In fact, its negation will not normally be a statistical hypothesis at all (i.e. of the form that a chance distribution is such-and-such).[8] [. . .] Conversely, one may accept ¬α while failing to reject α. One would do this if, while being convinced that ¬α is true, one acknowledged the possibility that it might be a dialetheia. Until this was ruled out, one would have to remain agnostic about α. [. . .T]he trichotomy accept/reject/be agnostic should now be clear (Priest, 2006b, p. 99).

We should be agnostic about something if we lack relevant information about it. Given orthodox assumptions about logic, we are agnostic about α if we do not know *which* of the pair α and ¬α is correct. Given either paracomplete or dialetheic assumptions, we could also be in the position of knowing the status of one member of the pair but being ignorant of the status of the other. (A paracompletist could take himself to know that α should be rejected, but not to know the status of ¬α. A dialetheist could consider herself to be in the opposite position.) Although in the case of (20), we do not seem to be in any of those positions. We have all the information about (20) we could want. There is no piece of evidence that we could learn in the future—or that was irretrievably lost in the past—that would clarify its status.

Recall from 1.5 that, wishing to differentiate his theory from "proto-paracomplete" ideas, Hartry Field prefers to describe paracompleteness as the view that there are "acceptance gaps" rather than "truth value gaps." Of course, a firm believer in the LEM might have "acceptance gaps" about all sorts of things. I have an "acceptance gap" about whether Alexander the Great's maternal grandmother's paternal grandfather was prone to bed-wetting as a small child. No one recorded the status of the boy's sheets for posterity, I do not have a time machine, and some metaphysicians have even argued that backwards travel through time is impossible in principle. This claim is controversial, but let us assume for the sake of argument that they are right, and as such that my inability to recover information lost to history was not just a contingent matter of no one having made the right technological breakthrough, or even a matter of nomological necessity, but a flat-out logical

[8] This example, it seems to me, relies on a straightforward equivocation between accepting or rejecting the proposition itself and accepting or rejecting it *as a statistical hypothesis*. If the statistician, in rejecting α as a statistical hypothesis is actually rejecting the content of α, then it hardly follows from the fact that ¬α will not be a statistical hypothesis in itself that the statistician has not accepted the proposition ¬α.

impossibility. Even so, *that* "acceptance gap" still would not be the kind that Field is talking about. He defined paracomplete theories as ones that "postulate acceptance-gaps even in states of complete information about the non-truth-theoretic world."[9]

This, of course, is precisely what we would be doing if we tried to insist on "agnosticism" about (20) or even (19).[10] In other words, a dialetheist who held the line here to avoid having to concede that rational acceptability and rejectability can overlap would be trading in dialetheism for paracompleteness.

Granted, hybrid positions are possible. (We saw in 1.5 that Beall briefly took such a position in his 2005 article "Transparent Disquotationalism.") The sort of hybrid possibility in question, though, would be considerably odder than that. Remember that even in 2005, Beall solved the Liar Paradox with gluts. His gaps were motivated by "vagueness, non-denoting terms, and other such familiar phenomena" (Beall, 2005, p. 15). The Liar and the Irrationalist certainly *look* like two heads of the same paradoxical beast, and it would be a considerably offense against the Principle of Uniform Solution (PUS)—same kind of paradox, same kind of solution—to solve one of them by jettisoning the LNC and to then turn around and sacrifice the LEM in solving the other. Moreover, logics that fall into the overlap between paraconsistency and paracompleteness are extraordinarily inferentially weak. Why amputate two limbs when one would do?

Finally, even a dialetheist willing to subject themselves to amputating different limbs to solve the Liar and Irrationalist's Paradoxes would not thereby be able to save Principle R. "Going paracomplete" does not just mean postulating acceptance *gaps*. It also involves rejection *gluts*. If neither (20) nor its negation should be accepted even in a situation of full information, then they must both be rejected. The same reasoning applies to (19).

5.3 Is the *Classical* Logician Entitled to Principle R?

In *Doubt Truth To Be A Liar*, having granted that the Irrationalist's Paradox makes Principle R unworkable within a dialetheic picture, Priest moves quickly to avoid the appearance that this puts him at any sort of dialectical disadvantage relative to his orthodox opponents. In fact, he argues, not only is the Irrationalist "just as much of a paradox for the non-dialetheist" as it is for the dialetheist, it is actually "*worse* for the classical logician" (my emphasis) (Priest, 2006a, p. 111).

[9] This particularly felicitous quote comes from private correspondence, but the idea recurs throughout his writings on paracompleteness.

[10] Someone could object here and say that we are not in such a state of full information about the non-truth-theoretic world, in so far as facts about rational acceptability and rejectability do not fall under the truth-theoretic heading, but making this objection work would require the objector to either (a) be able to provide some species of fact that we have not considered upon which rational acceptability or rejectability facts could supervene, or (b) make the rather desperate claim that reality contains brute rationality-facts, unexplainable in terms of truth, evidence, or anything else.

In at least one sense, this claim is flatly wrong. Remember, dialetheist is forced into rational dilemmas not so much by the Irrationalist's Paradox as by the combination of that paradox and the previously mentioned PUS (Principle of Uniform Solution). The core difference between the dialetheic solution to the Liar Paradox and *any* classically acceptable solution. The dialetheist embraces "Liar reasoning," holding both that all the premises are *true,* and that the paradoxical conclusion really follows from them. As discussed in 4.2, the first of those conjuncts is central to dialetheism's appeal. If the Liar, like any other paradox, is an intuitively good argument from seemingly unrejectable premises to a seemingly unacceptable conclusion, the major advantage of any classically acceptable solution (or, more generally, almost any non-dialetheic solution[11]) will be that the seemingly unacceptable solution can be avoided. The major disadvantage will be that one of the seemingly unrejectable premises (or one of the seemingly good inferences) must be rejected. The dialetheist's solution has the mirror image of these advantages and disadvantages, since none of the seemingly unrejectable premises need to be rejected. Given the extremely close parallels between both the premises of the Liar and Irrationalist's Paradoxes and the inferential steps involved, the dialetheist has very little room to maneuver in solving the Irrationalist.

For example, an important background premise of the Irrationalist's Paradox is that there *is* a proposition α such that $\alpha = \neg A(\alpha)$. This mirrors an equally important background premise of the Liar Paradox, which is that there is a proposition β such that $\beta = \neg Tr\langle\beta\rangle$. Harman and Horwich solve the Liar Paradox by rejecting the relevant instances of the T-Schema and thus denying that such a β exists. Similarly, a dialetheist *could* solve the Irrationalist's Paradox by finding a way to deny the existence of the equivalent α. . .but in doing so she would be considerably undercutting the case for dialetheism. If it is reasonable to avoid a paradoxical conclusion by saying that (19) and (20) do not "say of itself" what they seem to say, why not do the same for (1) and (2)?

Similarly, a dialetheist could argue that every proposition implicitly "says of itself" that it is rationally acceptable. Such a claim would reveal (20) to essentially be (21) in disguise.

(21) Sentence (21) is both rationally acceptable and rationally rejectable.

This would neatly avoid the problem. After all, if (14) is rationally rejectable, nothing rationally or logically paradoxical follows from this claim, any more than the truth of the second conjunct of (22) entails anything paradoxical.

[11] Hofweber's validity-is-generic solution, discussed in 1.6, is a borderline case for the way I diagnose different solutions here. Like dialetheists, Hofweber asserts that the premises of the paradoxical argument are *true* and that the inferential steps are good in the sense of being "valid," but unlike them he does not think the inference to contradiction is "good" in the sense of actually delivering its conclusion. More technically complex versions of a similar idea can be found in the work of Colin Caret, who has proposed an "inconsistent but not dialetheist" solution whereby the inferences are all valid, but the entailment relation works in unusual ways.

(22) Sentence (22) is both more than a thousand words long and less than a hundred words long.

Whatever *else* one thinks of this move, however, an obvious objection is that it could just as easily be made about truth. Indeed, many philosophers *have* made the same claim about truth, and thus argued that (1) is really just (17) in disguise, and hence a mere contradiction posing as a paradox.

(23) Sentence (23) is true and sentence (23) is not true.

I do not find the (1)-is-(23) solution to the Liar Paradox (and hence by parity of reasoning the (20)-is-(21) solution to the Irrationalist's Paradox) particularly plausible, but the point here is a narrow one. Whether these solutions are both reasonable or both unreasonable, they stand or fall together, and the point generalizes. For any Premise P of the Liar Paradox, there will not be any plausible way to argue that P is true but the parallel premise P* of the Irrationalist's Paradox is false, and the same point applies to the inferential steps that take us from the premises of each paradox to the contradictory or rationally dilemmic conclusion. By contrast, the classical logician *must* reject at least one P (or at least one inferential step) to (non-trivially) retain their commitment to classical logic.[12] Thus, the same principle of uniformity that commits the dialetheist to accept rational dilemmas based on the Irrationalist's Paradox actually forces the classical logician to reject a premise (or inferential step) of that paradox and thus to avoid dilemmas.

However, Priest provides three reasons for thinking that the dialetheist is at an advantage here after all. Let us take his three reasons one at a time.

> The Irrationalist's Paradox deploys the principle that if one believes that (α and to believe that α is irrational), then one's belief is irrational ($IB(\alpha \wedge IB\alpha)$). A dialetheist might contest this: it is not irrational if one *also* believes that to believe α is rational. The supposition that believing in α may be both rational and irrational is unlikely to appeal to a classical logician, however (Priest, 2006a, pp. 110–111).

Priest is clearly right that this move is unlikely to appeal to a classical logician. One of the best things about retaining both the LEM and the LNC is that the headaches involved in prying apart $A(\neg\alpha)$ and $R(\alpha)$ need not arise. Given classical logic and orthodox assumptions about the relationship between logic and reasoning, the circumstances under which α should be rejected are simply those under which $\neg\alpha$ should be accepted, which are entirely disjoint from the circumstances under which α should be accepted. This is a simple, elegant and immensely intuitively satisfying picture, which would be shattered by the introduction of rational dilemmas. Worse yet, "rational and irrational" sounds like not just a dilemma ($A(\alpha) \wedge R(\alpha)$) but a dialetheia ($A(\alpha) \wedge \neg A(\alpha)$). A classical logician who accepted the latter would either have to cease being a classical logician or become a trivialist.

[12] The combined product of dialetheism and classical logic, or indeed of dialetheism and any non-paraconsistent logic, would be trivialism.

More importantly, it is not clear how the dialetheist is supposed to be able to wring any sort of dialectical advantage from this move. Since his wording is ambiguous, let us go through both readings. If the phrase "irrational" in "it is not irrational to both believe α and believe that believing α is rational if one also thinks it's rational" means $\neg A(\alpha)$, then this true as far as it goes. . . .

1. $A(\alpha) \wedge \neg A(\alpha)$ Premise
2. $A(\alpha)$ 1 Conjunction Elimination
3. $\neg\neg A(\alpha)$ 2 Double Negation

. . .but what of it? This move certainly will not help with Principle R. Attempting to deploy it in that context would at best recapitulate Priest's "enough said" argument from the first edition of *In Contradiction* (which we examined in 4.5, and which Priest himself has quite rightly disowned). More generally, it does nothing to take away the sting of saying that we have rational obligations to believe irrational things. Earlier, we said that the conjunction "People turn into frogs when tapped on the shoulder and it is not the case that people turn into frogs when tapped on the shoulder" was (even for a dialetheist) as absurd as the internally consistent first conjunct would be on its own. Was that wrong? If not, why should it also go for "it's irrational to believe α and to believe that it's irrational to believe α, but it's also not irrational to do so"?

Perhaps the word "irrational" in this context should be read as $R(\alpha)$, but to get from there to $\neg A(\alpha)$, we need an "exclusion principle."

R to A Exclusion Principle: $R(\alpha) \rightarrow \neg A(\alpha)$

The obvious companion principle to this is:

A to R Exclusion Principle: $A(\alpha) \rightarrow \neg R(\alpha)$

It is hard to see how the two Exclusion Principles could come apart. Certainly, the assumption of dialetheism does nothing in itself to pry them apart. The further claim that rational dilemmas are possible might undermine the case for *both*, but it does not seem to undermine one more than the other. It the R to A Exclusion Principle holds even in the face of dilemmas, it is hard to see how the A to R Exclusion Principle could fail to be along for the ride.

With both in place, we can derive $\neg R(\alpha) \wedge \neg A(\alpha)$. In other words, it is "not irrational," but it is also *not rational*. This is essentially the same situation as we were in on the first reading, and the same considerations apply.

Priest's other two arguments are better. The second is that "not just dilemma but contradiction waits in the wings" (Priest, 2006a, p. 111). We have already seen one way whereby one might get from dilemma to contradiction—by means of the Exclusion Principles. There are other routes, but we can stop there, since rejecting the Exclusion Principles would mean rejecting the simple and satisfying orthodox picture of the relations between rejection, negation, and exclusion sketched out above, which would in turn considerable undercut the case for retaining classical logic.

Finally, Priest brings up the same connection between the Irrationalist's Paradox and more familiar paradoxes like the Liar we discussed above. He draws the opposite conclusion.

> [. . .S]ince this paradox makes no use of semantic or set-theoretic notions, none of the standard solutions to the paradoxes of self-reference get a grip on it (Priest, 2006a, p. 112).

We have already seen that at least two classical solutions to the Liar can be easily extended to the Irrationalist. Priest also mentions Russell's Paradox, which we covered in 2.3 and 3.2. While the application of the PUS to the Liar and Russell's Paradoxes is a lot less clear-cut than its application to the Liar and Irrationalist's Paradoxes—after all, both of the latter center around self-referential *sentences*, while (the canonical version of) Russell's Paradox is about *sets*—I am prepared to concede both points. All three paradoxes are similar enough that the right solutions to all three should be equally similar.

Even a partisan of classical orthodoxy less inclined than I am to concede the point about PUS should agree with the more modest claim that, in all three cases, *accepting the conclusion* of the paradoxical argument would be a disaster. This is clearest in the cases of the Liar and Russell's Paradox, because the conclusions are logically inconsistent, but accepting the rationally dilemmic conclusion of the Irrationalist's Paradox would not be much better, even without the threatened slide from dilemma to paradox.

To really get a sense of the unsavory consequences that would follow if the Irrationalist's Paradox forced the classical logician to accept rational dilemmas, let us return one last time to Principle R. We have seen that it is *dialetheically* acceptable. Is it classically acceptable?

It had better be. A classical logician who rejected Principle R would either be rejecting **Reject(U)** or forfeiting the right to use Disjunctive Syllogism in her reasoning. Either consequence would do tremendous damage to the case for the LNC.

To review, I have conceded to Priest that:

(i) The mixture of rational dilemmas and classical logic is unacceptable, so
(ii) Anyone who wants to defend classical logic (as I do) had better have a good solution to the Irrationalist's Paradox, and furthermore
(iii) Given the intuitive similarity between the three paradoxes, a good classical solution to the Liar Paradox, Russell's Paradox, or the Irrationalist's Paradox should apply in all three cases.

As such, when I argue for my preferred solution in Chap. 8, the reader is invited to use the question of uniformity as one of the criteria by which they judge it. *Does this apply to the Irrationalist's Paradox?*

I believe that it does, and that as such classical logicians are not forced into rational dilemmas (and are thus entitled to merrily continue to reason using Principle R and its equivalents for other classical inference rules without being forced into

accepting triviality). Even if I am right about all that, however, the dialetheist (or at least any dialetheist who wishes to use dialetheism to solve the Liar Paradox) is blocked by (iii) from simply taking on board my solution to the Irrationalist's Paradox. Principle R thus stubbornly remains dialetheically off-limits.

5.4 Re-Examining the Relationship Between Dialetheism and (Even "Informal") *Reductio* Arguments in Light of the Failure of Principle R

In Sect. 4.2, we saw how Priest argues that the question of whether RAA is *valid* can be separated from its legitimacy as a "critical weapon" in evaluating philosophical theories. In *Doubt Truth to be a Liar*, Priest separates RAA in the strict technical sense, the purpose of which is to "establish something of the form $\neg\alpha$ by deducing a contradiction from α" from a broader category of "*reductio* arguments" where "the point is not normally to establish something, but to force an opponent to give up a view" (Priest, 2006a, p. 86). We can think of these as RAA(N) and RAA(B), for "*Reductio Ad Absurdum* in the narrow technical sense" and "*Reductio Ad Absurdum* in the broad non-technical sense." While Priest's views on RAA(N) have evolved over time, he has been remarkably consistent in claiming that RAA(B) survives on the dialetheic picture. We saw a parade of quotations in 4.2. For now, notice that the way he describes RAA(B) in *In Contradiction* makes it abundantly clear that RAA (B) is to RAA(N) as Principle R is to Disjunctive Syllogism.

> A view can be criticized and made untenable if it can be shown to imply something that is rationally rejectable; for anything that implies something rationally rejectable is itself rationally rejectable (Priest, 2006b, p. 104).

Let us precisify these terms, to avoid any possible confusion later. RAA(N) is just the regular classical inference rule RAA. A definition of RAA(B) follows.

> RAA(B): If an entailment conditional is rationally acceptable and the consequent is rationally rejectable, the antecedent is rationally rejectable.

We saw in that earlier discussion that Contraposition in any form—whether RAA (N) or *Modus Tollens*—is untenable in a dialetheic framework, but in 4.2 we operated on the assumption that the dialetheist can at least make use of what we are now calling RAA(B). Although now that we have seen that Principle R is dialetheically untenable, that assumption needs to be re-examined.

One thing RAA(B) might seem to have going for it that Principle R did not is that the same difficulties do not seem to arise when it comes time to using it to reconstruct ordinary reasoning. As we saw in 4.5, trying to "substitute" Principle R for Disjunctive Syllogism means either implausibly attributing an implausible degree of sophistication to dogs and homophobic drill sergeants, claiming that people are not *really* reasoning from the premises that they think they are, or simply rejecting vast swathes of what certainly *looks* like good reasoning. No similar complications seem to arise

when we try to make sense of ordinary reasoning in terms of RAA(B). Indeed, Priest claims that we *would* run into considerable difficulties if we tried to make sense of ordinary *reductio* arguments in terms of RAA(N)! After all, no one thinks that "people turn into frogs when tapped on the shoulder" or "Graham Priest is a fried egg" are logically inconsistent claims, but we would all reject a theory that generated such absurdities. Critics of utilitarianism do not typically think that it is *logically impossible* that the right thing to do in the second version of the Trolley Problem is to push the fat man onto the tracks, but they think that the fact that utilitarianism generates this counterintuitive consequence gives us an excellent reason to reject it. Priest grants that dialetheists and non-dialetheists will differ about *which* claims are rationally rejectable, but since (a) non-dialetheists already take a number of consistent claims to be rejectable and merrily deploy RAA(B) against them, and (b) even dialetheists take some contradictions (indeed, *most* contradictions) to be rationally rejectable, the difference between RAA(B)-as-practiced-by-dialetheists and RAA(B)-as-practiced-by-non-dialetheists will be quite a lot narrower than it looks.

The first problem with all of this is that the claim about ordinary reasoning is flatly wrong, and relies on an equivocation about *consistency*. The claim that Graham Priest is a fried egg is not *internally* inconsistent, but it is inconsistent with the fact that he is not a fried egg but a human being. The claim that he will turn into a frog when tapped on the shoulder does not independently generate any sort of contradiction, but it does when combined with the background premise that people do *not* turn into frogs when tapped on the shoulder. The mere fact that utilitarianism generates the result that it is right to push fat men in front of trolley tracks in certain hypothetical circumstances would not give anyone a reason to reject utilitarianism (an act ordinary reasoners do not typically separate from accepting the negation of utilitarianism, although non-classical logicians think they should) if the people persuaded by this argument did not believe that it would *not* be right to do so.

To underline the point and see how this works, let us formalize this. Let α be the claim that utilitarianism is true and β be the claim that the hypothetical that fat-man-pushing would be morally right. The most straightforward reconstruction looks like this:

1. $\alpha \rightarrow \beta$ Premise
2. $\neg\beta$ Premise
3. α Assumption for RAA
4. β 1, 3 *Modus Ponens*
5. $\beta \wedge \neg\beta$ 2, 4 Conjunction Addition
6. $\neg\alpha$ 3-5 *Reductio Ad Absurdum*

It is true that, if you asked a non-utilitarian ethicist for a quick informal summary of her argument, she would most likely skip step 2. She might also skip step 5. People usually do skip Conjunction Addition in their casual explanations of their reasoning. One possible explanation of this communicative fact is that, psychologically, the difference between asserting 2 and 4 and asserting 5 does not usually amount to anything. Simon Evnine argues in his interesting book *Epistemic Dimensions of*

Personhood that, at least "in normal circumstances, what it is to believe a conjunction is simply to believe the conjuncts" (Evnine, 2008). I will not take a stand on the issue here, because it would take us too far away from our central focus here, and because the validity of Conjunction Addition lies in the overlap between the logical beliefs of Priest and other dialetheists and those of their orthodox opponents.[13] Whether Evnine is right about the explanation or not, it seems clear that once ordinary reasoners get to step 4, they can see that 6 follows.

But why? If not for the work being done by 2, what connection would be there be between 1, 3, and 4 on the one hand and 6 on the other? It seems far simpler and more plausible to say that 2 is left out of typical informal summaries not because it plays no role in the chain of reasoning, but because—like "Graham Priest is *not* a fried egg" and "People do *not* turn into frogs when tapped on the shoulder"—it is normally taken to be too obvious to warrant wasting the listener's time by making the premise explicit.

It could be objected that I am begging the question here by using negation-talk when rejection-talk would work just as well. Let us try out that suggestion, with operators R and A, where R(α) is read as "α is rationally rejectable" A(α) is read as "α is rationally acceptable." Now, one might think, the whole business can be re-run as an instance of RAA(B) rather than RAA(N).

1. A(α ⊨ β) Premise
2. R(β) Premise
3. A(β) 1, 2 RAA(B)

Is this not just as good a reconstruction—better, even, since it involves fewer steps? The problem is that it does not so much avoid the role played by the second step of the first reconstruction as obscure it. If R(β) is true, *why* is it true? The obvious explanation is that R(β) is the combined result of the intuition that ¬β and an implicit principle such as **Reject(F)** or (equivalently according to everyone but Priest) **Reject(U)**. Alternately, I suppose it could be claimed that it results from **Reject(R)** and. . .whatever it is that we need evidence of to be justified in rejecting things, but it is far from clear that this would count as an "explanation." It is undeniably true that critics of utilitarianism typically *do* believe that it would not be right to push fat men onto trolley tracks to save others. At the very least, it seems odd and unparsimonious to neglect to use this fact in one's explanation of the grounds of R(α).

Of course, even if I am right about this and the most natural and reasonable explanation of ordinary *reductio* arguments is that they are instances of RAA(N)—indeed, the most natural and reasonable explanation of the total situation is that RAA (B) is simply RAA(N) in disguise—all of the following *could* still be true:

[13] There are non-adjunctive paraconsistent logics, in which the inference from α and ¬α to β is blocked not by denying that (α ∧ ¬α) entail β but by denying that α and ¬α entail β, but the development of such logics is typically motivated by concerns such as modeling belief revision rather than by dialetheism.

(i) Dialetheism is true and hence
(ii) RAA(N) is invalid (See 4.2), but
(iii) RAA(B) is a good rule of reasoning

For (i)-(iii) to all be true, the basis of the rejectability of rationally rejectable contradictions would have to be a bit of a mystery. The dialetheist cannot reject them on any grounds that make use of principles such as **R(U)** without making *all* contradictions rationally rejectable. Moreover, even in cases where the route from **R(U)** and evidence for $\neg\alpha$ did ground R(α)—as it could not in any case where α is a dialetheia—the dialetheist *would* be faced with a trilemma between implausible ways of making sense of ordinary *reductio* arguments almost completely equivalent to the trilemma faced by a dialetheist trying to make sense of apparent uses of Disjunctive Syllogism in ordinary reasoning by appealing to Principle R. Either the reconstruction is implausibly complex, requiring extra steps to connect $\neg\alpha$ to R(α), or it involves implausibly claiming that ordinary reasoners' belief in $\neg\alpha$ plays no role in their chain of reasoning, or it involves regarding vast swathes of apparently good reasoning as fallacious. Still, the total epistemic situation could be such that the case for dialetheism was so compelling (and the need to "recapture" reasoning-by-*reductio* was so pressing) that it would be reasonable to bite these bullets.

The deeper problem is that the arguments given in 3.4 for the inescapability (on a dialetheic picture) of rational dilemmas makes the difference between RAA(B) to RAA(N) as superfluous as the difference between Principle R and Disjunctive Syllogism. Thus, all of the arguments for the dialetheic untenability of RAA (N) and Contraposition more generally given in 4.2 apply equally to RAA(B) or its equivalents for non-entailment conditionals. Take sentence (24).

(24) The sentence written in blue marker on the whiteboard in Room 214 of Scott Hall at the New Brunswick campus of Rutgers is rationally rejectable.

Assume that someone does in fact write (24) in blue marker on the whiteboard in Room 214. The Liar-style reasoning we used to derive dialetheically inescapable rational dilemmas for all options about (19) and (20) will apply just as well to (24). Thus, (25) (which should read as an entailment conditional) will be true.

(25) If sentence (25) has been written in blue marker on the whiteboard in Room 214, then that sentence is rationally rejectable.

In the scenario we are considering (and the dialetheic framework we are accepting for the sake of argument), the Liar-style reasoning we used to establish the truth of (25) will, combined with **Accept**, make (25) rationally acceptable. We have already seen that the consequent (i.e., (24)) is both rationally acceptable and rejectable. To review, so far, we have established (relative to the relevant assumptions) that:

(i) (25) is a rationally acceptable entailment conditional.
(ii) The consequent of (25) is rationally acceptable.
(iii) The consequent of (25) is rationally rejectable.

By (i), (iii), and RAA(B), the *antecedent* of (25) will also be rationally rejectable. But in this scenario, that antecedent is true. It is in fact, only if we know it is true that these consequences can be derived. Thus, in the distinctly non-fanciful hypothetical scenario in which a philosopher writes a paradoxical sentence on a whiteboard where many paradoxical sentences have in fact been written, Scott Hall 214 being a frequent site of philosophy courses, the dialetheist is forced to say that anyone who saw the sentence on the board (and is thus authorized by **Accept** to accept it) is also rationally obligated to reject the (*ex hypothesi*, accurate) testimony of their senses. Just as the combination of dialetheism and RAA(N) entails the falsehood of ordinary truths about the observable world, the combination of dialetheism and RAA (B) entails the rejectability of such plain (and plainly knowable) truths.

The same problem arises for any attempt to "recapture" *Modus Tollens* by substituting rejection talk for negation talk. Eliminate the qualification "(entailment)" from the definition of RAA(B) and it becomes MT(B) and all the same points apply.

The more general lesson should be clear. If a classical rule involving negation is dialetheically invalid, formulating a substitute rule involving rejection will not help. The same problems and paradoxes that led the dialetheist to reject the original rule will re-arise in the altered context of rejection-talk and sink the substitute.

5.5 The Broader Importance of the Failure of Principle R (II: Rejection-Talk in Beall and Field)

There is hardly any explicit talk of the notion of rational rejectability in Beall's book on dialetheism, *Spandrels of Truth*. There are couple of mentions of ideas that "must be rejected" (Beall, 2009 pp. 34 and 81) or "need to be rejected" (Beall, 2009, p. 9), but the closest thing to an in-depth discussion of the principles of rational acceptance and rejection comes in his discussion of an objection that dialetheism must by its own lights be irrational.

Recall from 1.3 that instances of the LNC (in the narrow technical sense, such an LNC instance is just $\neg(\alpha \vee \neg\alpha)$, not any sort of principle about contradictions being unacceptable) can actually be derived from contradictions. Thus, if dialetheism is the thesis that "there are true contradictions," it does not take too much logical maneuvering to get the result that dialetheists must be committed to the falsity of dialetheism.

> *Objection: False theory.*
> You are accepting a false theory! That is not rational.
>
> *Reply*. [. . .] Priest has sufficiently answered this question: that a theory is false doesn't take away from its truth. One might insist that any trace of falsity removes the rational acceptability of a theory, even if the theory is also true; however, such a claim begs the question or requires substantial argument (Beall, 2009, pp. 99–100).

Beall goes on to say that, while this response "is sufficient" on its own, he thinks he has a further defense against the worry given the details of his view. Unlike Priest's more expansive commitments (for example, to the Platonic existence of sets with inconsistent properties), the only true contradictions postulated by Beall's version of dialetheism are ones that result from sentences like (1) or (2). In Beall's terminology, these paradoxes are purely a consequence of the linguistic "behavior" of the truth operator. The resulting contradictions will not spill over into what the "base language" (i.e., sentences in which words such as "true" and "false" are not used). Since extra-linguistic reality will still be robustly contradiction-free, Beall's view is that the kind of dialetheias he does postulate do not cause trouble for any substantive form of reasoning. His way of putting the point is, "As far as I can see, there's no strong reason to think that rationality cares" about such a modest restriction of the LNC (Beall 2009, p. 100).

I will argue in later chapters that Beall's commitments leave him without a good way of avoiding base language contradictions. If I am right, his second reply is *wrong*. Although even if he is right, it is not clear how this point is supposed to intersect the original worry. Depending on how we parse the word "unacceptable," the worry either seems to be that **Reject(F)** dictates that dialetheism be rejected, or that what we can call **FailToAccept** dictates that we not accept it.

FailToAccept. One ought to not accept something if there is good evidence that it is false.

Either way, **Reject(F)** and **FailToAccept** (which are equivalent given classical logic and orthodox assumptions about the relationship between logic and reasoning) both sound like good rules of reasoning, and perhaps even like principles that might be partially constitutive of what we *mean* by calling a given piece of reasoning "good." Given the standard assumption that acceptance and rejection are psychologically incompatible, Beall's acceptance of dialetheism violates both.[14] This gives us at least a *prima facie* reason to suspect that dialetheism is irrational. What does the claim that "rationality doesn't care" about contradictions that fall outside of the base language have to do with anything? Surely the issue is whether the rules of rationality are being followed, not whether rationality "cares" about being disrespected in this way. To put the point less tendentiously and more carefully, whatever it means to say that "rationality doesn't care about the belief that α," either this claim entails "believing α is not irrational" or it does not.[15] If it does not, the reply simply does not intersect with the worry it is supposed to address. If it does, **Reject(F)** cannot correctly identify the circumstances under which a proposition should be rejected. (As we will see in a moment, given Beall's commitments,

[14] This would be the case *even if* Beall was willing to countenance contradictions about psychological states, which would presumably lie in the base language. If Beall both accepted and failed to accept dialetheism, the fact would remain that he accepted it, which would still violate **FailToAccept**.

[15] 2005 Beall would hold that there is an important sense in which this assumption is not warranted, but he reverses himself in *Spandrels of Truth*, unambiguously upholding every instance of the LEM.

moving from **Reject(F)** to **Reject(U)** would not help.) What principle *would* carve out those circumstances, then, and why is the alleged absence of base language contradictions relevant to it?

Beall's first reply, the one borrowed from Priest, is clearer. If we have good reasons to think that dialetheism is both true and false, then we have good reasons to think that it is true. Hence, by **Accept**, we should accept it. The obvious response is that its truth takes nothing from its falsity. Beall fudges this point slightly, framing the issue as one of the presence of "any trace of falsity" as if the issue were not one of *full-blown* falsity, but given Priest's commitments, the fudge is unnecessary. Priest explicitly rejects **Reject(F)** in *Doubt Truth To Be A Liar*, and he would presumably reject **FailToAccept** for similar reasons.

It is precisely here that Priest's moves become unavailable to Beall. Priest softens the counterintuitive blow of rejecting **Reject(F)**—surely, one might think, one of the constitutive goals of rationality is to *avoid falsehoods*, just as another is to *discover truths*—by accepting **Reject(U)**. This combination is only possible because Priest rejects the inference from falsity to untruth. Although as we have already seen, Beall accepts this inference. (He has to, given his core commitment to entirely unrestricted Capture and Release rules.) Thus, for Beall, accepting **Reject(U)** would entail accepting **Reject(F)** and believing that *dialetheism should be rejected*.[16]

He explicitly rejects that consequence in his discussion of the relationship between gluts and gaps. Any glut $\alpha \wedge \neg\alpha$ is equivalent by Double Negation to $\neg\neg\alpha \wedge \neg\alpha$, which can in turn be DeMorganed into $\neg(\alpha \vee \neg\alpha)$. Given the universal intersubstitutivity rules (and the analysis of falsity as the truth of negation), this entails the outright "truth value gap" $\neg(\mathrm{Tr}\langle\alpha\rangle \vee \mathrm{F}\langle\alpha\rangle)$.

> The point is correct. Any glut is also a gap, *so understood*. Of course traditional 'gap theorists', whatever else they might (or might not) say about 'gaps', are those who *reject* (what they want to call) 'gaps'. The gaps in my theory are ones that I fully embrace: they're ttrue and tfalse (Beall, 2009, p. 105).

Either this is a confession of irrationality or Beall embraces what we can call **Accept(G)** and **NoDilemmas**.

Accept(G). One ought to accept something if there is good evidence that it is a glut.

NoDilemmas. It is never the case that the same proposition should be both accepted and rejected.

As we have seen, **Accept(G)** is just an entailment of **Accept**. Beall's endorsement of Priest's reply to the objection about accepting a false theory makes it clear that he accepts it *because* it is entailed by **Accept**. But what about rejection?

Beall says not a word in *Spandrels of Truth* about **Reject(F)**, **Reject(U)**, or any other principle that might identify the circumstances under which rejection is rational. Perhaps he feels that it would be unnecessary, since (apart from the two paragraphs responding to the objection that knowingly believing a false theory is

[16] He would, in other words, be accepting the Rationality LNC which he endorses elsewhere in *Spandrels of Truth*, and which we discussed in Sect. 1.3.

irrational, and the aforementioned three passing mentions of specific claims that "need to be" or "must be" rejected) he does not make use of any rational-acceptability-and-rejectability talk.

He certainly makes use of *rejection* talk, though. For example:

> On my account, Curry sentences are false. I reject that they're ttrue (Beall, 2009, p. 33).

Although how literally can we take this? Are we really supposed to understand that JC as a discrete human individual has succeeded in forming psychological attitudes toward the infinite number of Curry sentences?[17] It is far from clear that this is even possible. (Think of intuitionist objections to infinitistic mathematics here.) Moreover, when he says that, or even something as innocuous as:

> I reject that vagueness involves gluts (Beall, 2009, p. 33).

.... he is not just reporting beliefs of his that he discovers through some introspective examination of the contents of his mind. If he were, why should anyone care?

This tension between the literal meaning of acceptance-and-rejection talk and what it is being used to indicate is even starker in the work of Hartry Field, a figure who (while not a dialetheist himself) has influenced Beall's version of dialetheism at least much as Priest. Given the way Beall approvingly cites Field's work throughout *Spandrels of Truth*, almost always accepting Field's way of framing things even when they come to different conclusions, Field's use of acceptance-and-rejection talk is an important clue about what is going in Beall's use of it. For example, after noting his rejection of both the claim that there are truth value gaps (understood as compound propositions of the form ¬(Tr<α> or F<α>)) and the claim that there are no such gaps, he says this:

> This might seem to indicate an unfortunate incompleteness on the part of the theory: it might seem that the advocate of the theory is remaining agnostic on the important question of truth-value gaps. I think this would be a mistake. Let's start with a simpler case, the case of the Liar itself. The theory doesn't accept *Q* and it doesn't accept ¬*Q*. Does that mean the theorist is agnostic about whether *Q*? (Field, 2008, p. 70)

Note the slippage back and forth between what "the advocate of the theory" or "the theorist" accepts, rejects, or is agnostic about and what "the theory" of paracompleteness accepts or does not accept. Something has gone seriously wrong here. Theorists can have psychological attitudes. *Theories* cannot. If "the theory doesn't accept *Q"* means anything, it means that the theory does not recommend the acceptance of Q. It is, in other words, a slightly odd way of talking about the claims about *rational acceptability* and *rational rejectability* entailed by a theory.

Similarly, when Beall says that "'symmetry' approaches" to Liar and Curry sentences "must be rejected—at least on my approach" (Beall, 2009, p. 34), he

[17] As we saw in 5.1, it is harder than it might look to precisely delineate the conditions for counting as such a sentence, but even on the narrowest definition, an infinite number of them will be constructable.

does not mean that it is somehow psychologically impossible to read *Spandrels of Truth*, be convinced by Beall's arguments for his version of dialetheism, and still fail to form an attitude of rejection to the claim that Liar and Curry sentences should be treated symmetrically. As Beall surely realizes, there is nothing impossible about such a reader forgetting all about the symmetry issue by the time she decides that she agrees with Beall about gluts, or remembering it but (while endorsing everything else he says in the book) remaining agnostic about that issue, perhaps mulling it over in the hopes of coming up with a better approach to that particular issue. What Beall means is that such an attitude would be irrational—that the claim that "Liar and Curry sentences should be treated symmetrically" is, in other words, *rationally rejectable*.[18]

The problems with Field's use of acceptance-and-rejection talk will have to wait for Chap. 7. Given the relevant differences between dialetheism and paracompleteness, the issues will be slightly different. In Beall's case, rejection-talk is used as an end-run around the problem with "(just) false" that we touched on briefly in our discussion of sentence (18). Since he regards truth and falsity (or, equivalently by his lights, truth and untruth) as compatible, Beall cannot say what he wants to say about Curry sentences and how they are unlike Liars by saying they are "false" or "untrue," so he turns to rejection-talk. This strategy becomes explicit when he turns to the "(just) false" issue. First, he argues that there *can't* be a difference between the contents of "true" and "(just) true" or between those of "false" and "(just) false." Remember that for Priest, falsehood and untruth are distinct and the former fails to entail the latter. As such, he can argue that "(just) false" *is* different from plain vanilla "false." The latter merely attributes falsehood to a proposition. The former attributes the conjunction of falsehood and untruth. Many critics (myself included) believe that this is an uninteresting distinction given the existence of sentences like (18), but it is at least *a* distinction. Given Beall's commitment to unrestricted intersubstitutivity, Priest's way of dividing up the concepts is unavailable to him. He does suggest, however, that there is a roundabout way to make "α is (just) false" do the work we want it to in "ruling out" α.

> While I maintain, as above, that *just true* is just *ttruth*, we can, if need be, appeal to a pragmatic difference. In the context of transparent truth theories, rejection is not simply acceptance of negation. In dialetheic theories, one may accept ¬α without rejecting α.[19] In dialetheic theories, one may accept ¬α without rejecting α. (In non-dialetheic theories, such

[18] His comment that the classical "'law' according to which $Pr(\alpha) + Pr(\neg\alpha) = 1$ is to be rejected" (Beall, 2009, p. 101) is perhaps even clearer. If there is a non-prescriptive way of glossing that comment, I am not imaginative enough to see what it could be.

[19] Beall frames the debate about the relationship between logic and the semantic paradoxes in just the way I introduced in 1.1, as a three-sided war between rejection of the LNC, rejection of the LEM and rejection of the principle that the truth operator should be "transparent" in the sense that the Capture and Release rules hold universally and we can thus always "see through" from Tr<α> to α and vice versa. While I find this way of introducing things initially useful, I ultimately reject it. As we will see in Chap. 8, I do not think we need to reject transparency to save the LEM and the LNC. Beall, however, takes it as a given that we do face a Sophie's Choice between classical logic and transparency, so he uses "transparent theories" to refer exclusively to the non-classical options.

> as Field's, one may reject α without accepting ¬α.) In this way, perhaps 'just true' carries pragmatic implicatures not carried by (an utterance of) 'ttrue', something to the effect that the speaker rejects the given negation. For example, if Max says that α is *just true*, he says nothing more or less than that α is ttrue; however, his statement carries the implicature that Max rejects ¬α. *Just true* and *ttrue*, on this account, remain equivalent; it's just that an utterance of the former carries 'autobiographical' implicatures that the latter doesn't (Beall, 2009, p. 52).

One might wonder, as above, whether such "autobiographical" information does the work Beall wants it to do. After all, even if one does not think that acceptance, rejection, and negation come apart in the specific way that the non-classical paradox-solver thinks they do, there might be a sense in which someone could be completely sincere in uttering a "Moorean contradiction" such as *α is true but I reject α*. Imagine someone who is aware that belief in ghosts is irrational and frequently utters the sentence *Ghosts aren't real* but is reduced to helpless terror by the very thought of entering into a haunted house. If we apply deeper tests for one's attitude toward a proposition than willingness to affirm or deny it in conversation, we might well say that this person does reject α. If the person in question has an epiphany about his own irrationality, "'Ghosts aren't real' but I reject 'ghosts aren't real'" might be an oddly phrased but admirably self-aware confession of irrationality. Closer to home, if dialetheism is true and there are devastatingly convincing reasons that we should all accept it, we could imagine a person in the equivalent position to the man who rejects "Ghosts aren't real." When she came out and *said*—never mind implicature!—"The Liar sentence is true and I reject the claim that it's untrue," she would be confessing her regrettable inability to get past the classical prejudices that lead her to reject dialetheias even when reason leads her to them. Surely the information conveyed by this sentence would be very different from what Beall means a speaker to be able to convey by an utterance of "α is just true." This seems like a perfectly good reason to think that Beall's use of acceptance-and-rejection talk must ground out in claims about rational acceptability and rejectability, and one that works even if we assume that a speaker like Max really is accurately reporting his psychological states when he says that various statements are "just true" and this carries the information that he rejects them.

Again, does this really work for sweeping statements about "rejecting Curry sentences" or "rejecting vagueness involving gluts"? Can an actual human individual *have* a specific psychological attitude toward all claims of the form *Such-and-such borderline application of this-or-that vague predicate is both true and false*, or all Curry sentences that *could* be constructed—even ones that are billions of words long? Clearly, we can form a psychological attitude such as rejection with regard to a sentence like (26) or (27).

(26) At least one of the infinite number of possible Curry sentences is true.
(27) At least one of the infinite number of possible sentences attributing gluttiness to a borderline application of a vague sentence is true.

In a certain way, sentences like (26) or (27) might be thought to be equivalent to the disjunction of attributions of truth to each sentence in the infinite list of sentences

in the relevant category, but there is also a fairly clear sense in which someone rejecting the (26) is *not* forming a specific mental stance to Curry sentences whose consequents would take millions of lifetimes to utter. (Recall the ambiguity about whether an agent can be said to be "thinking of" every instance covered by a universal generalization when they think of the generalization in our discussion Berkeley's Master Argument and Priest's dialetheic modification of it in Chap. 3) The important point is that Beall's strategy for using rejection-talk to recapture the lost powers of non-dialetheic negation fails *no matter what* one makes of this issue. Either humans *can't* reject each member of such infinite lists, or acceptance and rejection *can* overlap in individual human minds after all. Either way, the strategy is a non-starter.

Beall could avoid this result by rejecting either of the following claims.

(i) The transmission of psychological states across generalizations about truth works in both directions—in other words, if rejecting a sentence like (26) or (27) amounts to rejecting every sentence of the relevant type, then *accepting* a universal generalization like (28) amounts to accepting every sentence of the relevant type:

(28) All Excluded Middle instances for borderline applications of vague predicates are true.

(ii) It is possible to accept an attribution of truth to a set of sentences without being aware of the content of each one of those sentences.

The problem is that denying (i) looks hopelessly *ad hoc* and unmotivated, and denying (ii) would rob Beall's core assumptions about truth of much of their justification. Beall is fond of motivating his views on truth with a theological metaphor. God could describe reality in its every nook and cranny without making use of the truth operator, since that predicate does not identify some property like correspondence or coherence that would be left out of a "T-free" description. Humans are another story.

> We are unlike God in that respect we need a device that enables us to overcome finite constraints in our efforts to describe the world. That device is 'true' or, for clarity, 'ttrue' (for 'transparent truth'), a device introduced via the rules of intersubstitution [. . .] Even in daily discourse, let alone theoretical pursuits, we want to say (what, using *ttruth*, we say when we say) that all of So-and-so's assertions are ttrue, or that some claim in Theory X is tfalse (that is, that the negation of something in X is tttrue). As above, were we God, or even just beings with infinite time or capacities, we wouldn't need to use 'ttrue' in such generalizing contexts; we could simply assert each of So-and-so's assertions (or the negations thereof) (Beall, 2009, p. 1).

Let us pause to think about the specific nature of some of those limitations. The part about infinite time is clear enough. The first premise of the most famous syllogism in history is sadly true. Even if it were not, it would hardly matter. When Freddie Mercury asked "Who Wants to Live Forever?" on the *Highlander* soundtrack, one assumes that the implied appeal of immortality has to do with doing things that we *enjoy* forever. No one would want spend an eternal life reciting the infinite instances of Excluded Middle. That brings us to one finite human capacity.

Sooner or later, we get bored with things. What other limited capacities are relevant? One (closely related to our inability to stay interested in things forever) is our distinctly finite capacity to remember things. It is entirely possible to be guided through a twenty-step argument, to be individually convinced by every step, and to remain convinced ten minutes later that the argument is sound—and hence that *all of the premises are true*—without being able by then to recall the specific content of those premises. This is true even when one fully understands every detail of each premise when first hearing it. Given the sort of intellectual division of labor necessary for specialized theoretical pursuits such as the sciences to be a source of public knowledge, it had better be not only possible but epistemically respectable for a layman to believe in theories based on testimony even when he not only could not repeat every proposition that is part of each theory but would not be able to understand a great many of them when he did repeat them.

Well, if that is possible, so is overlap between acceptance and rejection! Ask a thousand college graduates (who did not major in physics) whether they accept or reject Einstein's Special Theory of Relativity. Then ask them if they accept or reject the possibility that two twins could have been alive for objectively different amounts of time. It would be a minor miracle if at least a few respondents in such a poll did not sign onto the former but reject the latter.

If such overlap is possible, rejection cannot do the work Beall wants it to do. Conversely, if acceptance and rejection really are mutually exclusive states, it must not be possible to "transmit" rejection from a sentence like (26) to each of the infinite number of Curry sentences, and "I reject all Curry sentences" must either be flatly false or a slightly off-point way of saying "all Curry sentences are rationally rejectable."

Either way, all of the considerations about the relationship between dialetheism and rational dilemmas that sank RAA(B) and Principle R as strategies for recapturing RAA(N) and Disjunctive Syllogism are equally fatal to Beall's attempt to use rejection talk to recapture the lost exclusionary power of negation.

5.6 Francesco Berto's Critique of Dialetheist Rejection-Talk

Everything that has been said about Principle R in the past four sections rests on the assumption that Priest, Beall and Field are right to make a sharp distinction between *accepting* $\neg\alpha$ and *rejecting* α. All three theorists assert that the former is psychologically compatible with accepting α and the latter is not. Both halves of this assumption can be (and have) been questioned.

Frege teaches us that the laws of logic are "the laws of thought" only in a normative sense—they are laws that we must follow to arrive at the truth—but the contrary "psychologist" view has deep roots in the history of philosophy. A typical "psychologist" view about logic is that believing logical contradictions is not just

wrong or irrational but psychologically impossible. As Francesco Berto points out, rebels against classical orthodoxy from Łukasiewicz to Priest have rejected such claims. . .and their reasons for rejecting them seem to transfer over quite neatly to the context of the alleged psychological incompatibility of acceptance and rejection!

> [Edwin Mares] calls rejection-consistency (RC) a "principle of coherence," and takes this incompatibility as simply constitutive of the notion: "by virtue of the nature of rejection, it is a necessary condition on a sentence's being rejected that it also *not* be accepted". Is there any such nature? [. . .] Husserl and Łukasiewicz argued against pragmatic and psychological versions of the (LNC) because of the weakness of their warrants. Such principles are, at most, inductive theses based on considerations of empirical psychology. They do not discover any *nature* or *essence* at all.
>
> Can the dialetheist reply that the mutual exclusiveness of mental acceptance and rejection is ascertainable by introspection, and link this to the infallibility of some First Person Authority? She may claim that our knowledge of our own mental states is infallible or incorrigible, at least in this respect: the fact that acceptance and rejection cannot co-occur simultaneously in the same mind, and with respect to the same sentence, is self-intimating. However, I doubt it. Against commentators who maintain that the Aristotelian (LNC) is a psychological law to be established by introspection, Priest observes that "the unsatisfactoriness of trying to establish psychological laws in this way hardly needs to be belabored" (Berto, 2007, pp. 298–299).

One way of reinforcing Berto's point here is to compare the Mares's claim about the relationship between the psychological states of acceptance and rejection to claims that a pre-Einsteinian physicist could make about the nature of time. Confronted with the idea that objectively different amounts of time could pass for two twins, a Newtonian might be tempted to say that "passing at the same rate for everyone everywhere" is part of the essence of time—it would not be "time" if it were not a universal invariant! Depending on what one makes of some sticky issues in the philosophy of language, this response might not even be entirely wrong. On the view that "time" is a rigidly designating natural kind term that always picks out the fourth dimension of space-time, Einstein showed that time does not work the way the Newtonian says it does, but on a view according to which the application conditions for a term such as "time" are internal to a given theorist's conceptual scheme, Einstein simply showed that "time" in the Newtonian sense does not exist. Only time in the Einsteinian sense exists.

Similarly, Berto's criticism of Mares can be framed as a dilemma. Assuming that the LEM applies to psychological states, theorists either can or cannot be mistaken about the application conditions for a concept such as "rejection." If they can, the warrant for the belief that rejection works in the way that Mares (and Priest and Beall and Field) think it does is very weak. It they cannot, the warrant for the belief that "rejection" in the Marest/Priest/Beall/Field sense *exists* is very weak!

It could be objected that we can offer a clear and precise definition of "rejection" that both satisfies the requirement of incompatibility with acceptance and is not susceptible to the worry that empirical psychology does not confirm the existence of such a thing. Start by letting rejection$_1$ be "rejection" in some sense whose application conditions do not include a rule against the same proposition being

simultaneously accepted and rejected in the same mind. (Perhaps to reject$_1$ α is simply to accept ¬α!) Now define rejection$_2$ as follows:

S rejects$_2$ α iff S rejects$_1$ α and S does not accept α.

This is robustly incompatible with acceptance just so long as the psychological realm is free of dialetheias. . . .and I can hardly claim that it is never the case that anyone rejects$_1$ α and does not accept α. Presumably, even if the psychological form of the LNC is dubious, it is at least true that in the *majority* of the instances in which someone rejects$_1$ some α, they also do not accept it!

However, on this way of defining rejection$_2$, the objection to Beall developed in 5.5 is strengthened. A dialetheist may be in a good psychological position to know that she rejects$_1$ all Curry sentences or all contradictions regarding borderline applications of vague predicates, but how can she be so sure that she rejects$_2$ them? If the truth operator is good for anything, it is for giving us the ability to signal our acceptance of claims that beings with our limitations cannot express directly. This creates the ever-present possibility of accepting claims we would not accept if we realized that they contradicted other claims of which we are more certain.

Put more simply, if rejection$_1$ is compatible with acceptance, so is rejection$_2$!

Chapter 6
Dialetheism, Rejection, and Probability Theory



Abstract If Principle R and related attempts to substitute rejection-talk for negation-talk fail, what remaining options does the dialetheist have at her disposal? Priest has suggested that, since the statistical frequency of true contradictions is quite low, classical laws that are not deductively valid if dialetheism is true nevertheless have probabilistic force. In Priest's terminology, such rules may, while being invalid in the strict sense, still be "quasi-valid" and thus appropriate for use in most contexts. When this strategy is held up to the light, significant problems emerge. First, if true contradictions exist in all the domains in which Priest says they exist, there is no particularly good reason to think their overall statistical frequency is very low. Priest has argued that the "observable world" would still be contradiction-free, but his arguments are weak. The probabilistic strategy appears to be as dubious as the strategy premised on emphasizing rejection. Other strategies, some proposed by Priest and others proposed by other dialetheists, which rely on non-monotonicity, logical pluralism, local validity, or a kind of super-negation postulated by Francesco Berto, are similarly shown to fail.

Keywords Probability · Pluralism · Monotonic · Non-monotonic · Validity · Quasi-validity

6.1 Probability and Quasi-Validity

Priest calls those classical inferences he hopes to recapture "quasi-valid." When he abandons the corpse of Principle R in the second edition of *In Contradiction*, he softens the blow by reminding the reader that in the first edition he proposed *two* strategies for recapturing such inferences.

> Section 8.3 uses principle *R* as part of the justification for using quasi-valid inferences in consistent situations. If principle *R* cannot be sustained, then neither can this justification. Note, however, that the probability considerations of Sect. 8.3 still stand (Priest, 2006b, p. 276).

What were those considerations? The core move consists in downgrading rules such as Disjunctive Syllogism from *deductively valid* to a merely inductive status,

© Springer Nature Switzerland AG 2022

B. A. Burgis, *Logic Without Gaps or Gluts*, Synthese Library 458,
https://doi.org/10.1007/978-3-030-94624-1_6

such that the truth of the premises of such an argument does not logically guarantee the truth of the conclusion but does make it highly probable. "Thus, one can employ quasi-valid inferences to take one from things with a high probability to other things with a high probability."

Applied to Disjunctive Syllogism, this gets us what we can call Principle P:

> Principle P: If there is a high probability that a disjunction is true and a high probability that one of the disjuncts is untrue, there is a high probability that the other disjunct is false.

Applied to RAA(N), it gets us what we can call RAA(P):

> RAA(P): If there is a high probability that an entailment conditional is true and a high probability that its consequent is false, there is a high probability that its antecedent is false.

Priest says very little about what notion of probability he is assuming, but the differences between the options matter a very great deal. To start with, notice that if we fully embrace Bayesian epistemology and regard the probability of a proposition as the logical relationship between the proposition and the evidence for it, such that what it is for it to be rational to accept a proposition just *is* for there to be a high probability that the proposition is true, then Principle P is just Principle R in disguise, RAA(P) is just RAA(B) in disguise, and Priest has really offered us not two strategies for classical recapture but two ways of talking about the same strategy. In fact, in the first edition, he suggests just such an equivalence:

> We have seen that we can allow ourselves the use of a quasi-valid inference provided we may reject its crucial contradiction, or assign it a sufficiently low probability (which, arguably, comes to the same thing) (Priest, 2006b, p. 115).

Although if they do "come to the same thing," Priest is wrong about "the probability considerations of Sect. 8.3 still stand[ing]" in the absence of Principle R. Perhaps he changed his mind about probability theory in between the first and second editions. In any case, any dialetheist who wants to recapture classical reasoning by reinterpreting "quasi-valid" classic inferences in a probabilistic way had better have some theory of probability other than the straight-up Bayesian view just considered.

We will return to the issue of what alternative theory might suffice for their purposes later in the chapter. First, let us look at the reasons Priest gives for thinking that Disjunctive Syllogism and the rest are at least *likely* to take us from truth to truth in most contexts.

> The reason is a simple one: The statistical frequency of dialetheias in normal discourse is low. Dialetheias appear to occur in a quite limited number of domains: certain logico-mathematical contexts, certain legal and dialectical contexts. . .and perhaps a few others (Priest, 2006b, p. 116).

As we saw in Chap. 2, the list of domains in which Priest takes himself to have discovered dialetheias is far more extensive than one might think from this passage. One of those domains is the domain of *things that change,* since he takes the "paradoxes of motion and change" to establish dialetheias. Given a Quine-like view, according to which mathematical objects are indispensable but other

apparently non-physical phenomena (such as mental states) can be understood in purely physicalist terms, one might think that reality is divided between the world of changing physical objects on the one hand and the changeless realm of mathematical *abstracta* on the other. Well, Priest has argued for the existence of dialetheias in that second realm too, based on both Gödel's first incompleteness theorem and the antinomies of naïve set theory. That would not seem to leave us with any contradiction-free domains!

Moreover, even if Priest *did* only argue for the existence of dialetheias in a few domains here and there, there remains a considerable logical leap from "there is only good evidence that some contradictions are true in domains X, Y, and Z" to "there is good evidence that true contradictions are *only* true in domains X, Y, and Z." What could ground this degree of confidence that the dialetheias discovered in the first generation of philosophical research on the subject are the only ones that exist? If dialetheias have only been discovered so far in domains X, Y, and Z, why not just remain agnostic about the consistency of A through W? The absence of evidence for the presence of dialetheias in these domains would only count as evidence of their absence in combination with some general reason to believe that, all else being equal, contradictions are unlikely to be true. . . . but that is exactly what the alleged low statistical frequency of contradictions was introduced to establish!

At this point, some readers might be troubled by something like the following thought. Perhaps dialetheias are like leprechauns. If we lack evidence for the existence of leprechauns, surely the rational course of action is not to suspend judgment but to actively believe that *there are no leprechauns*. Indeed, Priest has encouraged this sort of comparison, turning the neat phrase that "contradictions should not be multiplied beyond necessity" (Priest, 2006b, p. 71). Although why not? Priest's maxim suggests that something at least *like* ontological parsimony should be guiding our theorizing about the consistency of various domains, but there is certainly no literal issue of ontological bloat here. Take the following argument:

P1: John is either at home or at the library.
P2: John is not at home.
C: John is at the library.

Assume that the premises are both true and the conclusion is not—that, in other words, the first disjunct is a dialetheia and the second disjunct is (just) false. The items to which this diagnosis ontologically commits us are John and John's home. By contrast, a classical logician who diagnosed this instance of Disjunctive Syllogism as sound would be committed to John, John's home, *and the library*. Libraries may not fall into a different entity-category in the right way for the first list to be more ontologically parsimonious than the second, but the second certainly is not more parsimonious than the first.

There is also at least one fairly clear disanalogy between an Ockham or a Quine talking about multiplying entities and Priest talking about multiplying contradictions. It is trivially easy to generate an infinite number of Liar-like sentences, as in the following list:

(L1) The sentence labeled (L1) is false.
(L2) The sentence labeled (L2) is false.
(L3) The sentence labeled (L3) is false.

. . . . and so on, such that for every rank (N) sentence (LN) will "say of itself" that it is false. Given a dialetheist analysis of such sentences—and remember, no dialetheist has argued for dialetheism on grounds that do not at least *include* the Liar Paradox—it follows that there will be infinitely many dialetheias. *Whatever* moves one makes to limit the number of contradictions to which one is committed, given a dialetheic solution to the paradoxes, the number remains stubbornly infinite.[1]

This might not look like a disanalogy between Priest's and Ockham's principles if we take "entity" seriously. Anyone who accepts numbers into their ontology had better be prepared for an infinite number of numbers, for instance. In the case of the ontological principle, parsimony applies not over *individual* entities (such that a Platonist who thought the series of natural numbers went straight from two to four had an advantage in parsimony over the Platonist who accepted the existence of the number three) but *categories* of entities. When Priest invokes his Ockham-like principle, he seems to be talking about *individual* contradictions. For example, take his discussion of the "exclusion" principle ($Tr\langle\neg\alpha\rangle \rightarrow \neg Tr\langle\alpha\rangle$). He considers an argument for it based on the assumption that Contraposition is valid for whatever conditional is used in the T-Schema, but (sensibly enough, given his dialetheism[2]) he argues that this assumption is unwarranted.

> We have seen that the exclusion principle spreads contradictions, and in virtue of the failure of the above argument it would seem to do so beyond necessity [. . . .] On the basis of this, I tentatively reject the exclusion principle. (Priest, 2006b, p. 71)

The specific way that the exclusion principle spreads contradictions is that it allows us to go from an "internal" contradiction $Tr\langle\alpha\rangle \wedge Tr\langle\neg\alpha\rangle$ to an "external" contradiction $Tr\langle\alpha\rangle \wedge \neg Tr\langle\alpha\rangle$. At first glance, this might look like a case of Priest trying to avoid the multiplication of the *kinds* of contradiction to which he is committed. If so, there are two problems. The first is that, as Priest grants in several places, he is *already* committed to "external" contradictions. To generate one that *any* partisan of the dialetheic analysis of the Liar Paradox must accept, simply let α *be* $\neg Tr\langle\alpha\rangle$. We could easily rephrase every sentence on the list of sentences starting with (L1) above to generate an infinite list of sentences of that type. Thus, Priest is

[1] At this point, another reasonable criticism of Priest might occur to some readers. Given that both the category of propositions-in-general and the subcategory of glutty propositions will presumably be infinite, is it possible to make sense of this talk of a low statistical frequency in the first place? While I have said that this criticism is "reasonable," I am not confident that it is successful. Priest might be able to make sense of frequency-talk here with something like the following idea: What it means for there to be a low frequency of gluts is that, for any sufficiently large and properly randomized *finite* selection of propositions, the rate of gluts will be low. A lot more conceptual work would have to be done to make sense of this idea, but the project does not necessarily seem hopeless.

[2] See Sect. 5.2.

not trying to avoid the necessity of believing some (or even an infinite number) of "external contradictions" are true, he is simply trying to avoid adding more of them to the already transfinite pile.

The second problem is that if these are truly contradictions of different kinds (and, in the absence of the exclusion principle, they are not even logically equivalent), this undermines the uniformity of his solution even to different versions of the Liar Paradox. Some Liars get solved by postulating what we can call Type-1 Gluts (they are true and false) while others get solved by postulating Type-2 Gluts (they are true and untrue).[3]

In any case, the principle of "not multiplying contradictions beyond necessity" cannot do the work of rendering Priest's argument from the limited number of domains in which he takes us to have good evidence for the truth of contradictions to the claim that the statistical frequency of contradictions is low non-circular. When the curious reader flips to Sect. 8.4 to find Priest's argument for that principle, here is what they find:

> Given that the statistical frequency of dialetheias is low, and given a contradiction about which we have no particular information (or no time to take account of the further information there is), two things follow. First, this statistical frequency provided the epistemological probability of the contradiction. The probability that this contradiction is a dialetheia is, therefore, low. Secondly, it is rational to reject the contradiction. (See the discussion of rational rejection in Sect. 7.4.[4]) We might summarize this situation in the razor: contradictions should not be multiplied beyond necessity (Priest, 2006b, p. 116).

The order of explanation in Priest's epistemic picture goes from low frequency to low probability and from there (not even as a further claim, but as a summary of the situation) to the anti-multiplication principle. The third link in that chain cannot be made to support the first one without turning the whole thing into a loop.

He does offer two more arguments for the low frequency claim.

> The claim that the frequency of dialetheias is low seems fairly obvious, but, inevitably, it will be asked how one knows this once it has been conceded that there are some. The simplest argument for this is a head-count. The reader is invited to consider a random sample of the assertions he has met in the last few days and see what percentage might reasonably be thought to be dialetheic. If it's more than a handful, this probably means that the reader has been reading a book on paradoxes. (Maybe this one.) (Priest, 2006b, p. 116)

This, of course, is simply a restatement of the argument from the limited number of domains in which dialetheias "appear" to occur. The points made about the first version of the argument all continue to apply. The restatement does, however, draw out a further problem. When Priest argues that we should take seriously the possibility that truth and falsity can overlap, he is proposing a profound revolution in our concepts. . .as he himself is prone to point out in his bolder and more expansive modes. Here he is in the first edition of *In Contradiction*:

[3] Thanks to the anonymous reviewer for suggesting this last point.

[4] See the discussion of that discussion in Chap. 5.

> [The structure of inconsistency] is important because it bears centrally on an adequate account of sets, of semantics, of time and change, of normative systems, and *possibly many other things* [my emphasis]. All this I have tried to show in the foregoing pages. However, these investigations are but a start. The realm beyond the inconsistent is a continent on whose shore we have just alighted. This book does little more than forge some of the basic equipment required or the exploration, and test it on some important areas of the land. I do not claim that the equipment is perfect—far from it. But it allows us to make a start in exploring this new and unfamiliar terrain—the transconsistent (Priest, 2006b, p. 209).

And again nineteen years later in *Doubt Truth To Be a Liar*:

> The Law of Non-Contradiction has been well-entrenched in western thought—and so science—since the canonization of Aristotle, whose defence of the law has rarely been challenged. Hence, scientists and philosophers have not been prepared to brook the thought that an inconsistent theory of any kind might be true. But subscribing to the Law is not rationally mandatory [. . .] as the development of paraconsistent logics has showed a large role in showing. Once this fact is digested, scientists may well—justifiably—take a different attitude toward inconsistent theories of the appropriate kind. Indeed, they may even develop inconsistent theories, if these have the right empirical consequences, just as paraconsistent logicians have articulated inconsistent theories of semantics to handle the paradoxes of self-reference (Priest, 2006a, pp. 149–150).

If we have just begun to explore this vast new continent, if scientists have not yet "absorbed" the possibility of dialetheias, then why should anyone take a "head-count" of the dialetheias they have discovered so far as a meaningful indication even of what their beliefs would be once the absorption process is complete...much less a meaningful indication of the number of dialetheias that may lurk undiscovered in any number of domains? Indeed, the picture he paints above considerably understates the difficulty which would be involved in reasoners of all kinds fully "digesting" what Priest has called "the paraconsistent revolution" (Priest, 2006a, p. 150). Belief in the impossibility of contradictions may for the sake of argument be a *mistake*, but if so, it is a mistake with far deeper roots than deference to Aristotle. The near-universal (and usually implicit) assumption that the truth is consistent—that, for example, once it has been established that some claim is false, there is no further question to be asked about whether it is true—can no more be attributed to Aristotle than the (almost) equally widespread belief in the existence of a mind-independent external world can be attributed to Descartes. ("Almost" because while both assumptions are equally entrenched in pre-philosophical common sense, the existence of a mind-independent external world has been questioned far more in the history of philosophy and religion than the assumption that the truth is always consistent!) Even the small handful of human beings—mostly Australians—who are intellectually convinced of the possibility of exceptions to the LNC can hardly be said to have fully digested this revolution in their concepts. Most of them do not even follow Priest in accepting a dialetheic solution to Russell's Paradox, surely a quite modest extension of the dialetheic solution to the Liar Paradox. What can be learned from counting the number of dialetheias whose truth has been discovered in the course of reasoning—whether scientific, mathematical or everyday—that was guided by the assumption that inconsistent "possibilities" should always be rejected?

It could be objected here that *if* the probability of contradictions being true is sufficiently low, then we should not expect the assumption that the truth about any particular domain is consistent to have significantly distorted the results of our reasoning about that domain. That is so. However, if the "head-count argument" is being used to support the claim that the statistical frequency of dialetheias is very low, and the statistical frequency claim is in turn being used to support the claim that the probability of any given contradiction being true is correspondingly low, the low probability claim cannot (non-circularly) come to the rescue of the headcount.

Priest has one more argument for low frequency. As we will see, it comes to much the same thing.

> A more general argument for the low frequency of dialetheias in ordinary discourse can be given as follows. It is a fact, to be verified by simple observation, that people do commonly use quasi-valid arguments. DS, for example, is used commonly in mathematics, as well as in more common-or-garden reasoning. Moreover, it is also fact, to be verified in the same way, that such reasoning is by-and-large successful, or at least does not lead to recognizable errors. I know that you are either at home or in the supermarket. I ascertain that you are not in the supermarket. I go to your house and lo! You are there. Now if dialetheias were common, we would expect quasi-valid inferences to go wrong quite frequently. But they do not. Hence, they are not common (Priest, 2006b, p. 116).

The first thing to notice here is that the form of the argument is *Modus Tollens*, which in turn is just Contraposition plus *Modus Ponens*. As I argue in Sect. 4.2, Contraposition cannot be valid if the dialetheic solution to the Liar Paradox is correct. Of course, if Priest's argument above succeeds, Contraposition and hence *Modus Tollens* could be probabilistically quasi-valid, and Priest's argument would at least demonstrate that the statistically frequency of contradictions is *probably* low. However, to help oneself to the assumption that "quasi-valid" arguments have even probabilistic force in the course of *arguing* for the low statistical frequency (and hence low epistemic probability) of counterexamples begs the questions as thoroughly as using Disjunctive Syllogism in an argument against dialetheism.

Second, the core of the argument is just another headcount. In this case, we are counting counterexamples to DS and other forms of "quasi-valid" reasoning, not true contradictions, but even that distinction collapses when we consider that every true contradiction would by definition be a counterexample to some possible instance of Disjunctive Syllogism. As such, all the points made above apply here with equal force.

Finally, Priest's assumption that "Disjunctive Syllogism failure" is rare puts the epistemic cart before the horse. When mathematicians use Disjunctive Syllogism in their proofs, it is not as if they can visit Platonic Heaven afterward to make sure their conclusions correspond with mathematical reality. We believe the proof to have been a success *because* we are justified in believing the premises and we are confident that DS will never take us from true premises to false conclusions.

It might be thought that a partial exception to the general impossibility of discovering that DS has steered a mathematician wrong might come when a proof involving DS leads to a result that contradicts some other important result. As it happens, competing plausible assumptions do quite often deliver incompatible

results by means of equally impeccable-looking mathematical proofs. Instead of inferring the shakiness of the inferential machinery used to generate these results, however, the competing plausible assumptions are simply partitioned into distinct "axiomatic systems." Similarly, mathematicians quite often derive outright contradictions from seemingly plausible premises. . .on purpose, for the purposes of showing that those premises must be rejected. Sometimes DS or other rules considered by Priest to be "quasi-valid" are used in these proofs. If there are true contradictions, who knows how often such "proofs by contradiction" are mistakenly taken to establish the falsity of the premise rather than the truth of the relevant contradiction? Again, it is not as if mathematicians could check Platonic Heaven to see if they made a mistake.

Still, things might seem to be different when it comes to the "common-or-garden" everyday uses of Disjunctive Syllogism. Platonic Heaven might be empirically inaccessible, but the supermarket is not. Although even here, Priest's description of the situation fails to capture our ordinary epistemic practices. It happens quite frequently that we take ourselves to have good reasons to believe $\alpha \vee \beta$, we take ourselves to have excellent reason to believe $\neg\alpha$, but lo! β turns out to be false. Since we normally disregard the option of taking α and $\neg\alpha$ to be simultaneously true, the normal procedure in this case is to assume that we must have missed something in our initial inventory of the options. No matter how good our initial reasons for taking α and β to be jointly exhaustive, the revelation that $(\neg\alpha \wedge \neg\beta)$ always convinces ordinary reasoners that we must have been missing a third possibility γ—even if they have no idea what γ even *could* be.

In response to a similar criticism from Doris Olin, Priest says in the second edition of *In Contradiction* that, while this may indeed be the usual procedure in such situations, "the point is that, when the premises of the reasoning are well-grounded, we rarely have to backtrack in that way" (Priest, 2006b, p. 276 n. 39). This seems flatly wrong to me. Ordinary reasoners (and even scientific reasoners) backtrack in that way all the time. Sometimes a plausible γ is discovered. Sometimes, none is found and we all accept it as an instance of the routine incompleteness of human knowledge. In the case of "supermarket-or-home"-level reasoning, we usually do not think much about it one way or the other.

Of course, a dialetheist could grant this without granting that all or even most of these cases are instances of Disjunctive Syllogism failure. After all, there is well-grounded and there is *well-grounded*. We are rarely entitled to absolute certainty about the premises of even excellent arguments, and that goes double for Disjunctive Syllogism. Whether we are dogs chasing scents or mathematicians chasing interesting results, disjunctive lists of options usually play a role in our reasoning when we are initially undecided about which option is correct. So even if the premises of a Disjunctive Syllogism instance *can* be true without the conclusion being true, it is surely also at least *very often* the case that when a given instance of the argument form takes us from premises for which we have otherwise excellent evidence to conclusions that are clearly false, the fault lies not in Disjunctive Syllogism but in ourselves. False premises are (probabilistically) supported by good evidence all the time. From the perspective of the classical logician, *all* cases where a reasonably

well-grounded application of Disjunctive Syllogism takes us to false conclusions are cases where at least one of the premises was mistaken. All the dialetheist is entitled to say, at least if she restricts herself to cases where no suitable γ has been found, is that *some* percentage of cases are like that. She simply has no way of knowing how often Disjunctive Syllogism fails.

6.2 Dialetheism and the Consistency of the Empirical World

In the preceding discussion, I assumed that our inability to check to see if contradictions were true (and hence if our use of Disjunctive Syllogism or *Reductio Ad Absurdum* had steered us wrong) was limited to the mathematical realm. It is worth pausing to ask, however, whether it would be possible to empirically "check" to see if inconsistent states of affairs held even in the observable physical world.

If such checking turned out to be impossible, this would be a very bad result for the statistical frequency claim. Indeed, it would be a very bad result for the general compatibility of dialetheism with anything recognizable as ordinary human reasoning. To see why, consider the following (hardly unusual) situation. We have some evidence *for* some theory (or just for some garden-variety belief that does not deserve the honorific "theory") and some evidence *against* it. If we only had the first piece of evidence, we would be justified in concluding α. If we only had the second piece of evidence, we would be justified in concluding ¬α. Instead, we reserve judgment. But should we? Normally, if we have what would be sufficient evidence for α (when considered in isolation from our evidence for β) and we have what would be sufficient evidence for β (in isolation from our consideration of α), unless there is something about the idea of both being true that makes the conjunction less plausible than either of the conjuncts, we accept (α ∧ β). If β=¬α, the universal logical impossibility of α and β being jointly true would certainly be a good reason not to accept (α ∧ β). So, perhaps, would the low epistemic probability of true contradictions, grounded in their low statistical frequency—but right now we are trying to see if the dialetheist is entitled to the frequency claim in the first place. Given that the human epistemic condition is to be confronted with mixed and murky evidence on all sorts of subjects on a regular basis, things do not look good.

Moreover, even *if* they looked better, what of it? Would it matter? Let us assume for the sake of argument that we can help ourselves to Conditional C.

Conditional C: If there were true contradictions about the empirical world, we would be able to point to empirical evidence of their existence.

Let us further assume Empirical Claim E, despite everything we have said about the combination of dialetheism in the philosophy of logic with the real-world ubiquity of "mixed and murky" empirical evidence.

Empirical Claim E: It is not the case that we have empirical evidence of the existence of true contradictions about the empirical world.

The problem is that the inference from C and E to the claim that the empirical world is contradiction-free would still rely on *Modus Tollens*. As we have already seen, dialetheism and *Modus Tollens* do not mix. As such, there is a sense in which this section could end here. Graham Priest has made multiple arguments for Conditional C, but even if he is right, without *Modus Ponens* (and Empirical Claim E) they will not get him the result that he wants. Any reader sufficiently convinced of the dialetheic unavailability of *Modus Tollens* (and/or Empirical Claim E) to be uninterested in following a detailed refutation of Priest's argument for Conditional Claim C is thus invited to skip ahead to 6.5. Meanwhile, I hope to establish that *even if* I am wrong about the dialetheic unavailability of E and *Modus Ponens,* Priest is unsuccessful in his argument for C.

6.3 Priest's Case for the Consistency of the Empirical World, Part I (The Argument from Pierre and the Conjunction Principle)

Let α be a proposition describing an empirical state of affairs. We might think that if $(\alpha \wedge \neg\alpha)$ were true, we would only be able to observe that α. Priest disputes this, leaning on Sartre's example in *Being and Nothingness* about looking for Pierre in the café.

> When we perceive, we perceive that something *is* the case. Can we also perceive that something is *not* the case? Some have thought not. We can perceive only that something *is* the case. For example, we cannot see that something is not green. We can only see that it is red. Any judgment to the effect that it is not green has to be added to what we see by inference. This, as we now see, is false. I can see directly that something is not green. Or consider another example: you enter a room; the whole room is visible from where you stand; there is no one there. You can see that Pierre is not in the room. No Pierre-shaped objects meet the eye. Even the very distinction between seeing what is the case and what is not the case is a false one. Some seeings are both. With respect to actual physical objects, to be transparent is not to be opaque, and vice versa. But you can see that something is transparent and you can see that something is opaque (Priest, 2006a, p. 62).

Let us go ahead and assume that Priest and Sartre are right about Pierre. What one makes of Priest's claim that the *distinction* is false gets down to an issue about semantics, not the philosophy of perception. Is there a hard and fast distinction between "basic" predicates and their negations, or do English and other natural languages present us with pairs of interdefined predicates such that each is logically equivalent to the negation of the other, but there is no non-arbitrary way of designating one as "more basic"?

Whatever one thinks about this issue, I will just say that the pair transparent/opaque is a particularly unpromising candidate. Imagine explaining the meanings of "opaque" and "transparent" to an otherwise competent speaker of English who lacked both vocabulary terms. If you started with "opaque," you would still have to invoke the *concept* of transparency. ("*Opaque* means you cannot see through it.")

If you started with "transparent," not only would there be no reason to detour through the concept of opaqueness, but it would be difficult to do so without resorting to double negation.

Of course, for the purposes of Priest's point about the hypothetical observability of contradictions in the empirical world, he does not need the distinction to be false. He only needs us to be able to (non-inferentially) verify "negative" facts by perception.

> Next, if α and β are observable states of affairs, then so is their conjunction. We can see that a thing is a cat; we can see that the thing is black; hence we can see that it is a black cat (Priest, 2006a, p. 63).

If this principle held universally, Priest would have a point. If Pierre is at the café, we can observe that. If Pierre is not at the café, we can observe that. If we can observe that an object satisfies one predicate, and we can observe that it satisfies some other predicate, we can observe that it satisfies both. Hence, if Pierre were ever both at the café and not at the café, would be able to observe *that*. We never do observe such combinations, so (putting aside the question of how to justify our use of Contraposition in that "so") contradictions do not arise about such issues. Thus, if we are in a mixed-evidence situation, and we have a good reason to believe Pierre was at the café and a good reason to believe that he was not (and both reasons would be sufficient if not for the contrary evidence), the right procedure is the usual one. We should reserve judgment instead of simply endorsing the contradiction.

All this, as I have said, would go through *if* the principle about conjunction held universally. As Priest concedes, it does not.

> This principle needs appropriate qualification. If α and β record states of affairs that arise only at different times, or at places such that one cannot see both at once, then the principle is obviously not true. More importantly, it can be the case that one of the conjuncts occludes the others. Thus suppose that α is 'the street sign is blue' and β is 'it is night'. Both of those states of affairs are observable, but if it is indeed night, one cannot see that the street sign is blue (Priest, 2006a, p. 63).

At this point, one might think, Priest is in real trouble. After all, a natural thought here would be that if *anything* would occlude our ability to observe that Pierre is not at the café, it would be Pierre's visible presence at the café.

Priest will have none of this.

> [L]et α describe an observable state of affairs, that *x* is happening here and now. Then ¬α describes another observable state of affairs, that x is not happening here and now. But then the conjunction $\alpha \wedge \neg\alpha$ also describes an observable state of affairs. That is, if such a state of affairs were ever the case, it could be seen. Clearly, in applying the principle, we do not need to worry about α and ¬α being at different places or at different times. But what of the possibility of occlusion? [. . .] Might it not be the case that our cognitive functioning makes it impossible for us to see inconsistent states of affairs? Maybe our perceptual mechanisms impose a "consistency filter" on what we see, so that we only see one of the conjuncts, which masks out the other? This is certainly an empirical possibility. But there is no empirical evidence that I know of to suppose that there is such a filter (Priest, 2006a, p. 63).

It would be a mistake to let Priest shift the burden of proof this easily. To see why not, think about the first of Priest's standing exceptions to the principle. We cannot

observe the conjunction of observable properties if the properties are such that they only arise at different times. Take an obvious case. We can observe Jane turning thirty, and we can observe Jane turning forty, but we cannot observe her turning both thirty and forty.

Now, with a little creativity we can come up a way in which this combination *can* in some sense be observed. Imagine that Jane is one of the twins in the Twin Paradox. She went up in an interstellar spaceship with her twin sister Janice stayed on earth. Due to relativistic time dilation at the speeds at which her ship traveled, ten more years passed for Jane than for Janice. Jane returns to earth in time to celebrate Janice's thirtieth birthday. What age is Jane turning?

The obvious (and entirely reasonable) objection to this putative counterexample is that Jane is not turning thirty in the same sense that she is turning forty. The latter is her age if we calculate age by adding together the years someone has been alive in each of the reference frames in which they have existed. The former is her age if we calculate age by subtracting one's birth year from the current year. Given this ambiguity, the counterexample does not succeed in capturing what we mean by saying that she cannot turn thirty and forty. But why not? Is it because one or the other of the two senses of *turning a certain age* that create this ambiguity fail to capture what we meant when we said this combination was impossible? Perhaps so, but which one? This is a difficult question to answer, for the simple reason that for every human being in the history of the species before astronaut Jane, the results of the two definitions always coincide.

To see what the problem *is*, consider an even more outlandish case. A time-travel portal opens up in the middle of Edinburgh's Royal Mile. On one side, it is 2018. On the other, it is 1618. Gareth stands with his left foot in 2018 and his right foot in 1618. He looks down, and for a vertiginous moment his field of vision is like a split-screen television broadcast.

There is no ambiguity in this example. Gareth's right foot is in one time in precisely the same sense that his left foot is in 2018. It remains true that the state of affairs of one of his feet being in 1618 can only arise at a different moment in time than the state of affairs of one of his feet being in 2018. Nevertheless, if one can accept the possibility of time-travel portals, no additional obstacle is posed to accepting that Gareth is observing the conjunction by the general idea that it is impossible to observe the conjunction of two states of affairs that only arise at different times. This looks like a legitimate exception to the arising-at-different-times exception to the general principle that if two states of affairs are both observable, then so is their conjunction.

The intuitive impossibility of observing the joint satisfaction of predicates that are such that they arise at different times arises from the intuitive impossibility of such predicates being jointly realized. In treating contradictory states of affairs as a *new* proposed exception to the principle requiring special justification, Priest is getting things backward. The plausibility of the claim that temporally or spatially disparate states of affairs cannot be jointly observed relies on the assumption that *contradictory* states of affairs cannot be observed.

Priest would be better off conceding that this assumption underlies the other exceptions but arguing that we only assume that contradictories cannot be jointly *observed* because we assume that they cannot be jointly *present*. And *that* may well be true. However, even if it is, and even if we assume for the sake of argument that Priest is right to reject the latter assumption, it would not follow that those who take inconsistent states of affairs to be hypothetically observable have the burden of proof. Unless one side or the other can present more direct reasons to regard contradictory conjunctions of independently observable states of affairs as observable or unobservable, the right move here is to reserve judgment.[5]

6.4 Priest's Case for the Consistency of the Empirical World, Part II (The Argument from Perceptual Illusions)

As it happens, Priest does take himself to have such a reason. He claims that there are perceptual illusions that mislead our senses such that we have "perceptions with contradictory contents" (Priest, 2006a, p. 64). Thus, if things ever *really were* as they appear to be when we are in the grip of such illusions, nothing about our cognitive architecture would stop us from taking it in.

One of his examples has to do with an M.C. Escher-like drawing of staircase in which, starting from one corner, we seem to continuously ascend but arrive back in the same place. In a footnote, Priest reports that when he suggested this example in a paper, one of the referees made what seems to me to be exactly the right objection.

> One argued that we do not see that the corner is higher than itself, but we infer that it is. I think that this is a false dichotomy. *Seeing that* is always an interpretive process, and inference may well play a role in a rational reconstruction of how it proceeds. A trained eye might be able to see immediately that a person is suffering from jaundice, and no doubt it is the yellowness of the skin that informs this judgment. They see the jaundice nonetheless. Indeed, in the kind of pattern-matching that neural networks do, it is often hard to distinction between perception and analogical inference (Priest, 2006a, p. 59 n. 6).

[5] Sometimes in the Q&A after a talk, someone will suggest that if the "observable world" contained inconsistent situations, we would *by definition* be able to observe them—that is what "observable" means—but this is just an equivocation. Does the "observable" world consist of all "observable-level" states of affairs, including all conjunctions of such states of affairs, or does it exclude conjunctions of the ones we cannot jointly observe? For example, do blue signs disappear from the "observable world" when the sun goes down? Given the first definition, unobservable contradictory conjunctions of independently observable states of affairs *would* be part of the observable world. Given the second definition, they would be excluded from that world, but so what? The claim that "the observable world is free of contradictions" is philosophically interesting only in so far as it implies that true contradictions do not arise about the objects we normally take to populate that world. Even if approximately half of all library patrons at any given time were both at-and-not-at the library, the consistency of the "observable world" (in the second way of using the phrase) would remain glorious but pointlessly intact.

One can concede that Priest's analysis here is mostly right and still think it establishes only the existence of a messy gray zone in between the concepts, a zone in which it is difficult to know whether to classify a given cognitive event as a direct perception or an inference from a perception. Depending on how much is loaded into the definitions of "perception" and "inference," perhaps this gray zone could even be understood as the *overlap* between the extensions of the two concepts. None of this means that there are not clear cases on (just) one side or the other of the distinction, or that the conclusion that we have "ascended to the same height" in following the stairs from one corner of the drawing to the other is not a fairly clear case of an inference from what we actually perceive. Inference or something inference-like might indeed play a role in some events that we would otherwise be inclined to call "immediate perceptions," but whatever one says about such ambiguous events, this is not one of them. We do not even *think* we actually see a picture of a corner that is higher than itself. The inference wears its inferentialness on its sleeve.

The same reply applies to his second example, which has to do with motion. No doubt relying on the reports of others, Priest claims that "if one consumes enough alcohol, then before one passes out," the room around one will begin to spin.

> If one focuses one's attention on a point in the room, it appears stationary. The rest of one's visual field, however, appears to continue moving. But the internal spatial relations between the fixed part and the moving part do not appear to change. [. . .W]e have a perception of stationary motion (Priest, 2006a, p. 60).

Actually, what is being described is, at best, the *evidential basis of* a rather dubious *inference* to stationary motion. We do not see any part of the room as both being still and in motion. We perceive a still point, we perceive surroundings in motion, we perceive (assuming the correctness of Sartre and Priest's point about Pierre) that the relations between point and surroundings are not changing as they should, and, if (a) we did not understand perfectly well what was actually going on and (b) we were in any state to make complex inferences, we would perhaps infer a contradiction. If this counts as having a "perception with contradictory contents," then such perceptions are so ubiquitous that the excursion into perceptual illusions is unnecessary. *Any* case in which we observe something that gives us a reason to believe α and we observe something else that gives us a reason to believe $\neg\alpha$, would add up to a case of "perception with contradictory contents." (You see the train coming, and you see an electronically displayed schedule telling you that the train has not come yet. By the definition just given, you are having a perception with contradictory contents!) This would be a disaster for Priest's overall position. After all, in each of his cases of perceptual illusions adding up to "perception[s] with contradictory contents," Priest has plausible moves at his disposal with which to block any inference to the presence of *actual* contradictions in the observable world. We have, for example, plenty of evidence that excessive consumption of alcohol interferes with the proper functioning of one's perceptual apparatus. So far, so good. However, such special circumstances can hardly be expected to be present every time that the evidence of our senses is mixed. (Take a case hardly more interesting or unusual than the one about the early train. You are awake and sober. You hear

someone walking behind you, or think you do, and you turn around. No one is there. Were you imagining the sound, or confusing it with something else? Perhaps, but the only reason to think so is the visual evidence that no one was walking behind you. Far from independent reasons to believe that our senses were malfunctioning giving us a good reason to rule out the contradiction, here it is our everyday certainty that someone cannot both be there and not be there at the same time driving our belief that something must have gone wrong on a sensory level.) Putting the point more sharply, it looks like Priest has a choice here between abandoning the claim that examples such as the drunkard's spinning room add up to "perception[s] with contradictory contents" and abandoning the claim that such perceptions only arise under rare circumstances easily explained by the malfunctioning of our senses. If they arise frequently, and under circumstances under which our senses otherwise seem to be functioning properly, why should we believe that contradictory conjunctions of observable states of affairs are non-existent—or even particularly rare?

Priest's final example is the most interesting one, since it is the only one of the three that at least *seems* to involve direct perception of both halves of an impossible situation.

> The third example concerns colour. It is well known that the brain will "fill in" details of the optical field that are not there. For example, to fill in the fovea (blind spot). Now suppose a subject is shown a field, half of which is red and half of which is green, the two halves being separated by a black line. If the line is then removed, the brain fills in colours in the vacated space, and some observers report seeing that the boundary is now red and green (Priest, 2006a, pp. 60–61).

It is certainly true that we normally take it to be impossible for the same thing to be both "fully" red and "fully" green, but what kind of impossibility is this? Priest assumes that it would be a *logical* impossibility. It is certainly true that we normally assume that red objects are non-green and vice versa, and that this assumption, when combined with the claim that the same object is both red and green, generates contradictions about both the redness and non-redness and the greenness and non-greenness of the object in question, but this is just another way of saying that the assumption that nothing is both red and green all over generates a contradiction when combined with the assumption that something *is* red and green all over. Priest says that, because red and green are complementary colors, it is a "conceptual impossibility" for something to be both red and green, but should we go along with this? One plausible analysis of properties such as redness and greenness is that these are dispositional properties. In other words, what it means for an object to be red or green is that it will *appear* red or green under the right circumstances provided that our visual sense is functioning as it should. Various facts about the physical world, including facts about how our eyes and brains work, combine to get the result that something cannot both appear green and appear red under those circumstances, and thus ground the truth of *nothing can be red and green all over*. Thus, if we see something as both red and green, we are seeing something that is *physically* impossible, but to say that our perception has contradictory contents because no red thing can *really* be green or vice versa is like saying that a Mark Chagall

paintings of a flying goat depicts a *logical* impossibility because goats in the real world do not have the power of flight.

Let us assume for the sake of argument, however, that Priest's conceptual premise here is right. What about the empirical premise? Are the observers who report seeing the missing patch as both red and green are being precisely accurate? This is dubious. As aspiring phenomenologists, students of creative writing, and witnesses in criminal trials undergoing cross-examination all learn in their own ways, it can be terribly difficult to convey *exactly* what one is seeing under even the best circumstances. We can only speculate about what precisely may have been experienced by these experimental subjects. However, Priest himself is admirably meticulous in describing his own similar perceptions.

> It may be possible to obtain a similar experience in a different way. It is easy to construct a pair of glasses that have a red filter on one lens and a green filter on the other. (It is possible to obtain glasses of this kind for viewing old-fashioned 3D films.) If I put such glasses on, and allow perceptions to stabilize, everything appears a uniform brownish hue. But for a short space of time, until things stabilize, I have the very strange experience of seeing everything as red and green, though the red and the green seem, somehow, to be at different depths in the visual field (Priest, 2006a, p. 61).

Priest's meticulousness here undermines his intended conclusion. Even if one goes along with the highly dubious claim that the impossibility of redness and greenness coinciding is "conceptual" in some way from which it follows that an illusion of simultaneous redness and greenness is a "perception with contradictory contents," there is nothing conceptually, logically, or otherwise impossible about redness-at-one-level coinciding with greenness-at-another.

6.5 Why LPm, Logical Pluralism, and Local Validity Cannot Save Us

If, as I have been arguing, Priest's case for the consistency of the concrete world fails, what becomes of his probabilistic recapture strategy? His broader claim is that the statistical frequency of true contradictions is very low, and in light of this the epistemic probability of any particular contradiction being true is very low. If he succeeded in establishing that we have specific reasons to think that the observable world is contradiction-free, this would at least give him a domain-specific instance of his low-frequency premise. Since this argument does not succeed, we are left with the conclusion we came to at the end of 6.1—that, given the assumption that dialetheism is true, the only way to get to the conclusion that the statistical frequency of true contradictions is very low is to *start* from the premise that their epistemic probability is very low.

Of course, all of this only establishes that *Priest's* way of making the case for probabilistic recapture fails.[6] This leaves open the possibility that some other version of the probabilistic strategy could be successful. I will try to close that gap in the final section of the chapter, arguing that even if the low statistical frequency could somehow be established, *no* probability-based strategy could do the work that Priest wants it to do. Although it is worth pausing here to ask how much is at stake in this argument. In the absence of Principle R, is the probabilistic strategy really the dialetheist's only remaining alternative?

In his more recent work, Priest has often cast the question of recapturing classical reasoning in terms of his preference for a non-monotonic logic LPm, with LP as its monotonic core. As with any nonmonotic logic, LPm inferences can be undermined by the addition of new premises. For example, the premises $(\alpha \vee \beta)$ and $\neg\alpha$ jointly deliver the conclusion β. . . . unless we add that extra premise α. What it means for LP to be LPm's monotonic core is simply that any inference that is valid in LP goes through in LPm no matter which extra premises are attached.

This might seem to be an alternative way of dialetheically recapturing Disjunctive Syllogism and other lost treasures of classical reasoning. Alternately, some dialetheists (although not Priest) might think that logical pluralism, according to which different logics are appropriate for reasoning about different domains, might do the same job. A third route might involve the idea of a universal logic in which some rules, such as Disjunctive Syllogism, are merely "locally valid" in consistent domains. (Whether this adds up to a kind of logical pluralism is an issue of classification that need not detain us here.) A theorist unwilling to use the word "valid" for any rule that fails to *universally* preserve truth might express much the same idea by just saying that we are rationally entitled to use rules such as Disjunctive Syllogism in our reasoning about consistent domains because the instances in which it fails to preserve truth do not arise in such domains.

At this point, depending on how one individuates some closely related ideas here, one might think that even if both of the recapture strategies Priest presents in *In Contradiction* fail, the dialetheist has between two and four alternative strategies available to her. Although as it turns out, all 2-4 of these "alternative" strategies *rely* on Priest's probabilistic claim. If that claim falters, so do these "alternatives."

This is most obvious in the case of LPm. As Priest cheerfully acknowledges, the inference rules of LPm (outside of the LP core) do not universally preserve truth. If they did, it would not be nonmonotonic!

> *LPm* is a nonmonotonic (aka inductive) logic. And it is precisely the definition of such logics that they may lead us from truth to untruth. The point is as old as Hume ('The sun has risen every day so far. So the sun will rise tomorrow.') and as new as that much over-worked member of the spheniscidae ('Tweety is a bird. So Tweety flies.') If they did not have this property, these logics would be deductive logics, which they are not (Priest, 2012, p. 740).

[6] The reason I have focused on Priest is that, to his credit, he makes *arguments* for his conclusions about these matters. His fellow dialetheists such as JC Beall and Francesco Berto have a bad habit of just helping themselves to the unargued conclusion that this or that domain is contradiction-free.

In other words, the inference rules of the (non-core parts of) LPm have a merely probabilistic force. The reason $(\alpha \vee \beta)$ and $\neg\alpha$ are taken to even nonmonotonically deliver β is precisely the one Priest gives in his argument for the principle about unnecessarily multiplying contradictions.

> Given that the statistical frequency of dialetheias is low, and given a contradiction about which we have no particular information (or no time to take account of the further information there is), two things follow. First, this statistical frequency provided the epistemological probability of the contradiction. The probability that this contradiction is a dialetheia is, therefore, low. Secondly, it is rational to reject the contradiction (Priest, 2006b, p. 116).

That leaves us with the putatively alternate strategies about carving up domains of reasoning. These can be dealt with quickly. If Disjunctive Syllogism is locally valid in some domains—i.e., "universally" truth-preserving when we restrict the domain of quantification to those domains—the statistical frequency of true contradictions in those domains must be very low. (Specifically, there must not be *any* true contradictions in those domains!) Depending on the precise details of a pluralist account of logic, that restriction can perhaps be loosened, but given that we do not want to reason our way to triviality, for classical logic to be an appropriate logic with which to reason about the domains in question, it must be at the very least highly *unlikely* that *ex contradictione quodlibet* will be triggered by classically kosher but paraconsistently disallowed reasoning. In other words, for any of these moves to work, we need at least domain-specific versions of the low frequency claim, and as we have already seen, given dialetheism, the low frequency claim can only be established by *starting* from the premise of low epistemic probability.

6.6 The Failure of Classical Recapture, Part II (The Failure of the Probabilistic Strategy)

Let us take a harder look at the relationship between probability and evidential rationality. We have started the chapter by noting that if a probability just *is* a relationship between a proposition and the evidence for it, the probabilistic strategy is just the Principle R-based strategy in disguise. Given the failure of Principle R, then, any dialetheist who wants to make sense of the probabilistic strategy had better adopt a different theory of probability. But what alternate theory would get the job done?

One option is hinted at in Beall's (quite brief) discussion of probability in *Spandrels of Truth*.

> *Objection: Rejection and the like*.
> You accept some sentences of the form $\alpha \wedge \neg\alpha$ while also accepting the LEM. This clashes with classical probability theory and its role in guiding (or at least modeling) acceptance and rejection. So, either you need a deviant account of acceptance and rejection, or you have to reject classical probability theory.

> *Reply*.
> While nobody accepts and rejects something at the same time, acceptance and rejection are properly modeled non-classically, not in terms of classical probability. In short, since some claims of the form $\alpha \wedge \neg\alpha$ are ttrue, the 'law' according to which $Pr(\alpha) + Pr(\neg\alpha) = 1$ is to be rejected; instead, we need to allow that $Pr(\alpha) + Pr(\neg\alpha)$ may exceed 1 (Beall, 2009, pp. 100–101).

This is (almost) the sum total of what Beall says about probability. He goes on to note that paracomplete theorists must revise the probability calculus in the opposite direction, and he refers readers to the discussion of probability theory in *In Contradiction.* There, too, Priest merely says that $Pr(\alpha) + Pr(\neg\alpha)$ can be greater than 1. In the Field quote we looked at in Sect. 1.4, Field speculated that in a dialetheic version of the probability calculus $Pr(\alpha) + Pr(\neg\alpha)$ could go as high as 2. This seems right, in so far as the standard dialetheic line on contradictions arising from paradoxes such as the Liar is that the truth of such contradictions is "analytic," and we can presumably be certain in such cases of both α and $\neg\alpha$. Regardless, nothing I say in what follows relies on any particular estimate of the maximum value of $Pr(\alpha) + Pr(\neg\alpha)$.

Beall's distinction between the probability calculus "guiding" and "modeling" rational acceptance and rejection is a good starting point for clarifying the options. To say that the probability calculus provides us with rational guidance about acceptance and rejection is presumably to adopt epistemic Bayesianism or something like it and thus to collapse the difference between the probabilistic strategy and Priest's discarded Principle R. On the other hand, if all the probability calculus does is *model* rational acceptance and rejection, does that mean that it is possible for the model to break down in certain instances, such that something can have a low probability without being rejectable, or a high probability without being acceptable?[7] If so, it is hard to see how a probabilistic strategy—*any* probabilistic strategy—could do the work of recapturing the role of classical logic in regulating our *reasoning*. We want to be able to explain why men and dogs are being *rational* when they infer β from $(\alpha \vee \beta)$ and $\neg\alpha$. Neither of Priest's strategies can accomplish *that*.

6.7 Francesco Berto on Negation, Exclusion, and NOT

A third strategy has been proposed by Francesco Berto. As we saw in 5.6, Berto is rightly

skeptical of the ability of his fellow dialetheists to use rejection-talk as a surrogate for the lost

exclusionary power of classical negation. Still, he suggests, *some* surrogate is needed!

[7] Note that my talk of "high probability" and "low probability" here is neutral on the issue of whether, if "high probability" is sufficient for rational acceptance (or "low probability" is sufficient for rational rejection), the precise thresholds for acceptability and rejectability are universal or context-dependent.

> [E]xactly what is accepted by one who accepts a dialetheist's denial? A denial is supposed to convey some information. After all, as R.M. Sainsbury pointed out, "the expression of a denial on the page is just a sentence"; but "the appearance of 'A' on the dialetheic theorist's page, or of 'A is true only,' does not tell the informed reader whether A will [later] be said to be a dialetheia." The *what*-question is meant to stress the fact that [. . .] information is conveyed even by a pragmatic speech act, or by anything, only if it transmits a determinate content; that is to say, if it excludes something [. . .] (Berto, 2006, p. 291)

Priest has a standard response to this worry, which is to say that there is nothing that *anyone* can say about some α to "force" consistency and stop ($\alpha \wedge \neg\alpha$) from holding. Thus, the dialetheist and the defender of classical orthodoxy are in the same boat, and this problem cannot be used to argue for the superiority of one of the two rival positions.

The equivocation at the core of this response can be brought out by considering the case of denying trivialism. As we saw in 5.1, there is at least one sense of "begging the question," in which no one can beg a question against a trivialist. If you say, "trivialism is false," the trivialist will say, "I agree—but it's also true." If a dialetheist says, "trivialism is *only* false," the trivialist will say, "I agree—but it's also true." If the classical logician tries to squeeze some exclusionary power out of *ex contradiction quodlibet*, saying that it is false and that means that *if it is true, then everything is true*. . .then the trivialist will of course agree that it is false and agree with both the accompanying conditional and it is antecedent. They already think that everything is true. That is the trivialist's whole point!

At first glance, this looks like at least one case in which Priest is right that the dialetheist and the classicist are in the same boat. . .and there is a sense in which this is true. Crucially, there is also sense in which it is not. The classical logician cannot deny trivialism in a way that excludes an embrace of trivialism *by the trivialist's lights*. The dialetheist cannot deny trivialism in a way that excludes an embrace of trivialism *by his own lights*. The point generalizes. When a classical logician says that "there are *not* any sets that are both members and non-members of themselves," or for that matter "the moon is *not* made of green cheese," she is saying something that *by her own lights* excludes the possibility that there are such sets or that the moon *is* made of green cheese. Not so by the dialetheist's lights when he makes the same statements. This absolutely does have dialectical force. Moreover, as Berto points out, Priest and other dialetheists certainly talk about all sorts of things *as if* they were excluding certain possibilities.

> Priest too seems to talk as if *some things are impossible* or *some things are incompatible*, that is to say, they *cannot hold together*. Such would be, for instance, a situation in which the same person simultaneously catches and misses a bus. He can talk this way, only in so far as he has (as indeed anyone, obviously, has) a deeply rooted and unassailable intuition of *ontological content exclusion*, or of *material incompatibility* (Berto, 2006, p. 291).

According to Berto, some states of affairs are materially incompatible, and pairs of statements describing each of them therefore have the property of mutually excluding each other's content. The debate between dialetheists and non-dialetheists is thus a debate about whether *every* pair of logically contradictory statements has this property.

But how is the dialetheist to *express* that some pair of statements α and β "can't" both be true, if "$\neg \diamond (\alpha \wedge \beta)$" will not do the trick, and no rejection-based workaround is viable? Berto's suggestion is to introduce a kind of super-negation, NOT.

> To say that something is NOT P_1 is to say that it has some property, P_2, which is materially exclusive with respect to P_1 (Berto, 2007, p. 307).

Berto ends his book *How to Sell a Contradiction: The Logic and Metaphysics of Inconsistency* by recommending that other dialetheists join him in accepting "the (LNC) stated in terms of NOT" and expresses the hope that this will resolve "a hard conflict of intuitions between friends and foes of consistency" (Berto, 2007, p. 316). Perhaps, by analogy to semi-compatibilism about free will, Berto's position could be thought of as "semi-dialetheism." The LNC holds (in the sense that it does NOT have any true counterexamples) in one sense of "not" but not another.

Of course, this also means that the version of *ex contradictione quodlibet* formulated in terms of NOT will hold—or at least that it cannot be objected to on specifically dialetheist grounds (as opposed to, say, relevance-based grounds). The inference from "α" and "NOT α" to any arbitrary β will be universally and vacuously truth preserving according to the Bertonian dialetheist, for the same reason that the inference from "α" and regular "not α" will be universally and vacuously truth preserving according to the classical logician. An immediate worry arises with regard to what Beall would call the "inevitable spandrels" like (29).

(29) Sentence (29) is NOT true.

One might worry that (29) is either true and NOT true, and that the "familiar Liar reasoning" of Sect. 1.1 delivers the result that it is therefore both. . . .at which point Berto's view collapses into triviality. This worry would be misguided for the simple reason that a third option is that (29) could be not true without being NOT true.[8] All statements that are NOT true are not true, but NOT all statements that are not true are NOT true. So far, so good. . .but what this explanation brings into sharp focus is that,

[8] Rubin and Roffé 2018 and Arenhart 2019 argue that sentences like (29) lead to revenge paradoxes for Berto after all. According to Rubin and Roffé, such a sentence only fails to deliver the kind of "absolute contradiction" Berto is so keen to avoid—in other words, it only avoids being true and NOT true)—because Berto does not tell us very much about the rules governing NOT. A plausible rule is that if some sentence S entails an absolute contradiction, P is NOT true, and they show that (29)-style sentence create a real revenge paradox when this rule is assumed. Whether Berto has a good non-*ad hoc* way of resisting the introduction of this rule is unclear, at least to me, since the proposed rule is essentially RAA for NOT. I argued in 4.2 that dialetheists generally have good reasons to reject RAA, but perhaps these do not apply to this case since Berto is a dialetheist about sentences being true and not true but not about sentences being true and NOT true. Arenhart argues that, despite initial appearances, Berto cannot avoid something like an LEM for NOT. Any given sentence is either just true, just false, or glutty. If some sentence S is either just true or glutty, it is true, and "S is NOT true" will be just false. (Parallel points hold for falsity.) Thus, Arenhart thinks, every sentence will be either true or NOT true and sentences like (29) will be paradoxical after all. Whether Arenhart succeeds would seem to hinge on whether Berto is forced to (or indeed able to) make sense of a notion of "just" true or "just" false that is not defined in terms of NOT.

while the LNC formulated in terms of NOT comes out as true (and NOT not true) on Berto's account, the *LEM* formulated in terms of NOT does NOT.

In other words, Berto's proposal amounts to a sophisticated version of the position Beall put forward in his 2005 paper "Transparent Disquotationalism" and abandoned in his later work—according to which natural language contains two concepts of negation, one of which delivers "gappy" results and the other "glutty" results in the face of the "inevitable spandrels." (See 1.5 for a discussion of this iteration of Beall's dialetheism.) With regard to the LEM formulated in terms of "gappy" negation, Berto's attitude will perforce be the same as Hartry Field's attitude to the "normal" LEM—he will reject it "all the way down." For example, he will reject (by saying that it is NOT true) the disjunction, "Either (29) is either true or NOT true or else it is NOT either true or NOT true." If he does not reject this, he is in trouble, at least assuming that the DeMorgan equivalencies hold when formulated in terms of NOT.

One might wonder how this unparsimonious multiplication of negations can be justified on Berto's account. If Berto is comfortable with solving (some versions of) the Liar Paradox in a paracomplete manner, then unless he endorses one or more of the arguments for "extra-semantic" contradictions canvassed in Chaps. 2 and 3, or something like them, why not just hold negation to be unequivocally paracomplete?

In any case, in the next chapter I will argue that such paracomplete solutions to the paradoxes are as problematic as their paraconsistent duals.

Chapter 7
Hartry Field and Paracompleteness



Abstract The paracomplete solution to the semantic paradoxes championed by Hartry Field might seem to be more promising than its paraconsistent competitor. Embracing some logical contradictions is more initially implausible than rejecting both disjuncts of some instances of the LEM. On closer investigation, however, this intuitive disadvantage disappears. The LEM turns out to be as important to our reasoning practices as the LNC. The importance is simply less obvious because implicit appeals to the LEM tend to have in the background of arguments rather than at flashy and obvious points such as the discovery of a contradiction in an opponent's views. Moreover, while incompleteness is normally far less of a strike against a theory than inconsistency, *paracompleteness* is a very different beast than mere incompleteness. In fact, to even make sense of what it means to reject propositions without accepting their negations we have to abandon the probability calculus that is essential to our understanding of subjects ranging from weather prediction to the mysteries of quantum mechanics.

Keywords Paracomplete · Paracompleteness · Field · Probability · Maudlin

7.1 Paracompleteness and Truth Preservation

In classical logic, everything follows from contradictions, and everything entails Excluded Middle instances.

- $\alpha \wedge \neg\alpha \,//\, \beta$
- $\alpha \,//\, \beta \vee \neg\beta$.

Classical logicians take all LEM instances to be true and all contradictions to be false. Since one of the assumptions baked into classical logic (as opposed to, for example, relevance logics) is that *validity is truth preservation*, all inferences *from* contradictions and *to* LEM instances are classically valid.[1]

[1] More precisely, validity, at least in the strict sense captured by the rules of symbolic logic, is *necessary truth preservation by virtue of logical form*. An interesting question outside of the scope

© Springer Nature Switzerland AG 2022

B. A. Burgis, *Logic Without Gaps or Gluts*, Synthese Library 458,
https://doi.org/10.1007/978-3-030-94624-1_7

Dialetheists who want to avoid trivialism necessarily reject the part about inferences *from* contradictions and thus reject classical logic in favor of a "paraconsistent" framework in which some inferences from inconsistent premises are invalid. Similarly, paradox-solvers such as Hartry Field who reject some statements of the form $(\beta \vee \neg\beta)$ but want to avoid inconsistency necessarily reject the part about inferences *to* LEM instances.

The most natural way to express Field's position would be to say that he accepts all instances of the LNC but rejects some instances of the LEM and thus thinks that the logical system that gets all the facts about truth preservation right is paracomplete. Oddly enough, it turns out that he cannot say any of that (except the part about rejecting LEM instances) without throwing away his paracomplete solution to the paradoxes.

Let L be a Liar sentence. Field cannot accept the LNC instance $\neg(L \wedge \neg L)$ without affirming (by way of DeMorgan's) the LEM instance $\neg L \vee \neg\neg L$. Since this instance would be entailed like all the others by accept the LNC as expressed as an unrestricted universal generalization, he cannot do that either.

As to the claim that his own paracomplete rules preserve truth, recall our discussion of Curry's Paradox in 5.1. Like Priest, Field solves the paradox by rejecting Conditionalization and Contraction. So far, so good—or so bad, if you see the loss of Conditionalization as staggeringly counterintuitive.[2] This move suffices to block the inference from Curry sentences' truth conditions to the truth of their absurd consequents, but a different problem remains.

Consider our go-to Curry sentence, (8).

(8) Tr<8> $\rightarrow \perp$.

The reason that rejecting Conditionalization is attractive to non-classical paradox-solvers such as Priest and Field (and *Spandrels*-era Beall)[3] is that it allows them to block triviality without having to sacrifice Modus Ponens. But if Modus Ponens is valid, and all valid inferences preserve truth, it looks like we have to say that if the antecedent of (8) is true, so is the consequent—but to say that is just to assert (8). At this point, if nothing is done, it looks like Conditionalization was sacrificed in vain.

Priest's solution to this second-round version of Curry's Paradox is to use classical conditionals to make sense of valid inferences preserving truth and non-classical conditionals to make sense of Modus Ponens. As we saw in 5.1, dialetheists cannot accept that Modus Ponens is valid for conditionals that obey

---

of my project here is whether it makes sense to talk about a broader sense of validity in which, for example, the single-premise inference from "Person P is a bachelor" to "Person P is male" is "valid" or whether the only reason why it looks like the conclusion deductively follows from the premise is that we are reading the argument as an enthymeme with the definition of bachelorhood hovering discreetly in the background.

[2] It is.

[3] As we saw in 5.1, Beall has more recently moved to a very different solution to Curry.

the classical equivalence between $\alpha \rightarrow \beta$ and $\neg\alpha \vee \beta$, instances of Modus Ponens for such conditionals are (at least given Double Negation) equivalent to instances of Disjunctive Syllogism.

One might wonder whether "conditionals" that do not honor Modus Ponens really capture any intuitive sense of "if. . .then," and in particular whether they get at what we mean when we say that the conclusion of a valid argument must be true if the premises are true. Recall from 5.1 Priest himself, in the first edition of *In Contradiction* (written before Field pointed out the second-round Curry problem) was quite insistent on this point—"any conditional worth its salt" must satisfy Modus Ponens because this is "analytically part of what implication is" (Priest 2006b, p. 83).

Whatever is true about this point, Field goes in a different direction. Since he accepts Disjunctive Syllogism, and in any case it is a point of pride for him that his conditionals act like classical conditionals in contexts where the LEM is assumed, he has no neutered, non-detachable conditionals lying around with which to express that the idea that his logical rules preserve truth. Instead, he gives up on this claim.

> I'm inclined to state my conclusion by saying that the validity of a rule does not require that it generally preserve truth. However, some may think that this simply violates the meaning of the term 'valid': 'valid', they may say, simply means 'necessarily preserves truth', or 'necessarily preserves truth in virtue of logical form', or some such thing. I don't think it does mean this [...] but I don't want to fight about semantics: if one insists on using 'valid' to mean that, then my point is that every serious theory of truth employs rules whose 'validity' (in this sense) it rejects (or else can't express). This seems initially surprising, but becomes less so when one reflects that the rule might still preserve truth 'when it matters' (Field 2009, p. 266).

Field also suggests that the failure of validity to (always) preserve truth "is somewhat less counterintuitive" for paracomplete theorists than it would be for any of their classical brethren who came to the same conclusion "because they don't declare their rules *not* to be truth preserving" (Field 2015, p. 40). To test this last claim, let us consider three sentences.

(30) Hartry Field is a Catholic priest.
(31) Either Hartry Field is a Catholic priest or this sentence is false.
(32) Modus Ponens is universally truth preserving.

Field is not a Catholic priest, which makes (30) false and (31) a "contingent Liar" sentence of the sort which are sometimes brought up to make trouble for various proposed solutions to the paradoxes. If I find (32) obviously true, does Field's rejection of it lose any of its counterintuitive sting because he takes it to have the semantic status he attributes to (31) rather than the semantic status everyone attributes to (30)? It is not clear to me that it does, but whatever one makes of that issue, there is a dilemma for Field here. If rejecting something on the basis that it has the semantic status of a Liar sentence is somehow a "milder" form of rejection than rejecting it on the basis that it is false, this might make Field's rejection of the truth preservation claim less counterintuitive, but it should make his inability to assert the negation of every contradiction (although he can *reject* every contradiction) *more* counterintuitive. Similarly, *reductio ad absurdum* cannot come out as valid in Field's paracomplete framework, or else Liars would come out as *not true*.

He can mitigate the intuitive damage of this by saying that the derivability of a contradiction from a proposition always gives us a reason to *reject* that proposition, but not if rejection-without-attribution-of-falsity is rejection-lite. Either way, the result that "all valid inferences preserve truth" has the same status as "some contradictions are true" makes the package of views less plausible.

Field makes another, at least initially more promising move, which is to claim that he can still make sense of saying that valid inferences preserve truth "when it matters." In his 2015 essay "What is Logical Validity?" Field provides a convenient list of the contexts in which he takes it to matter.

> [O]n my own approach to the paradoxes and some others, the following all hold:
>
> (A) Logically necessary truth-preservation *suffices for* validity in the regulative sense; for example, if $\Rightarrow \mathrm{True}(\langle A \rangle) \rightarrow \mathrm{True}(\langle B \rangle)$ then one's degree of belief in B should be at least that of A.
> (B) The validity of sentences coincides with logically necessary truth: it is only for inferences with at least one premise that the implication from validity to truth preservation fails.
> (C) If an argument is valid, there can be no *clear* case of its failing to preserve truth.
> (D) If an argument is valid, then we should believe that it is truth preserving to at least as high a degree as we believe the conjunction of its premises.
>
> This collection of claims seems to me to get at what's right in truth-preservation definitions of validity, without the counterintuitive consequences (Field 2015, pp. 41–42).

(A) definitely represents *part* of the intuitive appeal of the traditional view that validity is truth preservation. So does (B), although it is worth pausing to note that given Field's commitments, (B) comes with a cost. Field must either say that $\alpha \rightarrow \alpha$ is not a logically necessary truth or that not *all* Curry sentences ($\mathrm{Tr}\langle\alpha\rangle \rightarrow \beta$) should be rejected, given that in a situation where $\alpha = \beta$, ($\mathrm{Tr}\langle\alpha\rangle \rightarrow \beta$) will be equivalent via the intersubstitutivity rules to $\alpha \rightarrow \alpha$. This dilemma could be short-circuited if Field held that the intersubstitutivity rules can sometimes fail to preserve truth, but this would nullify his frequently stated commitment to *unrestricted* "transparent" truth operator.

(C) is just Field's first move discussed above. Saying that a case where we reject a proposition without attributing falsehood to it does not count as a *clear* case of that proposition failing to be true merely sharpens the dilemma he faces on that point. Does the conjunction of all contradictions not *clearly* fail to be true?

While it is true that (D) captures *one* important reason why truth being preserved by valid inferences matters for our reasoning practices, it is certainly not the only one. Take an extremely common epistemic situation—one where we have no idea whether α is true, but we do know that α entails β and γ. Normally, we take this as a logical guarantee that if α is true, so are β and γ. Facts about entailment are part of how we learn what *would* be true in various scenarios, and it is quite often an important part of the way that we move from having no idea whether α is true or false—the situation represented in the standard probability calculus by assigning a probability of 0.5 to both α and $\neg\alpha$—to at least tentatively accepting one or the other of the possibilities. Think of both the sort of "informal" version of *Reductio Ad Absurdum* discussed in 4.2, whereby we reject theories because they entail

implausibilities that are not actually logically impossible, and the way that we decide between theories not just by holding their false predictions against them but by giving them credit for their fecundity in predicting a large number of true things. This process seems to involve knowing that the prediction would be true *if* the theory was true before we actually decide whether it is true. But if we are only obliged to believe that a theory's entailments would be true if the theory is to the extent that we believe the theory itself, in a situation where $P(\alpha) = 0.5$ and $P(\neg\alpha) = 0.5$, $P(Tr\langle\alpha\rangle \rightarrow Tr\langle\beta \wedge \gamma\rangle) = 0.5$ and $P(\neg[Tr\langle\alpha\rangle \rightarrow Tr\langle\beta \wedge \gamma\rangle]) = 0.5$. It is true that Field rejects the standard probability calculus in favor of a paracomplete alternative (see 7.3, below), but the differences between the two calculi do not matter here. In the brave new probabilistic world, *however* one should correctly model the situation of having no idea whether α is true, false, or has whatever status Liars have, we are in the same situation with regard to $Tr\langle\alpha\rangle \rightarrow Tr\langle\beta \wedge \gamma\rangle$). In other words, given (D), whether your logic is classical or paracomplete, if you have no idea whether α is true, you should equally have no idea whether the hypothetical truth of α would necessitate the hypothetical truth of β and γ. That is a serious enough problem that surely (A)–(D) do not capture *everything* that is intuitively important about the idea that valid inferences always preserve truth.

7.2 Paracompleteness and Ordinary Reasoning

As discussed in Chap. 2, the structural parallel between paraconsistent and paracomplete approaches to paradox means that they have similar advantages in enabling uniform solutions to structurally similar paradoxes. Unsurprisingly, it also means that they share a lot of the same disadvantages.

Just as several of the deductive logical laws that play the largest role in ordinary reasoning, such as Disjunctive Syllogism and *Reductio Ad Absurdum*, rely on the LNC (in the sense of "the LNC" discussed in 1.3), lots of ordinary reasoning rests on the assumption that all instances of the LEM hold. Take the "married/unmarried" puzzle Alex Bellos offered the readers of his math column in the online edition of *The Guardian* on Marh 28th, 2016.[4]

> Jack is looking at Anne, but Anne is looking at George. Jack is married, but George is not. Is a married person looking at an unmarried person?
>
> A: Yes
>
> B: No
>
> C: Cannot be determined (Bellos 2016)

Most *Guardian* readers picked Option C. It certainly does not *look* like we have enough information to solve the puzzle. What often does not click with people until they have seen someone walk through the reasoning steps is that, if Anne is married, a married person (Anne) is looking an unmarried person (George), and if Anne is

[4] Bellos says that the puzzle was designed by Hector Levesque.

unmarried, a married person (Jack) is looking at an unmarried person (Anne). Since these are the only possibilities, the right answer is A. In my experience, the universal reaction of a person who initially picked C to hearing this explanation is along the lines of, "Oh right! I can't believe I didn't see that."

But if Field is right that the crucial premise that "Anne is either married or unmarried" is not a *logical* truth, *why* is A the right answer? While it does not of course follow from the foundational paracomplete claim that the LEM sometimes fails that the LEM fails in any given case, it *does* mean that we cannot derive any given instance of it "from thin air" in the way we are entitled to if we hold LEM breakdowns to be logically impossible. To help ourselves to any given instance, we need some general story about why we can safely assume instances *of this type* to hold.

One might claim, *a la* JC Beall's version of dialetheism, that Excluded Middle breakdowns never result from how "the world" is but rather arise because of the specific features of the "spandrels" created by the introduction of the truth operator into our language. One problem with such moves was explored in Sects. 6.2, 6.3, and 6.4. Once we revise our theory of what is logically possible, this should give us a reason to take a long second look at our ordinary beliefs. There are a great many ordinary cases in which we have no particular reason to believe some α and none to believe $\neg\alpha$, but we simply assume that one of them must be the case. If there is no logical guarantee that such assumptions are always accurate, they require some new justification.

In any case, the Beall-like view just considered is not Field's position. He offers a paracomplete solution to the paradoxes of vagueness, such that Excluded Middle fails for borderline applications of vague predicates. Well, "married" is a vague predicate. It has plenty of borderline instances. Imagine, for example, that George and Jack are guests at Anne's wedding. She and her beloved have both said "I do," but no one has gotten around to signing the marriage certificate. Or perhaps Jack is looking at Anne and Anne is looking at George several years down the line from that happy day. Anne's marriage has collapsed. She and her husband (ex-husband?) now live in different countries with different legal requirements for divorce. Since the two paltry sentences of information given to us Bellos' puzzle column do not rule any of this out, is C the right answer after all?

Field says that worries of this sort are "greatly exaggerated."

> There is, I admit, a worry about this. It's evident that a non-classical logic appropriate to the paradoxes is going to have to be a bit complicated, especially as regards the conditional. And maybe that's OK in the case of the paradoxes [...] But for vagueness (it might be said), the matter is different. Given that almost every term is somewhat vague, and given the complexity of the logic appropriate to the paradoxes, wouldn't applying that logic to the case of vagueness make ordinary reasoning difficult?
>
> I think this worry is greatly exaggerated. It might be useful to compare this case to geometric reasoning. We all know that space is not quite Euclidean, and indeed fails to be Euclidean in a quite complicated way; nonetheless we are safe in using Euclidean reasoning except in special contexts, because the error in doing so is so slight. That seems a reasonable methodology in the case of logic too: reason classically, except for those situations where there is reason to think that the errors induced by such reasoning are significant (Field 2008, pp. 105–106).

Two points are in order here. First, many of the situations in which the non-Euclidean geometry of Einsteinian space-time stops looking approximately Euclidean are a lot less mundane than the "married/unmarried" puzzle. As such, they are hard to miss. No one starts daydreaming while driving home from work, accidentally presses down from the gas pedal too hard, and finds themselves approaching the speed of light, but innocent everyday uses of vague predicates can easily involve the accidental application of a vague predicate to a borderline case. Jack is looking at Anne and realizes that he has developed a little crush on her. Anne is looking at her platonic friend George because they are having a conversation. When Anne leaves, Jack asks George, "Is she married?" George says, "No, she's divorced." Is George' statement true or false? If Field is right, it floats in a paracomplete negative zone where all attributions of truth, untruth, falsity, and unfalsity should be rejected. Moreover, *as Field himself has pointed out*, it is similarly possible to stumble unwittingly into asserting a paradoxical sentence. His amusing example is the one that arises if he claims "that what the least intelligent man in the room is now saying isn't true" and (unbeknownst to him) he *is* the least intelligent man in the room (Field 2008, p. 76). This may be unlikely to happen to Hartry Field, but can the rest of us be so confident?

Whatever makes of this point, a second and deeper problem for Field's move is that it amounts to a paracomplete version of Priest's probabilistic strategy for recapturing classical reasoning. Whether Field's version succeeds where Priest's failed is an open question.

7.3 Paracompleteness and Probability

The reason we think it is rational to apply Euclidean spatial reasoning to our four-dimensional Einsteinian reality in ordinary everyday cases where the distortions will be very slight is that it is very unlikely that such slight distortions will lead us seriously astray in any given case. We can model this, of course, using our familiar classical probability calculus—the one we use for purposes ranging from reasoning about the Monty Hall problem to predicting weather patterns. As we saw in 1.4, Field's rejection of classical logic inexorably leads him to reject one of the central axioms of that classical probability calculus, the one that tells us that if probability of rain is 0.78, the probability of not-rain is 0.22. To start to see the problem, notice that a sufficiently light sprinkle may well count as a borderline application of the vague "raining" predicate.

The distinction between "subjective" and "objective" probability might look helpful in limiting the damage here, but Field makes it clear in his article "What is Logical Validity?" that, although he recognizes some sort of distinction between the two, he does not seem much daylight between what will be true of the objective

probability of Object O having Property P at Time T and the subjective probability which a sufficiently well-informed and epistemically rational agent should assign to O having P at T.[5]

> I take it that no one would want to say that a claim such as
>
> The chance of this atom decaying in the next minute is 0.02
>
> is equivalent in meaning to
>
> You ought to believe to degree 0.02 that the atom will decay in the next minute.
>
> The first claim, unlike the second, is not literally about what we should believe. On the other hand, there seems to be no prospect of a reductive account of chance. And it seems clear that an important part of understanding the first claim lies in its conceptual ties to the second: what would we make of someone who claimed to accept the first claim but thought it was rational to disbelieve strongly that the atom will decay in the next minute? (Field 2015, p. 55)

Of course, Field quite likely wishes to say that we should be confident that the classical axioms he rejects do not fail in contexts such as evaluating the probability of an atom decaying in the next minute, and thus the subjective probability we should assign to this happening. Although, how should he express that confidence with a probability assignment in his paracomplete probability calculus? Will there be a straightforward route from expressing that high degree of confidence in "local" versions of these axioms to engaging in something that looks like normal classical reasoning about decaying atoms?

The word "probability" appears a total of four times in *Saving Truth from Paradox*, the word "probabilistic" once. He refers more often to "degrees of belief," expressing doubt that our beliefs come in precise degrees but using the idea as a "useful, albeit somewhat crude" way to model acceptance and rejection, also offering a continuum-valued version of the same idea and expressing the hope that an as yet unformulated "still less crude" model might work (Field 2008, p. 74) Only a few passages scattered around the book go into any kind of detail about what probabilistic axioms would govern rational people's degrees of belief (or continuum-valued partial beliefs, or whatever) and only two of them add anything to the basic point explored in 1.4. One is a straightforward application of that general point to juggling subjective probabilities, and the other is a tentative promissory note on the subject of conditional probability.

[5] He allows that, given inevitable human epistemic limitations, any given person's beliefs *will* be probabilistically incoherent, and that even the claim that an individual's degrees of belief *should* obey all the "standard probabilistic laws" can fail for two reasons. One reason is, of course, that some of the "standard" laws rely on classical assumptions that Field rejects. Even given a suitably paracompletized probability axioms, however, Field says that the claim that any given person's degrees of belief *should* follow these axioms fails "on senses of 'should' that take into account the agent's computational limitations and faulty logical theory" (Field 2015, p. 45). I do not think the issue of computational limitations is going to be relevant to the difference *between* classical and paracomplete probabilistic reasoning, however, and all the problems I am raising here should apply to any probabilistic reasoning engaged in by agents who have read *Saving Truth from Paradox* and thus revised their previously "faulty" logical theories.

> It's easy enough to get a non-standard theory of degrees of belief [. . .] that retains the general law from classical subjective probability theory
>
> $$P(A \vee B) = P(A) + P(B) - P(A \wedge B)$$
>
> When B is $\neg A$, this gives
>
> $$P(A) + P(\neg A) = P(A \vee \neg A) + P(A + \neg A)$$
>
> The paracomplete theorist (who I assumed not to also be a dialetheist) takes P($A \wedge \neg A$) to be 0 but allows P($A \vee \neg A$) to be less than 1, so P(A) + P($\neg A$) must be less than or equal to 1. The paraconsistent dialetheist, if he accepts excluded middle as Priest does, takes P($A \vee \neg A$) to be 1 but allows P($A \wedge \neg A$) to be greater than 0, so P(A) + P($\neg A$) must be greater than or equal to 1. In that case, rejecting A (having degree of belief below $1 - \tau$, where τ is the threshold for acceptance) requires acceptance of $\neg A$, but the converse fails (Field 2008, p. 364).
>
> On the simple model, to accept B conditional on the assumption A is to have one's conditional degree of belief $P(B|A)$ be over the threshold; for conditional rejection it must be under the co-threshold. (One needs to think through how these conditional degrees of belief work when degrees of belief obey the non-standard laws above, but I think that the usual rule P($A|B$) · P(B) = P($A \wedge B$) is still defensible.) (Field 2008, p. 75)[6]

This last point is an important one for the role played by both deductive logic and probability in our reasoning.

> [O]ne of the reasons for the importance of the notion of logical consequence is that if B is a logical consequence of A, then we should conditionally accept B on the assumption A; and the above definition of conditional acceptance provides for this (Field 2008, p. 75).

This is exactly why the subject is far too important for non-classical logicians to do exactly what Field, Priest and Beall all do for their respective paracomplete and paraconsistent revisions of the probability calculus and essentially leave the whole business as an exercise for the reader. Once one thinks through what the new non-classical axioms look like, *would* the usual rule about conditional probability still be defensible? Once the system is fully worked out and we take it for a spin, applying it to such cases as modeling the behavior of vague predicates (where any given application might be a borderline application), would the results really be as close to the classical results as Field's analogy to Euclidean spatial reasoning suggests? Perhaps they would, but before even asking whether the general idea of a probabilistic strategy for recovering those parts of classical reasoning that rely on Excluded Middle will work, we would need to know much more about what a fully worked-out paracomplete probability calculus would look like.

Meanwhile, let us turn from the complexities of juggling *degrees* of belief to the superficially clearer case of what a paracomplete paradox-solver can say about "full" acceptance and "full" rejection.

[6] Some of this is said in the context of Field's discussion of Kripke's KFS, but it seems clear in context that he means the point to carry over to his own paracomplete system. (If he *didn't* mean that, then the outline of paracomplete probability would be even more underdeveloped.)

7.4 Paracompleteness and Rejection

Recall Priest and Beall's difficulties regarding rational rejection. The principle **Accept** counsels that one should believe something if there is "good evidence for its truth." Of course, "good" is vague, but the higher we set the bar for "good evidence," the harder it is to reject the principle. The limit case is what we can call **Accept(A)**.

Accept(A). One ought to accept something if its truth is analytic.

This is, however, one of the cases that gets dialetheists into trouble about acceptance and rejection. Priest, Beall, and all other dialetheists with whom I am familiar regard the truth of sentences like (1) and (2) as analytic. Thus, such sentences should surely be accepted. But where does that leave us with regard to rejection? Priest rejects **Reject(F)** on the grounds that true statements can be false, but holds the line at **Reject(U)**. As we saw in 5.2, this solution is problematic for at least two reasons. One is that the idea that a proposition can be false without being untrue is philosophically dubious even given dialetheic assumptions, and indeed all dialetheists except for Priest reject it. As we saw in 6.1, Priest's "spreading contradictions" objection to that entailment is an odd one, since it is difficult to make sense of the prohibition against "multiplying contradictions beyond necessity" in terms of counting individual contradictions, and Priest already accepts (and cannot help but accept, given a dialetheic solution to the Liar Paradox) contradictions of the *kind* that dialetheists who accept the falsity-to-untruth entailment must take on. The other problem is that, even lacking this entailment, the move from **Reject(F)** to **Reject(U)** is made irrelevant by the Strengthened Liar. Thus, the dialetheist who accepts **Reject (U)** is by her own lights faced with a rational dilemma about whether to accept or reject the very sentences at the heart of the most powerful argument for dialetheism.

Technically, this problem could be avoided if the dialetheist retreated from **Reject (U)** to **Reject(R).**

Reject(R). One ought to reject something if there is good evidence that it should be rejected.

This would, however, be spectacularly unsatisfying. We are justified in accepting propositions because we have good reason to believe that they are *true*. Similarly, the parallel attitude should be justified by some parallel fact about the sentence, such as falsehood or (if there is a difference) untruth. Consider the case of a commission that advises the Queen about which British citizens should be knighted. They might recommend one person for his contributions to theoretical physics, another for her military valor, and so on, but even a person as inexpressive as the Queen would surely raise an eyebrow if they tried to claim that a third person deserved a knighthood for deserving a knighthood. Homer Simpson's "Award for Outstanding Achievement in the Field of Excellence" comes to mind.

Happily, the paracomplete theorist can accept all three of **Reject(U), Reject(F)** and **Accept**. For every α, if Tr<α>, α should be accepted, and if F<α>—or, equivalently, ¬Tr<α>—α should be rejected. Notice, however, that the paracompletist cannot regard **Reject(U)/Reject(F)** as telling us *everything* there is to know about which propositions a rational agent would reject in a situation of complete information. After all, the core of the paracomplete solution to the paradoxes is to reject sentences like (1), and *to also reject their negations*. Given the T-Schema, one cannot (consistently) reject α without also rejecting ¬Tr<α>, and in any case Field is explicitly committed to "rejecting all the way down," such that for every rejected α, its Bivalence and LEM instances are rejected, the Bivalence and LEM instances *about* those Bivalence and LEM instances are rejected, and so on into infinity.

This raises an obvious worry about Field's theory of paradox: What can he say about *why* paradoxical sentences should be rejected? Can he do better (with regard to *those* sentences) than **Reject(R)**?

Note that the same question can and should be asked about Field's paracomplete probability assignments. In fact, since Field defines rejection in terms of degrees of belief (whether precise or continuum-valued) governed by paracomplete probability axioms, for Field this should not even count as a distinct question.

On the *classical* version of this picture, an agent should revise her degree of belief in some proposition α downward to the degree that the evidence suggests that α is *not true*. On the paracomplete version, what should we substitute for the phrase "is not true"? Routing one's theory of rejection through the probability calculus does not make this problem any less pressing.

7.5 Paracompleteness and Revenge

A different way to ask the question is: What is it about these sentences in virtue of which they should be rejected? Let us start with our go-to Liar, sentence (1).

(1) Sentence (1) is not true.

According to the paracomplete approach, (1) should be rejected. Why? Field cannot say that it is because (1) is *untrue* without sacrificing the distinction between paracompleteness and the theory of truth value gaps. (See the discussion of "proto-paracompleteness" in Sect. 1.5.) Instead, he says that it is because (1) is not *determinately* true. A is determinately true (DA) iff $A \wedge \neg(A \rightarrow \neg A)$. (See Field 2008, p. 236) Thus for a statement to *not* be determinately true is for it to satisfy $\neg[A \wedge \neg(A \rightarrow \neg A)]$, which is equivalent through DeMorgan's and Double Negation to $\neg A \vee (A \rightarrow \neg A)$. Falsehoods are not determinately true because they satisfy the first disjunct. (1) is not determinately true because it satisfies the second. (That it satisfies the second is part of what gives rise to the paradox in the first place.)

Now, it sounds like we have a sufficient universal principle for paracomplete rejection, **Reject(D)**.

Reject(D). One ought to reject something if there is good evidence that it is not determinately true.

An obvious objection (for which Field is well-prepared) is that we can use **Reject (D)** to construct a revenge Liar.

(30) Sentence (30) is not determinately true.

Should sentence (30) be rejected? If so, why? If it is because it is not determinately true, then do not the standard intersubstitutivity rules (Capture and Release) quickly generate a contradiction? The reason they do not has to do with the formal features of the conditionals in Field's paracomplete logic. We can say of (30) that it is not determinately true without affirming (30)'s attribution to itself of "ground-level" failure to be determinately true. Iteration of this problem and the solution to it yields an infinite chain of determinateness operators.

As such, it looks like **Reject(D)** should be more precisely expressed as **Reject (DD)**, where the extra D stands for "Disjunctive."

Reject(DD). One ought to reject something if there is good evidence that it is not determinately true on at least one of the infinite number of levels of determinacy.

The obvious problem is that if Field could legitimately refer simultaneously to all the level of the hierarchy at once to express **Reject(DD)**, the revenge-monger could legitimately do so to formulate a sentence like (31).

(31) Sentence (31) is not determinately true on at least one of the infinite levels of determinacy.

Field devotes enormous energy in *Saving Truth from Paradox* to arguing that there is no "top" or "absolute" determinacy operator—that such a final operator will not be generated by his hierarchy, that we should not just postulate one, and that the impulse to do so is misguided—but I cannot help but thinking that all of this misses the point. If I make the odd but (at the time of this writing) true assertion that *the number of Presidents is identical to at least one odd number*, I am simultaneously talking about every odd number in the infinite series of odd numbers, but I am not thereby postulating a "top" odd number.

One of Field's core philosophical commitments is to a disquotational view of truth in which asserting, for example, that the statement "'Snow is white' is true" does not contain any information about the world not conveyed by "Snow is white." Depending on how exactly the metaphor of truth as disquotation—as the removal of quotation marks—is cashed out, when we say "'Snow is white' is true" we are not attributing a "substantive" property to "Snow is white," and perhaps not even a non-substantive one. The communicative function of truth is rather to allow us to express agreement and disagreement without having to spell out exactly what we are agreeing or disagreeing with, and in particular without having to *know* what we are agreeing or disagreeing with. As JC Beall often says—see, for example, the quote we looked at in 5.5—if there is a God, He could have asserted all the "base language" facts without needing a truth operator. We resort to talking about truth and falsity

precisely because we are finite and fallible beings. It would take an (uncountably) infinite amount of time, for example, for an enemy of paracompleteness to assert every instance of the LEM, but only a few seconds to say, "All LEM instances are true." (To say it all, we would have to, for example, asserting of every individual real number that it either is or is not a cucumber.) This is exactly the sort of thing a truth operator that unrestrictedly plays Capture and Release is good for. Denying that it is *possible* to refer to every member in an infinite series with this sort of generalization would be an extremely odd move for a disquotationalist to make.

7.6 Paracompleteness and Hierarchies: Field Vs. Tarski and Maudlin

Let us assume for the sake of argument that there is a good solution to the problem raised in the previous section, and that Field's infinite hierarchy of determinateness operators can yield a revenge-immune solution to the paradoxes without undermining his core commitment to a deflationary view of truth. How satisfying should we find this result?

As a point of comparison, consider two other solutions to the paradoxes—Tarski's solution and the one offered by Tim Maudlin in his interesting book *Truth and Paradox: Solving the Riddles*. Maudlin's view is a sophisticated variant of the family of T-Schema-denying views we looked at in Chap. 1. The Schema logically decomposes into (T-IN) and (T-OUT).

(T-IN) $\alpha \rightarrow \mathrm{Tr}\langle\alpha\rangle$.
(T-OUT) $\mathrm{Tr}\langle\alpha\rangle \rightarrow \alpha$.

Maudlin takes Liars to be gappy (neither true nor false), but he accepts classical logic and (T-OUT), so he denies instances of (T-IN). Tarski, on the other hand, famously "regiments" natural languages such as English into an "object language" in which the truth operator is not used and an infinite hierarchy of "meta-languages", each of whose truth operators can only be used to attribute truth to sentences of languages beneath it in the hierarchy. An object language sentence such as "Cats are animals" can be True_1 in the metalanguage, a metalanguage sentence such as "'Cats are animals' is True_1" can be True_2, and so on. Each truth operator obeys both halves of the unrestricted T-Schema (which, after all, Tarski originally formulated!) "at its level," playing "Capture" with sentences in languages below it in the hierarchy and releasing back into those lower-level languages.

While this solution has been historically popular, and my sense is that many contemporary analytic philosophers with only a passing interest in these issues assume that something like it must be right, a great deal of subsequent work on paradox has been motivated by dissatisfaction with it. Some of the objections to Tarskianism can only be discussed by going into the finer points of his account, but

others are straightforward.[7] To start with, it might seem unparsimonious to solve a problem about an apparently singular entity, Truth, by postulating an infinite number of Truths. Even more troublingly, it is not clear how a Tarskian can explain, assert, or even express their own theory of truth without somehow being able to talk about all the languages and truth operators in the hierarchy at the same time. . . which would seem to defeat the point. If this is possible, what is to block the expressibility of a Liar like (32)?

(32) This sentence is not true in any of the infinite senses of "true."

Tim Maudlin has the opposite problem. Instead of being unable to assert sentences he takes to be true, he is committed to assert sentences he takes to be *untrue*. Since he asserts that Liars are neither true nor false, he is necessarily committed to asserting that our go-to Liar (1) is not true. But, of course, this assertion *just is* (1)! Things get even stranger when we consider (9), the Lying Curry sentence used in 5.1 to make trouble for the view that Curry and the Liar are cleanly separable paradoxes.

(9) If sentence (9) is true, sentence (9) is false.

Given the contracting conditionals of classical logic, this is equivalent to (9)'s (T-OUT) instance! Thus, gap theorists such as Maudlin—who Field simply calls "(T-OUT) theorists" since they retain all instances of (T-OUT) but not (T-IN)—are committed to declaring that certain instances of their own *axioms* are false![8]

As Field points out, all of this creates a strange problem for "(T-OUT) theorists" who wish to express disagreement with one another's views.

> Suppose [. . .that Jones] puts forward quite an elaborate gap theory involving (T-OUT). And suppose that I disagree with this theory overall, but can't quite decide which specific claims of the theory are problematic. It is natural for me to express my disagreement by saying "Not everything in Jones' theory is true." But this doesn't serve its purpose: since Jones himself, as a gap theorist, believes that important parts of his own theory aren't true, I haven't succeeded in expressing disagreement (Field 2008, p. 140).

Maudlin solves this problem by introducing a notion of "permissibility." Sentence like "(1) is not true" and "If (9) is true, then (9)" are untrue, but permissible.

[7] Tarski scholars diverge in all sorts of ways about how to interpret those finer points, and any description of it as short and breezy as this one is likely to cause *some* Tarskian to adjust the glasses on his or her (but usually his) nose and say, "Well, *actually,* what Tarski meant was. . ." I can only express the modest hope that I have at least conveyed the core of "the Received Tarski." For my limited purposes, the question of the relationship between the Received Tarski and the actual philosophical and logical views of Alfred Tarski can be put to one side. Having spent the last several chapters trying to represent the views of Priest, Beall, and Field with hair-splitting accuracy so that I can argue against them without being vulnerable to the suspicion that I have missed something, any Tarski devotee who may be muttering to themselves as they read this about some aspect of Tarski's account that I would not have gotten wrong if only I had attended to a key passage underlined three times in their copy of "The Concept of Truth in Formalized Languages" has my sympathy.

[8] See Field 2008, pp. 138–141.

Crucially, this is not a purely epistemic notion. If I write an obviously true "object language" sentence on a chalkboard, and a very quiet and sneaky person manages to erase it and replace it with "The sentence Ben Burgis is saying right now is false" without alerting me in any way, then my witlessly paradoxical statement "The sentence on the chalkboard is true" will be impermissible on Maudlin's account.

Field finds this move unsatisfying.

> This does indeed handle the problem of agreement and disagreement, but only by inventing another notion to do the job that truth has been standardly thought to do. (Indeed, I suspect that many people will regard their own conception of truth as closer to Maudlin's conception of objective permissibility than to Maudlin's conception of truth.) And obviously, if we allow a notion of objective permissibility in addition to truth then it will raise paradoxes of its own: e.g., sentences that assert their own objective impermissibility. Maudlin of course is well aware of this, and proposes a solution. In this case what he proposes is not a (T-OUT) solution (or rather, a (Permissibility-OUT)) solution; instead, he offers a solution of a more standardly Tarskian kind (Field 2008, p. 141).

What Field seems to mean by accusing Maudlin of Tarskianism is that Maudlin ends up multiplying permissibility operators in a way that is reminiscent of Tarski's hierarchy of truth operators. It is worth noting that Maudlin goes out of his way to deny this analogy. His view is that various rules for permissibility are all desirable (some of which will, as Field suggest, look an awful lot like standard rules governing our ordinary understanding of *truth*), but as the Rolling Stones have taught us *you can't always have what you want*. Given some of the hard choices between different rules that cannot co-exist within the same set of rules without inconsistency, Maudlin thinks that multiple distinct sets of rules will be equally satisfactory. Liar-like sentences constructed with one sort of impermissibility operator might not be impermissible relative to the set of rules defining *that* operator, but can still be impermissible relative to another set of rules. When he introduces his multiplicity of permissibility operators, he stresses the ways in which the resulting picture is *un*-Tarskian.

> It is certainly correct that the language we have been discussing up to now is importantly incomplete: it does not contain "permissibility predicates", and must be expanded to include them if it is to become as expressively rich as English. And it is also correct that expanding the language to include these predicates will cause us some headaches. But [. . .] it is not so obvious that much is gained by characterizing the expanded language as a metalanguage, at least not if that tacitly suggests that there ought to be an even more expansive meta-metalanguage. [. . .W]e have expanded the language to include an infinitude of such predicates, all at one fell swoop. And this collection of predicates does not, in any obvious way, have a structure or form a hierarchy. There are not different "levels" of predicates (Maudlin 2004, pp. 170–171).

Putting aside the issue of whether Field's charge of Tarskianism is a fair one, let us ask a more basic question. Assuming for the sake of argument that Field is right to say that (a) Maudlin can only disagree by introducing a new notion to do the work that falsity normally does, and (b) Maudlin can only solve the paradoxes that arise given this new notion with a move that looks distinctly Tarskian, and inherits many of the problems of Tarskianism, does all this add up to a good reason to reject Maudlin's solution to the paradoxes?

The reader is invited to let that question marinate in the back of his or her mind while we note some relevant features of Field's paracomplete account. First, as we have already seen, his "not determinately true" operator is often supposed to do precisely the disagreement-expressing work played by falsity in the classical picture. (Field cannot say that all contradictions are false, but he can say that no contradiction is determinately true. He cannot say that all propositions that logically entail contradictions are false, but he can say that none of them are determinately true. . .) Second, when paradoxes arise about the "other notion" that he substitutes for falsity, Field's solution precisely mirrors Tarski's infinite hierarchy of truth operators. Each Liar attributing lack of "determinate truth" to itself lacks determinate truth. . .at the next level. Whether or not this is true of Maudlin's permissibility operators, Field's "not determinately true" operators most definitely have a structure and form a hierarchy. Finally, Field's hierarchy inherits both of the problems for Tarski's hierarchy sketched out above. In both cases, an infinite hierarchy unparsimoniously replaces a single apparently unitary operator, and in both cases, it is not clear how the solution can be explained, asserted, or even expressed without at least implicitly talking about all the levels at once. When he says, for example, in relation to the problem examined in 7.1, that it "matters" that valid inferences preserve truth "when the premises are not pathological" (Field 2008, p. 200), how can we understand his use of "pathological" except as "not determinately true at some level or other of determinate truth"?[9]

Perhaps a sympathetic critic could argue here that, even if Field is in the same boat as Maudlin on both points, Field's analysis of just what it is that makes Liars "pathological" is sufficiently compelling and insightful that his paracomplete theory is still the best solution to the paradoxes on the market. Let us turn to that question now. Putting aside all difficulties about revenge and expressive limitations, let us assume that what we called **Reject(DD)** above—"One ought to reject something if there is good evidence that it is not determinately true on at least one of the infinite number of levels of determinacy"—is a reasonably good approximation of what Field is getting at. How satisfying should we find it?

[9]The context of this remark is a discussion of the advantages and disadvantage of glut theories, but it seems clear in context that by "when it matters," he does not just mean "when the glut theorist thinks it matters." Comparing this passage to the more extensive discussions of the same idea in Field 2009 and Field 2015 makes it clear that he is attributing to the glut theorist the advantage of being able to say that valid inferences preserve truth when it *actually* matters.

7.7 Paracompleteness and the Exhaustive Characterization Project

JC Beall usefully divides up the two "projects" at stake in any solution to the semantic paradoxes as the NTP (Non-Triviality Project) and the ECP (Exhaustive Characterization Project).[10] A traditional definition of paradox is that a paradox is an argument from premises that seem too obvious to accept to a conclusion that seems obviously unacceptable, but which nevertheless seems to follow from those premises. In the case of the Liar Paradox, the steps go like this:

(1) There is a proposition L such that $L = \neg L$.

(2) The T-Schema holds universally.

(3) The LEM holds universally.

From (1) and (2), we get:

(4) $L \leftrightarrow \neg L$.

From (1) and (3), we get:

(5) $L \vee \neg L$.

From (4) and (5), we get:

(6) $L \wedge \neg L$.

. . .and from (6) we classically derive triviality:

(7) $\perp$.

The NTP concerns which premise (or bit of reasoning) the paradox-solver rejects to avoid (7). The dialetheist accepts (1)–(6) but resists the move from (6) to (7). (When dialetheism is not one of the solutions under consideration, the Non-Triviality Project just becomes the Consistency Project.) The paracomplete theorist accepts (1) and (2) but rejects (3). The solution I suggest in the next chapter rests on rejecting (1). As with any paradox, the question of what a given paradox-solver discards and what kind of argument they can give for discarding it is an interesting one. After all, whether the paradox is being solved by rejecting at least one premise, arguing that the conclusion does not really follow, or claiming that the conclusion is an acceptable one after all, *some* widely held intuition must be going by the wayside, or else—at least given the definition just given—it would not be a paradox!

Still, the question of how triviality is avoided and what must be forfeited in exchange is not the *only* question that should interest us when evaluating solutions. There is also the question of mapping the post-solution landscape. What is it *about* Liars and other such sentences that gives rise to the problem? If such "defective"

[10] See Beall 2009, p. 66.

sentences have peculiar features by virtue of which the one of the seemingly true premises in the argument for triviality turns out to be false, what exactly are those features? This is the Exhaustive Characterization Project.

To pick an obvious example, the Harman/Horwich solution discussed in 1.1 is unsatisfying not because it fails to avoid (7), or even because their move of rejecting (2) has an obviously higher intuitive price tag than rejecting (3). The problem is that the maps are so uninformative. If Liar sentences do not have the standard T-Schema truth conditions, do they have some other, non-standard truth conditions, or are there no conditions under which such sentences could be true? If we pick the last option, are they false or neither true nor false or what? No such answers seem to be forthcoming.

Beall thinks that Field's paracomplete approach to the paradoxes scores very well indeed on the ECP test. He favorably contrasts it to Saul Kripke's approach, which anticipates Field in many ways.

> What Field provides is a way of consistently classifying all Liar-like sentences (would-be paradoxical sentences) definable in the 'hierarchy of determinately operators'. Unlike Kripke's account, we have it that, for example, D-free Liars are simply not 'determinately true' or 'determinately false'. In turn, D-ful Liars similarly get classified: they're not determinately determinately true or determinately determinately false. In turn, DD-ful Liars similarly get classified. And so on.
>
> Accordingly, inasmuch as ECP calls only for an account of how, if at all, we can ttruly[11] characterize (or classify) the 'semantic status' of any (definable) Liar- like sentences, Field has provided an answer to ECP. It is hard to ask for much more (Beall 2009, pp. 92–93).

Is it, though?[12] To start to get a handle on this, note that Field's phrasing suggests that the defectiveness of Liar sentences is not being analyzed "directly" (e.g., in terms of such sentences having a third truth value). Rather, they are characterized as defective by virtue of lacking the more basic property of "determinate truth." But this is misleading. The "determinate truth" (DA) of some statement A is officially defined as $A \wedge \neg(A \rightarrow \neg A)$. If Field were a dialetheist, this would be an interesting and important subcategory of truth, since by dialetheic lights not all truths are determinate truths. . .but since Field is not a dialetheist, the positive notion of "determinate truth" does no work whatsoever. It is just truth simplicter with a wholly redundant coda—the equivalent of defining determinate dog-hood as the state of

[11] Readers who missed the footnote in 1.3 may be thrown off by Beall's odd habit of adding an extra "t" to words such as "true," "truth," and "truly." In Beall's parlance, the phrase "transparent truth" ("ttruth") sometimes refers to his particular deflationary theory of truth and sometimes more refers more broadly to any notion of truth that obeys "transparency" (i.e., the unrestricted Capture and Release rules discussed in 1.2).

[12] To be fair, Beall goes on raise, "on the other hand," a problem about revenge paradoxes. We will cover that territory below, but the sense in which Beall seems quite satisfied with Field's answer to the ECP is precisely the sense with which I take issue. Even ignoring the thorny questions about Field's infinite hierarchy and whether and in what sense we talk about it "all at once," has Field adequately addressed the semantic status question *even for any individual level* of that hierarchy? The question can be put like this: Ignoring the existence of "D-ful" Liars, has Field provided an adequate characterization even of "D-free" Liars?

being a dog and not being a subatomic particle. Since no dogs are subatomic particles, this adds nothing to the ordinary notion of being a dog. Note, however, that all subatomic particles have the odd property of "not being determinately a dog." More on that point in a moment.

Given that Field's notion of determinate truth adds nothing to his notion of truth, we can safely discard the framing of $\neg$DA as D's *lack* of this feature. If $\neg$DA does any conceptual work, it does it only in its "negative" mode, when it has been DeMorganed into $\neg A \vee (A \rightarrow \neg A)$. *This*, in effect, is Field's analysis of defectiveness—a statement is defective if it is either false or it implies its own negation. The second disjunct is of course half of what makes a paradoxical sentence paradoxical—the other half being that it is *implied* by its own negation.

It is worth briefly spelling out the logical significance of this within a paracomplete framework. In a classical context, all and *only* false sentences imply their own negations, since all conditionals whose antecedents are false come out as true in the truth tables of classical logic. While Field's conditionals are of course non-classical, they act just like classical conditionals $A \rightarrow B$ whenever they are combined with both $(A \vee \neg A)$ and $(B \vee \neg B)$. (Field often presents this as a selling point of his conditionals.) However, since their non-classical features come to the fore when one or the other of these LEM instances is not assumed, it is possible for $A \rightarrow \neg A$ to be true without it following that A is false. This is the case whenever it is *also* true that $\neg A \rightarrow A$. Thus, for Field, $A \rightarrow \neg A$ is logically equivalent to $\neg A \vee (A \leftrightarrow \neg A)$. Strictly speaking, then, the first disjunct of Field's de facto analysis of defectiveness is wholly redundant, since on his logical views, to say of a statement A that $\neg A \vee (A \rightarrow \neg A)$ is to say of it that $\neg A \vee (\neg A \vee (A \leftrightarrow \neg A)$.

Thus, in effect, Field's definition of defectiveness (the feature by virtue of which a statement should be rejected) is the disjunction of being false and being paradoxical. But a disjunctive analysis is no analysis at all! This is like thinking we have said something informative about subatomic particles when we say that they are not "determinately dogs" in the sense that they are not quantum particles and being "not a quantum particle" was part of the definition of determinate doghood given above. The classification of the *sense in which Liars are defective* is simply that they have the property of logical equivalence to their negations (which is the starting point of the paradox). Since *lacking truth* is a well-understood feature of statements by virtue of which they should be rejected, lacking "determinate truth" sounds promising. Although on closer inspection, that Liars lack determinate truth has nothing whatsoever to do with the notion of truth. It would be equally informative to analyze the defectiveness of such sentences by saying that they have the disjunctive feature of either being fried eggs or being logically equivalent to their own negations! The fried eggs are not doing any work.

7.8 Paracompleteness and Properties: Field Solution to Russell's Paradox and Mine

Recall from 2.3 that Field (and Beall) distinguish "Russell's Paradox for sets" from "Russell's Paradox for properties." This brings us to a final point about paracompleteness worth considering before I present an alternative, classical solution to the paradoxes in Chap. 8. The paracomplete solution to the paradoxes is supposed to have the advantage over any *classical* alternative that it can retain the "naïve theory of properties," according to which every property has an extension, without this generating a contradiction in the case of the property of *not instantiating itself*.

What can be said about this from a classical perspective? I certainly do not want to deny that every property has an extension. So, what about the "Russell property" *not instantiating itself*? The structurally obvious move, by analogy to the classical solution to the set-theoretic version of Russell's Paradox, would be to deny that there is such a property, but would not that require us to say that sentences like (33)–(37) were all false?

(33) The property *being red* does not instantiate itself.
(34) The property *being a coffee cup* does not instantiate itself.
(35) The property *being a property* does not instantiate itself.

Intuitively, (33) and (34) are true and (35) is false. This looks like evidence that there is an *instantiating itself* property, which *being a property* has and *being red* and *being a coffee cup* do not.

My view on this is that (33) and (34) *are* true, and (35) *is* false, but that we can explain all of this quite adequately without postulating a property of instantiation (or lack-of-instantiation). Instead, we should think of "does not instantiate itself" (or its equivalent for predicates, "is not true of itself") as a kind of semantic placeholder, inheriting its content from the phrase to which it is applied. Thus, (33)–(36) express the same propositions as (36)–(38). (39) does not express any proposition at all, like an instance of "This is that" said by someone putting words together at random without any intended real or even imaginary referent for either of the indexical expressions.

(36) The property *being red* is not red.
(37) The property *being a coffee cup* is not a coffee cup.
(38) The property *being a property* is not a property.
(39) The property *does not instantiate itself* does not instantiate itself.

Of course, if someone insisted that they meant something by (33) above and beyond what (36) means—that they really meant that *being red* does not have that particular *property*—and the same for the rest, I would shrug and grant that *their* utterances of these sentence-tokens really were all false, but I would still insist that other people's utterances of them be interpreted in a more charitable way. That goes double for use of the more common phrase "true of," which is at the heart of what's

usually called "Grelling's Paradox" but could be called, using the Field/Beall terminology, "Russell's Paradox for predicates." (40) expresses the same (true) proposition as (41), (42) expresses the same (true) proposition as (43), (44) expresses the same (false) proposition as (45), and (46) expresses no proposition at all.

(40) The predicate "is red" is not true of itself.
(41) The predicate "is red" is not red.
(42) The predicate "is a coffee cup" is not true of itself.
(43) The predicate "is a coffee cup" is not a coffee cup.
(44) The predicate "is a predicate" is not true of itself.
(45) The predicate "is a predicate" is not a predicate.
(46) The predicate "is not true of itself" is not true of itself.

Field finds the conclusion that sentences like (46) could in any straightforward way be "meaningless" deeply implausible. He thinks they are so obviously meaningful that anyone who claims otherwise must mean the term "meaningful" "in a special technical sense" (Field 2008, p. 12). In terms of the specific issue of whether the *predicate* of the sentence is meaningful, he anticipates the general outline of the move I make here, but argues that such a maneuver would be both unmotivated and unhelpful.

> The most obvious possibility is that a 'meaningful predicate' in the technical sense is a predicate that expresses a property. If we adopt this line we must rely on a theory of properties that avoids the Russell paradox by the Non-Existence Solution to that. (If only we had an attractive theory of properties of this sort!)
>
> Moreover, the claim in (I) that (TO) holds for "meaningful" predicates (or better, predicates that express properties) doesn't settle the application of 'true of' to "meaningless" predicates (predicates that don't express properties) (Field 2008, p. 13).

After spelling out various possible stipulations about whether meaningless-in-a-special-technical-sense predicates do or do not apply to anything, he allows that someone like me "might well say that the question of whether "'meaningless' predicates are true of things is itself 'meaningless', at least in some cases: in particular, it is 'meaningless' to ask whether the 'meaningless' predicate 'is not true of itself' is true of itself." This does indeed strike me as the correct position. Field's objection is that this amounts not a non-existence solution that (like the parallel solution to "Russell's paradox for sets") steers clear of both gaps and gluts, but to "a kind of paracomplete view, one which withholds assent to excluded middle for (some or all) sentences containing predicates that are "'meaningless' in the technical sense" (Field 2008, p. 13).

Let us take these charges one at a time. First, to say that (46) is meaningless in the sense that it does not express a proposition, rather than, e.g., in the sense that the words being used to form the sentence are themselves meaningless, is admittedly to make a somewhat technical distinction. (If I started asking random people on the train to New Brunswick tomorrow morning what the difference was between a sentence and a proposition, most would be stumped, and some of them, being residents of New Jersey, would use language that cannot be reprinted in an academic tome.) It is, however, about as familiar a distinction as there is in the philosophy of language.

Second, it seems to me that to deny that the phrases such as "is true of itself" and "instantiates itself" express distinct properties (as opposed to functioning as semantic placeholders, inheriting their content from the phrases to which they are applied, and being semantically orphaned when one attempts to apply it to itself) one does not need a global theory of properties, any more than to deny that the truth operator expresses a substantive property like correspondence or coherence rather than acting as a device for generalization, one need have in one's back pocket a complete theory of how all substantive properties work. The "semantic inheritance" and "device of generalization" explanations are simply more parsimonious and attractive on their merits.

Finally, I do not grant that saying that "the question of whether 'meaningless' predicates are true of things is itself 'meaningless', at least in some cases: in particular, it is 'meaningless' to ask whether the 'meaningless' predicate 'is not true of itself' is true of itself" (i.e., that neither "yes it is" nor "no it doesn't" expresses a proposition) amounts to "a kind of paracomplete view, one which withholds assent to excluded middle for (some or all) sentences containing predicates that are 'meaningless' in the technical sense." Briefly, if there is no proposition there, there is nothing to which we *could* assent or withhold our assent. A more expansive answer to the "disguised paracompleteness" objection will have to wait until the next chapter. As astute readers will have doubtless already guessed, my preferred solution to the paradox about "is not true of" doubles as a preview of coming attractions for my preferred solution to the more familiar paradox about "is not true."

Chapter 8
How to Solve the Liar Paradox Without Sacrificing Classical Logic



Abstract After seven chapters of seeing how important proposed solutions to the paradoxes can go wrong, several important things learned along the way are pulled together to propose a positive alternative (rooted in a non-standard way of parsing disquotationalism about truth called Content Inheritance Disquotationalism) that retains the full resources of both classical logic and a fully transparent truth predicate. A wide range of possible objections are considered and resolved. The CID-based solution is compared to its paracomplete and paraconsistent rivals on a crucial standard emphasized by Graham Priest—the Principle of Uniform Solution. Finally, several loose ends are tied up through a survey of objections to classical logic rooted in quantum weirdness, borderline applications of vague terms, and Edwin Mares's claim that predicates can be not only "underdefined" (leading to truth-value gaps) but "overdefined" (leading to truth-value gluts). Some of these, particularly concerns about vagueness, are difficult philosophical problems not resolved by anything said here, but none add up to a good reason to abandon classical logic.

Keywords Disquotationalism · Liars · Truth · Inferentialism · Vagueness · Quantum

8.1 Some Preliminary Points About Propositions

In what follows, I will argue that sentences like (1) and (2) fail to express propositions, and that as such that it is a category mistake to symbolize such a sentence with a Greek letter and try to perform logical operations on it. The central premise of this argument will be a version of disquotationalism about truth, which I will call CID for Content Inheritance Disquotationalism. Although before getting to that, it is worth highlighting two of the most controversial background assumptions without which my solution does not get off the ground.

(i) The Greek letters of what is usually called propositional logic but sometimes called sentential logic properly symbolize *propositions* rather than sentences.
(ii) It is possible for syntactically "well-formed" declarative sentences to fail to express propositions.

© Springer Nature Switzerland AG 2022

B. A. Burgis, *Logic Without Gaps or Gluts*, Synthese Library 458,
https://doi.org/10.1007/978-3-030-94624-1_8

Before going further, it may seem necessary to clarify my use of the term "propositions." Any survivor of a graduate-level seminar in the philosophy of language should be able to rattle off several distinct theories of positions. Choosing one would limit the appeal of the proposed solution of the paradoxes to partisans of that theory. Happily, I do not need to do that. Any respectable theory of propositions should validate at least one core claim about these entities and their relationship to sentences, **Prop**.

Prop. Any given sentence expresses a (atomic) proposition iff it attributes a particular property (or properties) to a particular object (or objects).

This can be clarified by supplementing it with **CompoundProp**, which simply restates what students learn from any introductory textbook on propositional logic:

CompoundProp. All propositions are either basic propositions or propositions that have been *built out of* basic propositions (or compound propositions themselves built out of atomic propositions, or compound propositions built out of other compound propositions built out of atomic propositions or. . . .etc.)—for example, by negating them or combining them in conjunctions, disjunctions, or conditionals.

Someone might worry here about sentences that apparently refer to non-existent objects, but **Prop** fares perfectly well on either a Russellean or a Meinongian approach to such sentences. On the Meinongian view, "the Present King of France is bald" refers to a non-existent object (the present King of France) and attributes a property (baldness) to him. On the Russellean view, the sentence refers to every object in existence and says of them that at least one of them has the properties of (a) baldness and (b) being the present King of France. Nor does it matter whether the property is positive ("is bald") or negative ("is not bald"), or whether it is monadic ("is bald"/"is not bald") or relational ("is older than the future King of France").

I do not claim that **Prop** constitutes any sort of informative analysis of propositions. On the contrary, I take it as a virtue of **Prop** that it is extremely uninformative—it does not do much to narrow down the list of acceptable analyses of "proposition." Of course, some philosophers love desert landscapes enough to banish properties from their ontology, but it is hard to imagine a landscape simultaneously austere enough to exclude properties and lush enough to allow propositions. As such, any theorist of propositions who rejected either the left-to-right or the right-to-left half of the **Prop** biconditional would be put in a philosophically awkward position. Starting with the right-to-left component, if a sentence can attribute a particular property to a particular object without expressing a proposition, then what work are propositions doing to earn their place in our ontology. Similarly, while philosophers of language from the verificationists straight up to David Kaplan and his successors have acknowledged a sub-propositional level of meaning, a low-level sense of "means" such that someone can know what a sentence "means" without knowing exactly which proposition it expresses, the latter piece of knowledge is supposed to be the maximum possible amount of knowledge about a sentence's meaning. If it is possible to lack such crucial pieces of information about a sentence's meaning as exactly who or what the sentence is about and exactly what is being said about that object (or objects), then again, what are propositions *for*?

With all of this in mind, claims (i) and (ii) both strike me as obviously true. Even so, since there is no consensus on either point, arguments are necessary.

While there are both interrogative and imperative "logics," these clearly are not logics in quite the same sense as the competing systems of logic discussed in this book.[1] Even theorists who take sentences rather than propositions to be the primary "bearers of truth" do not typically claim that interrogative or imperative sentences can be "true" (or "false"). Thus, the issue is whether the Greek letters properly symbolize propositions or *declarative* sentences.

An obvious reason to favor the proposition view is that declarative sentences are often ambiguous. One token of "Lucy went to the bank" might signify that Lucy went to her local Bank of America branch to cash a check, while another token of the same sentence type might mean that she went to the riverbank to lie in the sun and read poetry. Since she cannot "go to the bank" in both senses at the same time, symbolizing both claims about her activities with the same letter would open the door to strange and rather silly version of dialetheism.

One way around *this* version of the ambiguity objection to the sentential view would be to take the Greek letters to symbolize sentence tokens rather than sentence types, but as Laurence Goldstein points out, this move comes with very similar problems.

> Token sentences are concrete marks or sounds; they are of the same type if they are typographically similar (enough). But consider this: I am preparing sandwiches for tea after the village cricket match, and have thoughtfully attached a label with the inscription 'these are cheese sandwiches' to a toothpick, thereby making a little flag. If I stick the flag into a pile of cheese sandwiches, then I (via the label) say something true. Were I mischievously to stick *the same token* into a pile of egg sandwiches, then, courtesy of the label, I say something false (Goldstein, 2001, p. 116).

This can technically be reconciled with most standard assumptions about truth and falsity by embracing tense logic, so that we can symbolize the label as α and use a pastness operator P to say without contradiction once α has been stuck on the pile of egg sandwiches that $\neg Tr<\alpha> \wedge P(Tr<\alpha>)$, but there are at least two serious problems with this move. The first is that the claim that there are irreducibly tensed utterances is deeply controversial on other grounds. The second is that similar problems can be generated where the ambiguities are not *temporal*.

After all, sentence tokens containing words such as "these" do not always come considerately toothpicked to their referents. Transpose the sandwich-talk from the English village to Richard Nixon's White House—where everything is taped. Decades after the conversation, in a bit of recently released audio, we hear Ehrlichman tell Haldeman, "These are the cheese sandwiches."

Is Ehrlichman accurately referring to some cheese sandwiches as cheese sandwiches, inaccurately referring to some egg sandwiches as cheese sandwiches, or are

[1] This is not a knock on the imperative or interrogative calculi or those who work on them! The set of subjects that fall within the purview of this book is, if I have succeeded in doing something of interest, still just a proper subset of the set of interesting and valuable subjects.

the conspirators using "cheese sandwiches" as a code for files detailing a series of dirty tricks the Committee for the Re-Election of the President was planning against Democratic nominee George McGovern? These three possible interpretations all have very different logical implications. Is it not more natural to think that logic concerns the relations among the claims the sentences are being used to make than among the sentences themselves?

Let us take a closer look at the distinction. Imagine that we feel confident that this particular conversation was really about sandwiches and not election-related conspiracies. At this point, it would seem that we have all the "internal" linguistic information we could ask for regarding the meaning of the words *as they are being used in this particular sentence token* and the way in which the rules of English string together those word-meanings into sentence-level meaning. Nevertheless, there is something we do not know about what "She went to the bank" means. This is captured by David Kaplan's distinction, to which we have already alluded, between "linguistic character" and "propositional content." In Kaplan's terminology, the character of a sentence is a function from context to content. The information we lack about the meaning of "These are cheese sandwiches" is not about what the word "these" means. It is about how the context of the utterance fixes its reference and thus tells us what the claim is *about*.

Of course, a skeptic about propositions could (and should!) acknowledge that the semantic phenomenon Kaplan has his finger on is real. She could nevertheless argue that it is possible to regard sentences as the sole "bearers of truth" and the proper referents of the Greek letters of what she prefers to call "sentential logic." Perhaps she could make sense of all this by saying that the Greek letters only symbolize *precisified* sentences.

Arguing for or against this approach would take me well outside the scope of the present project. For now, I will simply note that, relative to this approach, **Prec** has pretty much all the same things going for it that **Prop** has relative to the theory of propositions.

Prec. An (atomic) sentence has been precisified iff it attributes a particular property (or properties) to a particular object (or objects).

This can be supplemented with a **CompoundPrec** in precisely the same way that **Prop** was supplemented by **CompoundProp**. Can precisified sentences do the work of propositions? Is a precisified sentence simply a sentence that expresses a particular proposition, or is the notion of precisified sentences a useful and meaningful way of cutting down on our ontological commitments? At least for present purposes, my own position on both issues is agnostic. I assert merely that for the theory of precisified sentences to be a viable competitor to the theory of propositions, **Prec** must be true. Meanwhile, I will continue to defer to tradition (and convenience) by talking about whether this or that sentence "expresses a proposition" rather than whether it "is a precisified sentence," but proposition-resistant readers should feel free to mentally substitute (i)* and (ii)* for (i) and (ii) and **Prec** for **Prob**.

(i) * The Greek letters of what is usually called propositional logic but sometimes called sentential logic properly symbolize only *precisified* sentences.

(ii) * It is possible for a sentence to be syntactically "well-formed" without being precisified.

With that in mind, we can turn to (ii)—or if you prefer (ii)*. With Kaplan's distinction in our back pocket, it is easy to generate syntactically well-formed sentences with well-defined linguistic characters but no propositional content. After all, a "These are cheese sandwiches" sentence-token can come into the world in all sorts of ways without the world being kind enough to supply a context that, toothpick-like, fixes the expression "these" to a specific object (or objects).[2] A cat could walk over a computer keyboard and accidentally spell out the sentence. Someone could spell it out in refrigerator magnets with no particular intention except for finding words to make with the available letters and turning those words into something "well-formed." Or we could encounter it in a book such as this one, as in (47), below, in which the author simply stipulates that the "These" is not meant to refer to anything. That last case, of course, parallels that of a standardly presented Liar sentence where it is simply stipulated that the phrase "This sentence" is self-referential.

(47) These are the egg sandwiches.

Is (47) false? Is it neither true nor false? Does it hover in some sort paracomplete negative zone where all such judgments should be rejected? We will return to that issue in 8.5. For the moment, let us just say that, since neither express a proposition, whatever we say about (1)'s relationship to concepts such as truth and falsity should be what we say about actual egg sandwich.

8.2 More Preliminaries: Disquotationalism and Inferentialism

Deflationists deny that truth is anything like correspondence or coherence. This is not to deny there is some set of propositions with which all true propositions will cohere. (On any non-dialetheic view of truth, this will turn out to be trivially true so long as whatever propositions get to go in the privileged set are themselves true.

[2] It could be objected that such sentences have "some" content rather than none. Even a token of "I am here" formed in one of the context-free ways sketched out below comes with a sort of ghostly trace of propositional content, since any instance of "I am here" that succeeds in expressing a proposition will entail the further (not terribly informative) proposition that some person is in some location. However, given **Prob**, if "having propositional content" is a way of saying "expresses a proposition," it is binary. The linguistic character of a sentence may be such that the *property* any proposition the sentence is used to express is "already filled in," as in the egg sandwich case. Alternately, the particular object being referred to can "already be filled in," as in "Nixon is here." However, these conditions—specifying a particular property to be attributed to whatever object the proposition is about, specifying the particular object—are individually necessary but only *jointly* sufficient for expressing a proposition.

Even dialetheists, who believe that not all truths are *consistent*, believe that there is *some* sense in which all truths cohere with each other.) Nor do deflationists necessarily have to deny (nor, bracketing any number of thorny metaphysical issues, do many even necessarily *wish* to deny) that true propositions can be said in some sense to "correspond to" facts or states of affairs "in the world." Rather, what makes a deflationist a deflationist is that she denies that the attribution of correspondence, coherence, or any other such "substantive" property to true propositions is any part of what a competent speaker of a natural language ordinarily *means* when she attributes truth to a proposition.

This is a very broad family of positions. (To start to get a sense of a few possible lines of division, note that to describe the consensus between different deflationists, some of whom deny that truth is a property and some of whom are comfortable with it being a property as long it is the right kind of property, we have had to resort to the suggestive but hopelessly vague term "substantive.") The specific branch of the family with which I identify is *disquotationalism*. Although this term is only slightly more informative, given the wide variety of deflationary views that fall within it, I will use it from this point forward unless I am quoting or referring to the views of others who prefer "deflationism." One reason for this choice is that some but not all deflationist views about truth are motivated by more generally "deflationary" metaphysical and/or linguistic views. This is not at all true in my case, and in my experience describing a view of truth as deflationist invites this misinterpretation. My deflationism about *truth* is motivated by the intuition that "is true" is very *different* from normal predicates and operators that pick out properties in the normal way, so tying the view about truth to a more general sort of "deflationism" would erase precisely the distinction that motivates that view in the first place.

A disquotationalist is a deflationist who combines the negative thesis about the meaning of "true" sketched above with the positive thesis that the meaning of the term *is* given in some sense by the relationship between quoted and unquoted sentences given by the T-Schema. When the phrase "is true" is appended to a quoted sentence, the effect (or perhaps the most important effect, or "the point") is "disquotation"—i.e., the removal of the quotation marks.

One of the most popular ways to cash this out this metaphor is by tying it to inferentialism. To be an inferentialist about Domain D is to think that the meaning of the terms in D stem from their inferential rule. The form of inferentialism that is easiest to understand is inferentialism about the meaning of the logical constants. There are inferentialists even about *moral* terms, and some philosophers are global inferentialists, but inferentialism's "home turf" is the logical constants. Many logicians seem to have picked up a certain amount of inferentialism about this domain almost like philosophical second-hand smoke—hence the common sight of a section of a paper devoted to "giving the language" or "defining" a new constant through the enumeration of inferential rules—and even many philosophers who would be skeptical about, e.g., moral inferentialism find the view plausible on that home turf.

According to inferentialism about logical constants, the meaning of, for example, conjunction, is given by certain logical rules involving conjunction. These almost certainly include (and perhaps are limited to) the introduction and elimination rules.

Conjunction Introduction: α / β // $\alpha \wedge \beta$
Conjunction Elimination: $\alpha \wedge \beta$ // α

The disquotationalist inferentialist simply treats "true" (and hence its negation) in the same way, holding its meaning to be entirely constituted by the "introduction and elimination" rules **Capture** and **Release**.

Capture: α // Tr<α>
Release: Tr<α> // α

If inferentialism is a generally reasonable view, this way of cashing out the "disquotation" metaphor may be appealing. It seems to get us most of what we want out of such a view about truth—a way of validating every instance of the T-Schema and making sense of the communicative role of truth attributions without having to say that every time we make such an attribution, we are claiming that true sentences have some complex, philosophically controversial property like correspondence or coherence. If, on the other hand, we have good reasons to reject inferentialism even about logical terms such as conjunction, we should be on the lookout for a different way of cashing out the metaphor.

The inferentialist analysis of terms like conjunction according to which canonical inferences involving such terms are "meaning-constituting" reverses the older and simpler view about the relationship between meaning, truth, and validity, according to which the meaning of a logical constant like "and" contributes no more and in no different a way to determining whether the sentences it appears in are true or false than a term like "red," and a logical rule such as Conjunction Elimination is valid *because* it never takes us from truth to falsehood. To see how the two views come apart, consider the case of disjunction. The introduction rule is, of course. . .

Disjunction Introduction: α // $\alpha \vee \beta$

. . .but the classically valid inference that it would be most natural to call an elimination rule, letting us go from a disjunction to one of the disjuncts, just as Conjunction Elimination takes us from a conjunction to one of the conjuncts, is our good friend **Disjunctive Syllogism**!

Disjunctive Syllogism: $\alpha \vee \beta$ / $\neg\alpha$ // β

One of the most basic assumptions underlying my project here is that (contra Quine) it is entirely possible for philosophers with differing logical views to debate the merits of those views without talking past one other. Anyone who finds that assumption attractive thus has at least one reason to be skeptical of inferentialism even about inferentialism's "home turf"—the logical constants. This may not be an indefeasible reason, since there are sophisticated ways of evading what Quine called "the deviant logician's predicament"—"when he tries to deny the doctrine, he only changes the subject" (Quine, 1970, p. 81)—without denying the underlying assumption that the inferential rules define the meaning of the logical constants. Still, it is *a* reason, since undeniably the simplest and most intuitive way of resisting this "predicament" is to simply deny inferentialism.

Another reason to deny inferentialism even about logical constants is the classic one given by A.N. Prior in 1960. If, rather than inferences being valid *because* they preserve truth, with "the world" and the meanings of the various logical and non-logical terms (whether those meanings are primitive or analyzable into more basic concepts) conspiring in the usual way to make the premises or conclusions themselves true, the reason that the conclusions of logical inferences are true (given that the premises are) is that we define the logical terms into existence by simply stipulating which rules will be valid, what is wrong with the operator "tonk," which combines the introduction rule for disjunction with the elimination rule for conjunction?

> '2 and 2 are five,' for instance, [can be derived] from '2 and 2 are four.' It is done in two steps, thus:
> 2 and 2 are four.
> Thus, 2 and 2 are four tonk 2 and 2 are five.
> Therefore, 2 and 2 are five (Prior, 1960, p. 39).

The dominant response of inferentialists to this problem is to introduce "logical harmony" constraints to the ability of speakers to legislate logical terms into existence by stipulating which rules governing it will be valid. As Roy Cook puts it, such accounts attempt to hold onto "at least the spirit of the inferentialist view attacked by Prior" without "allow[ing] tonk legitimacy." Several decades after Prior's article came out, the "project of demarcating the bad operators (or their rules) from the good ones" is ongoing (Cook, 2005, p. 217). Still, however complicated and murky the details may be, at least some of the general intuitions underlying this project are clear enough. Inferentialists who accept classical orthodoxy take connectives whose governing rules would generate inconsistency (and hence triviality) to be unacceptably unharmonious. Dialetheist inferentialists such as JC Beall accept inconsistency (which, as both a dialetheist and a disquotationalist inferentialist, Beall takes truth to generate, given the combination of propositions that attribute untruth to themselves and truth's meaning-constituting rules **Capture** and **Release**), but draw the line at constants whose rules would generate triviality directly—such as, perhaps, contracting conditionals.[3]

An alternate strategy for dealing with the tonk problem, endorsed by Cook, seems to rely on a sort of meta-linguistic pluralism according to which both the expressive power of our "language" of logical terms and what counts as acceptable deductive reasoning are context-relative. Cook points out that by abandoning the transitivity of logical consequence, we can have tonk without triviality—or more precisely that in the "context" of non-transitive *Tonk-logic*, tonk can be part of the logical language without triviality ensuing.

> The existence of *Tonk-logic* suggests that discussions of the legitimacy of logical connectives should focus not only on which ones are acceptable *simpliciter*, but should also take

[3] Beall takes this line on such conditionals in *Spandrels of Truth*, but as we saw in 5.1, he was since moved to a very different solution to Curry's Paradox.

> into account which connectives are acceptable in which contexts (i.e. particular formal approaches to logical consequence) (Cook, 2005, pp. 223–224).

Sadly, this move faces a revenge problem. In his paper "Connectives Stronger Than Tonk," Heinrich Wansing both generates innovative formal rules with which to "define" a super-tonk connective that would generate triviality *even without* the transitivity sacrificed in the move from classical logic to *Tonk-logic*, and generalizes the conclusion, providing a way to generate ever stronger tonk-like connectives to generate triviality problems within ever weaker logical frameworks. "Whereas *Tonk-consequence* is a cure for tonkitus," Wansing concludes, "no non-trivial notion of consequence is a cure for super-tonkitus" (Wansing, 2006, p. 659).

If Wansing is right, the inferentialist is left with the dominant "logical harmony" strategy. A full evaluation of what different formal attempts to implement that strategy have going for them, how these attempts have been criticized, what the prospects are for responding to these criticisms, and how we should think about the underlying philosophical motivation would involve at least as many chapters as I devoted earlier to the question of how dialetheists attempt to recapture the use of inference rules such as Disjunctive Syllogism. I hope someone as unsympathetic to the "logical harmony" move as I am writes a serious long-form critique of the idea soon, but this is not that book. For the moment, I will simply point out that such a move would be a very odd fit with the sort of "deviant" logical moves Priest, Beall, and Field make with regard to the paradoxes. If one accepts that (a) terms can simply be exiled from our language for being logically disharmonious and that (b) there are propositions that attribute untruth to themselves, why (c) take one's commitment to a single "transparent" truth operator as a suicide pact? Instead of sacrificing all sorts of logical rules that play an important role in ordinary reasoning (think of Disjunctive Syllogism, or the use of the LEM in simple bits of reasoning such as the married-unmarried puzzle), why not just reject the notion of a unitary "transparent" truth operator as unharmonious? Why not, for example, follow Tim Maudlin in restricting T-IN in the name of logical harmony?[4]

I propose, instead, that we should take the tonk problem and the Liar Paradox equally seriously. The former shows that logical rules do not "constitute" meaning in the way that the inferentialist thinks they do, although they often illuminate it, and the latter shows that *something* must have gone wrong in the way we think about truth. However, in puzzling out *what* has gone wrong, we are not simply free to stipulate our way out of the problem by deciding which rules about truth and falsity will be valid. Some rules simply *are* valid. There is semantic evidence to be considered. English and other natural languages clearly contain a unitary truth operator that can be and frequently is used to form meaningful sentences. Any plausible understanding of how this works must validate all instances of T-IN and T-OUT and—equivalently, since Modus Ponens is valid—make sense of unrestricted Capture and Release rules. If possible, such an account should not load much of anything into our understanding of truth *beyond* the bare-bones

[4] See 7.6 for a detailed discussion of Maudlin's view.

semantic and logical facts just mentioned, so as to accommodate the disquotationalist intuition that the communicative role of the word "true" is simply to allow limited and fallible beings like us to make generalizations that would otherwise lie beyond our expressive power.

8.3 How To Cash Out the Disquotation Analogy and Solve the Liar Paradox in the Bargain: Content Inheritance Disquotationalism

I will call the version of disquotationalism I prefer Content Inheritance Disquotationalism (CID). I make no claims about CID's uniqueness or originality. In fact, I take it as a mark in favor of its plausibility that several other theorists have arrived at the same conclusion for reasons quite independent of paradox-avoidance. Ramsey's view in his classic 1927 paper "Facts and Propositions" clearly falls within the same narrow branch of the disquotationalist family—although Ramnsey's view is often described as a "redundancy" theory of truth and for reasons I will spell out in a moment I take the use of the truth operator to be very frequently non-redundant. Similarly, some contemporary philosophers who call themselves "fictionalists" about truth mostly say things I agree with, although I do not see any particularly fruitful analogy between CID and the sort of views about other domains usually called "fictionalism." The reason to apply a fresh label to these views is simply to "take ownership" of my understanding of some common disquotationalist claims without creating unnecessary confusion. To the extent that other forms of disquotationalism already "on the market" map onto CID, I welcome the company.

CID is simply the view that the propositional content of any sentence S that uses the truth operator is inherited from the content of the sentence(s) to which S attributes truth. (Emphasis on *use*—both this sentence and the sentence before it, for example, mention the truth operator without using it.) Thus, a sentence like our go-to "Truth-Teller," (4), is semantically orphaned—it has nowhere from which to inherit its propositional content.

(4) Sentence (4) is true.

Even if the Liar Paradox did not exist, the most natural way to extend this view to falsity would be to say that sentence that use the negation of the truth operator (or equivalently the phrase "is false") inherit their propositional content from the negation(s) of the sentence(s) to which they are applied. As such, (1) is just as much of a semantic orphan as (4).

(1) Sentence (1) is not true.

A different way of saying all of this is that if any sentence S_1 takes the syntactic form "Sentences $S_{2\text{-}n}$ are true," S_1 expresses all and only the proposition or

propositions expressed by Sentences $S_{2\text{-}n}$. In a simple case like the pair formed by (48) and (49), both sentences express he same proposition.

(48) Snow is white.
(49) Sentence (50) is true.

In a (slightly) more complex case like (50)–(55), (55) expresses the conjunction the propositions expressed by (52), (53), and (54).

(50) Grass is green.
(51) Snow is white.
(52) If Senator Bernie Sanders had won the 2016 Democratic nomination, he would have beaten Donald Trump in the Presidential election.
(53) The Normans conquered England in 1066.
(54) Led Zeppelin's discography represents the aesthetic pinnacle of the great Anglo-American tradition of blues rock.
(55) All of the last five sentences are true.

Similarly, (59) expresses the disjunction of the propositions expressed by (56)–(58).

(56) If Bernie Sanders had been the 2016 nominee of the Democratic Party, he would have beaten Donald Trump in the general election.
(57) If Martin O'Malley had been the 2016 nominee of the Democratic Party, he would have beaten Donald Trump in the general election.
(58) If my miniature schnauzer Lucy had been the 2016 nominee of the Democratic Party, she would have beaten Donald Trump in the general election.
(59) At least one of the last three sentences is true.

The next three cases are more interesting.

(60) All instances of the LEM are true.
(61) No contradictions are true.
(62) Everything Professor Einstein said in that lecture on relativity theory was true.

While (51) expresses a proposition that could just as easily be expressed without use of the truth operator, and (55) and (59) express propositions that *could* be expressed (with a lot more words) without it, (62) expresses a very complex conjunctive proposition that could not be expressed without the truth operator, especially if it was uttered by a first-year physics student who could not hope to reconstruct every detail of the lecture that so impressed him. This example alone shows that the use of "true" is not always redundant, or merely (as Ramsey sometimes says) a "stylistic" choice. And if the first-year physics student could not express what (62) expresses without "true," *no* finite being could express the uncountably infinite conjunctions that (60) and (61) express without the use of "true."[5]

[5] See 7.5 for a discussion of why such conjunctions are *uncountably* infinite.

Examples like (62) point to an additional problem. The explanation of CID above ignores the fact that sentences that use the truth operator often do more than just attribute truth. This presents no difficulties in a case like (63), where the conjunctive proposition expressed by the sentence is simply the conjunctive proposition expressed by (61) with the extra conjunct "grass is green."

(63) Grass is green and (61) is true.

(62), on the other hand, clearly contains propositional content above and beyond the conjunction of all the propositions about relativity uttered by Professor Einstein in the lecture. It also credits him with uttering them all!

The core claim of CID clearly requires an "add-on" to deal with these cases of "truth-with-attribution." The simplest add-on would be one according to which, if Sentence S_1 says that a Sentence S_2, e.g., "said by so-and-so" or "written down on such-and-such blackboard" or some such is true, the proposition expressed by Sentence S_1 is the conjunction of the proposition or propositions expressed by Sentence S_2 and the proposition *that* this proposition was uttered by such-and-such speaker or written down on such-and-such blackboard.

This works well in simple cases such as a three-way conversation between Duncan, Richie, and Methos.

Methos: I think that the time of the Gathering is upon us.
Richie: Duncan, what do you think?
Duncan: I think that what Methos said is true.

If you start out thinking that CID basically gets the function of truth talk right, it is natural to gloss the propositional content Duncan's statement as a conjunction of "According to Methos, 'The time of the Gathering is upon us' and 'The time of the Gathering is upon us'." Although this reading might start to feel wrong, when we apply it to certain cases of blind endorsement—which is (along with infinitely conjunctive propositions, cases where we do not remember everything we were endorsing, and similar cases) a big part of the way that having a truth operator enhances our communicative powers. Imagine a sentence like (62) being uttered not even by a starry-eyed first year physics student who attended the lecture and followed along to the best of his ability (but could not *reproduce* it to save his life) but by a close friend of Professor Einstein who was proud of Einstein's genius and fame but does not know anything about physics, does not pretend to know anything about physics, and was not present at the lecture. Perhaps two people who *do* know about physics are arguing about some contentious detail of the lecture. (Maybe Professor Einstein touched on the vexing question of the relationship between relativity and quantum mechanics.) The close personal friend of Einstein overhears the argument and says, oh, I am sure everything the professor said in his lecture was true! This is surely a case in which the truth operator allows someone to express propositions they could not hope to utter without it. For example, the friend does not know that simultaneity is relative to reference frames and would be confused by the idea if it were explained to him, but by co-signing the lecture with his use of the truth operator, he has still managed to utter a sentence that *expresses*

(among other things) that simultaneity is relative to reference frames. This is exactly the sort of thing the truth operator is good for—expressing propositions we could not express without it since we do not fully remember them, did not know them in the first place, could not say them in one lifetime, etc. Even so, it might seem wrong to say that the friend is expressing the proposition that *Einstein says* that simultaneity is relative to reference frames. This account does not quite capture the spirit of the friend saying, that what Einstein said, *whatever that may have been*, is true. To handle this, we could switch to a reading of the attributive part of the friend's sentence as a universal generalization quantifying over all proposition, so that he is asserting that *for every proposition α,* if Einstein said α, then α.

Do we need a unified account of "truth-with-attribution" or could we plausibly read sentences in the first way when the speaker knows which claim she is endorsing and the second way when she is in the situation of Einstein's friend? I take no position on these questions here, since either (a) reading all such sentences in the first way, (b) reading all such sentences in the second way, or (c) reading some in the first way and some in the second would all be consistent with everything I have to say about the paradoxes.

8.4 Comparing CID With Other Disquotationalisms

If CID is the right way to think about truth talk, truth is not a property. Many philosophers of language with disquotationalist leanings see this as a desirable result, but they end up being forced by their preferred ways of cashing out the disquotation metaphor to concede that truth is at least a "logical property" like existence. "Quantum particles exist" and "epistemic norms exist" are both true, and two sentences do not seem to have much of anything in common. Still, this line of thought holds, they both say the same thing (even if it is not a "substantive" thing) about their respective subjects, and thus attribute the same minimalistic "property" to it. According to CID, however, when we say, "It's true that, 'Snow is white'," we are not really saying *anything* about the sentence or proposition "Snow is white." We are saying something about *snow*—that is, it is white.

Moreover, every instance of the T-Schema straightforwardly follows from this account—and given classical logic, the T-Schema and the Capture and Release rules are a package deal. If α and Tr<α> express the same proposition, then Tr<α> ↔ α is just a disguised instance of the logical truth α ↔ α. It might be objected that this makes the T-Schema trivial, when disquotationalists want it to be informative—indeed, to tell us everything there is to know about truth—but two points are in order on that front. First, the triviality is in a sense desirable. When cashing out the disquotation metaphor, we do not want a "deep" or "substantive" theory of truth. Second, as an analogy to the indexical case shows, the T-Schema *is* informative. It tells us how to read sentences that use the truth operator. If Justin Trudeau is the intended referent of "He is a Canadian," someone who does not know that will find (64) quite informative.

(64) *He* is a Canadian iff *Justin Trudeau* is a Canadian.

Just as (64) tells us how to read the "He" in "He is a Canadian," the T-Schema informs us that the meaning of "It is true that, '. . .'" is entirely given by the removal of the quotation marks. Applying truth talk to propositions does not direct us to some property of the propositions themselves like correspondence or coherence. It is simply a linguistic mechanism for converting mention into use, one that lets us say lots of things about external reality that we could not say without it. These are precisely the benefits that, for example, inferentialist disquotationalism claims for itself—except that the inferentialist disquotationalist has trouble denying that *some* sort of property is being attributed to propositions by uses of "true" and "false." Adopting CID lets us deny this. It also avoids the tonk problem. It also solves the Liar Paradox and does it without begging the question against any non-classical views, since as I hope the last two sections have established, CID is motivated by various considerations totally independent of paradox avoidance.

8.5 Objections and Replies

Objection. This solution does not really solve anything—at least if we are concerned with the NTP and not just the ECP. Even accepting for the sake of argument that such sentences do not express propositions, the question of whether they are true, false, both, neither, or something else is still on the table.

Reply. It straightforwardly follows from CID that sentences that attribute truth or falsity (or any conjunction or disjunction of these properties) to semantic orphans like (1) and (2) are themselves semantically orphaned. We are tempted to say of sentences such as "(1) is true" or "this egg sandwich is true" that they are *not true* to the extent that we think of truth as a property of sentences, propositions, or some other "truth bearers." According to CID, it is not. To say of a sentence that it is not true is to assert the negation of the proposition expressed by that sentence. If a sentence does not express a proposition, there is no such thing as the negation of that proposition, any more than there is a negation of an egg sandwich to be expressed by, "This egg sandwich is not true."

Objection. You say that CID validates all instances of the T-Schema but that is not so. If (1) and (2) are semantic orphans, so are their T-Schema instances. What you have done is at best to provide an independent motivation for the kind of T-Schema-denying solution you dismissed in Chap. 1.

Reply. At a first pass, it is worth noting that even *if* there were genuine T-Schema instances that were not validated by CID, the same would be true of LEM instances, LNC instances, etc. "(1) is either true or untrue" and "It's not the case that (1) is both true and untrue" also come out as semantic orphans. If one insists on describing this as a situation in which the T-Schema sometimes fails, it at least does not fail any more often than (or under different circumstances than) logical truths such as the LEM and the LNC. "Parity with other basic logical truths" is surely about as much as

any disquotationalist could hope to be able to say for the T-Schema. That said, I deny that *any* of these logical generalizations fail given CID.

Presumably, everyone would agree that "Egg sandwich is true iff egg sandwich" does not express a proposition. Of course, that "T-Schema instance" is not a syntactically "well-formed" sentence, but as I argued in 8.1, everyone should accept that some "well-formed" sentences do not express propositions either. If the sentence "This is that" is formed by my cat Shabazz walking over my computer keyboard, I do not see how the (undeniably "well-formed") sentence (65) can express a proposition.

(65) "This is that" is true iff this is that.

If someone "asserted" (65) and "deduced" the "conditional" "the quoted sentences is true if this is that," a reasonable follow-up question would be, "If *what* was *what*? I don't understand what 'this is that' means." There just *is* no answer to this question—nor does our inability to answer it stem from a lack of some relevant piece of information. If we know literally everything there is to know about the sentence, we do not know what it means—because (at least if "meaning" is used in a propositional rather than purely linguistic sense) it does not mean anything.

As I argued in 8.1, propositional logic is about propositions. Those are what the Greek letters should properly be taken to symbolize. The universal generalization that Tr<α>↔α no more "fails" with regard to a "case" where α is a sentence that does not express a proposition than it "fails" where α is a rock. Rather, someone who attempts to represent a rock or an uninterpreted sentence with a Greek letter and perform logical operations on it is making a fundamental mistake.

Objection. Even if a sentence that inherits falsity or untruth to *itself* does not express a proposition, since sentences like (1) and (2) cannot "self-inherit" propositional content, there are many paradoxical sentences that are not self-referential. Take the "Card Paradox" where the sentence on one side of a card says, "The sentence on the other side of this card is true," and the sentence on the other side of the card says, "The sentence on the other side of this card is false." Worse yet, take the series of sentences postulated in Yablo's Paradox, which do not even seem to be *indirectly* self-referential. Each sentence in the series takes the form:

Sentence (N) Every sentence with a rank of (N+1) or higher is false.

Reply. It is controversial whether the Yablo series involves disguised self-reference. Graham Priest has argued that it does, while Otávio Bueno and Mark Colyvan have argued that it does not.[6] I am happy to stay neutral in that debate, though, because my reasons for thinking that (1) and (2) are semantic orphans have nothing to do with self-reference per se. If I have a large amount of money, and someone asked me where I got it, "I inherited it from myself" obviously would not count as any sort of explanation at all, but if I say, "I inherited it from Person 2" and Person 2 says "I inherited it from Person 3" and so on, none of this will add up to an

[6] See Priest, 1997 and Bueno & Colyvan, 2003.

explanation until you get to a Person N who got money by means of something other than inheritance. If Bueno and Colyvan (and Yablo himself) are right about the Yablo series, that case is analogous.

The reason that sentences of the form "Such-and-such sentence(s) are true" *need* to inherit their propositional content from other sentences is attributing a specific property to specific object(s) is both necessary and sufficient for expressing a proposition and "is true" does not name a property. As such, no such sentence can inherit its propositional content from another sentence that uses "is true" (or "is false," or "is untrue," or whatever) unless *that* sentence directly or indirectly inherits its content from some sentence in which truth talk is not used. Every sentence in the Yablo series *would* have propositional content, if for example it ended with sentence (1,533,234), and instead of being another falsity-attributing sentence like the rest of the series, (1,533,234) was, "Snow is white."

Nor is this a finitist constraint. Every sentence in a Yablo-like series of sentences that *started* with "Snow is white" would have propositional content if every sentence from N=2 onward had the format.

Sentence (N): Every sentence with the rank (N-1) or lower is false.

Objection. You say that your version of disquotationalism about truth is not tied to any larger deflationary program, but by advancing a theory of truth that only references propositions and not the facts that make those propositions true, you are failing to accommodate the basic correspondence intuitions underlying any realist view of metaphysics.

Reply. It is true that nothing about accepting CID *forces* the acceptance of metaphysical realism, but it also does not rule it out. If we take true propositions to refer to objects that really exist in external reality and attribute properties to them that they really have, and other propositions that refer to the same objects but attribute properties to them that they *don't* have to be false, we have accommodated the realist intuition that true propositions correspond to reality without loading any account of correspondence into our understanding of the *meaning* of the term "true."

Objection. This solution collapses on content with a revenge paradox like the one revolving around (66).

(66) Sentence (66) is either false or meaningless.

If (66) is meaningless, it is true and hence either both true and false or both true and meaningless, which is just as much of a contradiction. In fact, (66) just reminds us of a bigger problem, which is that false sentences and meaningless sentences are both types of sentences *other than* true sentences. As such, as Graham Priest explains in *In Contradiction*, "non-existence" solutions like the one based on CID do not even get around the Strengthened Liar sentence.

> A desperate way of blocking a couple of steps in the argument is to say that, if *m* is the code of a meaningless sentence, T*m* is itself meaningless, and that the meaninglessness of a part spreads to the meaninglessness of the whole. This line cannot be maintained: if m is the code of a meaningless sentence, then, since it is not true, ¬T*m* is true in the intended interpretation, not meaningless (Priest, 2006b, p. 136).

Reply. (66) is ambiguous, but if "meaningless" means "meaningless in the sense of not expressing a proposition," then (66) is indeed meaningless. To see why it does not follow from this admission that (66) is *true*, consider sentence (67).

(67) Either snow is white or blork gork geblork.

Clearly, someone who said based on the true "first disjunct" that this disjunction was "true" would be making some sort of mistake. More controversially, I would say that the specific mistake they would be making would not be that they were saying something *false* but that they were saying something that did not express a proposition. Of course, (67) is not "well-formed," but as should almost go without saying at this point, we can always substitute in an "empty" use of "This is that."

A crucial move here is one that JC Beall and his co-author Bas C. Van Fraassen mention as part of a discussion of solutions like mine (which they take to face "numerous problems") in their book *Possibilities and Paradox*.

> The idea is nicely captured by a Chinese proverb: *A single jot of rat's dung spoils the soup.* In other words, if A is meaningless then any conjunction is likewise meaningless if A figures as one of its conjuncts; the same goes for all other connectives—the rat's dung spoils the entire soup (Beall & van Fraassen, 2003, p. 126).[7]

The Rat Dung Principle is not some sort of "desperate" ad hoc move resorted to to avoid further paradox. Even if the paradoxes did not exist, and the only semantic orphans were non-paradoxical ones like (4), I do not see how anyone could make sense of saying that (4) does not express a proposition without saying the same about (68) and (69).

(4) Sentence (4) is true.
(68) It is true that "Sentence (4) is true."
(69) Grass is green or sentence (4) is true.

The idea that it is a "desperate move" to say that if (4) is meaningless (in the sense of "meaningless" at issue) so is (68) seems very odd to me. The two are pretty clearly a package deal. If (4) expressed a proposition, (68) would surely express *exactly the same proposition.* This is true both in the sense that given the correct account of the truth operator (CID), (68) could only express propositional content if it got it from (4) and in the sense that, if "is true" named a property of "truth-bearers," (4) and (68) would both be attributing the same property to the same truth-bearer!

Conversely, anyone already inclined to doubt that (4) expresses a proposition should *of course* be equally inclined to doubt that of (68) or (69). Anyone "asserting" either sentence after asserting that (4) did not express a proposition would be faced with the question of what proposition (68) was attributing truth to, and how the

[7] In context, Beall and Van Fraassen were considering Kripke-like views that accord with mine in that that they take sentences like (1) and (2) not to express propositions but diverge from it in that they take the Greek letters of symbolic logic to properly symbolize *sentences*, even sentences that do not express propositions. Such views thus end up with "gappy" formalizations. On the issue of whether I have really or only apparently evaded gappiness, see below.

propositional content of (69) differed from that of the atomic sentence, "Grass is green." Grass is green or *what*? Of course, anyone friendly to the idea that (4) *does* express a proposition would find "Grass is green *or (4) is true*" perfectly satisfying, but I do not see why anyone who was not in that camp when they considered that combination of words on its own should suddenly be convinced that they express a proposition when they are presented by it in the context of this "disjunction."

With the Rat Dung Principle in hand, we can see that (70) poses no special problems.

(70) Sentence (70) does not express a true proposition.

This is just a disguised form of (66), since it is meant to cover both the possibility that (70) is not true and that it does not express a proposition at all.[8] The same points apply to (71), if the implied list of types of mistakes includes the possibility that it would be a *factual* mistake—i.e., calling a false sentence true.

(71) It would be a mistake of some sort to say of (70) that it is true.

The same goes for (72), if "other than" means "false or. . .", no matter how that disjunction is filled out.

(72) Sentence (72) is in a category other than the category of true sentences.

The idea that sentences that do not express a proposition are in a category "other than" true ones trades on the idea that truth is a property. If it is a property, and if we take the hard line according to which "category mistakes" (like "ideas are green") are just *false*—if ideas are not the right kinds of things to have a color, then they *aren't green*—we get the result that meaningless sentences are not true. Of course, we also get the odd result that the cup of coffee on my desk as I type this is not true, which might be counterintuitive enough of a result to count as an extra reason to accept CID, but the larger point is that while I am perfectly happy to take the hard line on "category mistakes," it does not matter because I deny that truth is any sort of property.

One way to try to revive the idea that sentences that do not express propositions are "other than" true might be to think of all the true sentences as forming a set. If so, sentences that do not express propositions are non-members of that set, and the "doesn't express a true proposition" can be reformulated in terms of membership in that set. This gets us (5), from our discussion of Russell's Paradox in Chap. 2.

(5) Sentence (5) is not a member of the set of all true sentences.

If (5) does not express a proposition, surely it is (like every egg sandwich and coffee cup in the world) *not a member of the set of all true sentences*. On the assumption that there is such a set, this is a serious revenge problem for my solution. So—is there such a set?

[8] If we make it "does not express a false proposition," then the disjuncts are not expressing any proposition at all and being not false—i.e., true—and substantially the same points apply.

If naïve set theory were true, such a set would have to exist, but helping oneself to such an assumption while arguing about whether the paradoxes give us a reason to abandon classical logic would beg the question against me. That said, it might not seem like I have any obvious way to rule it out. Nominalism would certainly do the trick, but that is far too large and controversial an assumption to haul out at this point in the argument. The assumption that the set-theoretic universe is exhausted in by the cumulative hierarchy of some particular post-naïve set theory like Zermelo-Frankel would also rule out the existence of the set in question, but as I argued in 2.3, this assumption is unwarranted.

Happily, there *is* a reason, quite independent of revenge-avoidance, to reject the existence of a "set of all truths." In a 1984 article in *Analysis*, Patrick Grim showed that such a set was ruled out by a fundamental result that, far from being an artifact of some particular post-naïve set theory, goes back to Cantor. The set of all the subsets of a set S—the "power set" of S—will always have a greater cardinality than S.

> To each element of the power set will correspond a distinct truth, and thus there will be at least as many truths as there are elements of the power set FF. But by Cantor's power set theorem the power set of any set will be larger than the original. There will then be more truths than there are members of F. Some truths must be left out, and thus F cannot, as assumed, be a set of all truths (Grim, 1984, p. 207).

Most of the philosophical debate generated by Grim's result in more recent decades seems to have focused on the implications for the theology of omniscience.[9] The only attempt I know of to actually undermine the set-theoretic result comes from JC Beall, who argued that there could be a set of all truths with no violation of the power set theorem if we adopt what he calls the "S-View"—the view that "propositions are functions from worlds to truth values" and that as such all *necessary* truths (truths that holds in all worlds) are numerically identical (Beall, 2000, pp. 38–39).[10] Whatever one makes of this move, Beall is clearly talking about true *propositions*, not the sentences that express them. Even if we adopt (what Beall admits is the fairly counterintuitive) position that sentences (73) and (74) express the same proposition, there is no denying that the sentences themselves are distinct.

(73) All bachelors are unmarried.
(74) No squares have six sides.

As such, even if the S-View is right and there can be a set of all the true propositions, there cannot be a set of all true *sentences*, which is what we need for the revenge problem revolving around (5) to get off the ground. To see why, let us try the version formulated in terms of propositions.

---

[9] For the record, I think that the idea that there is some sort of problem there is confused. Theists can use the truth operator to attribute to God an infinite series of conditionals of the form *If α is true, then God believes that α,* and God, if He exists, can simply believe the conjunction of all truths, even if those truths do not form a set.

[10] Beall also hints at a dialetheic response to Grim's result, which would be consistent with Priest's dialetheic approach to the rest of the paradoxes of naïve set theory discussed in 2.3. My response to Priest on that point in 3.2 applies just as well to Beall on the set of all truths.

(75) Sentence (75) does not express a proposition that is a member of the set of all true propositions.

The expression "does not express a proposition that is a member of the set of all true propositions" is just "does not express a true proposition" with an extra disjunct. A sentence can fail to express a true proposition either by failing to express a proposition or by being untrue. A sentence can fail to express a proposition that is a member of the set of all true propositions in both of these ways or because there is no such set. As such, the response to (70) above applies to (75)—whether or not Beall is right that there could be a set of true propositions.

One final note is in order before finishing with this objection. Sentence (76) is just *false* and nothing in CID gives us any reason to think otherwise. Sentence (77) is just *true*. "Expressing a proposition" is a regular predicate that names a property (the property of expressing a proposition), not an expressive device like "is true." Sentences that say of some sentence that they express a proposition do not need to inherit their propositional content from some other sentence. As such, any sentence that refers to any sentence (even itself) and says of it that it expresses or fails to express a proposition does indeed express a proposition.

(76) Sentence (76) does not express a proposition.
(77) Sentence (77) does express a proposition.

Objection. By denying that truth is a property and classifying paradoxical sentences as "semantic orphans" that do not express propositions, this solution fails to accommodate the intuition that at least some such statements *could* have been unproblematically true. For example, in the Card Paradox, the statement on the front of the card ("The sentence on the other side of this card is true") *would have been true* if the sentence on the back ("The sentence on the other side of this card is false") were erased and replaced with a true sentence like, "The Normans conquered England in 1066." While CID technically captures this result in terms of the *sentence*, it does not capture it in a fully satisfying way. According to CID, the sentence would be true at this point because although its linguistic character had not changed, it now expressed a proposition (that the Normans conquered England in 1066) that it did not express before. This is a pretty minimal sense of "could have been true." For example, the sentence "I am a dinosaur" (as uttered by a human being) *could have been true* if it had somehow been uttered by an actual dinosaur—since it would have expressed a proposition (that Dinosaur D is a dinosaur)—that it does not express when uttered by a human being. Surely, when we say that a paradoxical sentence *could have been true*, we mean it in a richer sense than that they would have been true if they had hypothetically been used to express true propositions.

Reply. It is true that CID cannot capture the intuition that some Liars express *propositions* that could have been true, but I am skeptical that anyone who is not deeply invested in the specialist literature on paradoxes has this particular fine-grained technical intuition. A more plausible candidate for the "starting intuition" that drives people to such philosophical views is the far more coarse-grained idea

that there is at least one paradoxical sentence P spoken by speaker S at time T in circumstances C (where C considers all the facts about the speaker's mental states, her level of epistemic access to the sentence to which she is attempting to attribute truth, etc.) that could have been true *when spoken by S at T in C*. This is going to come out true for many paradoxical sentences. The case of the sentence on the front side of the card being repurposed by replacing the sentence on the other side of the card, which closely parallels Goldstein's egg sandwich example from 8.1, is one of these cases. Hartry Field's example of the least intelligent man in the room arrogantly claiming "that what the least intelligent man in the room is now saying isn't true" is another (Field, 2008, p. 76). Such sentences *could* have been true even when still spoken by S at T in C. The same is not true of "I am a dinosaur."

Objection. It still feels like *something* is wrong with your treatment of these cases. Perhaps it is that, if P does not express a proposition at all (when uttered by S at T in C) but *could* have expressed one (while still being uttered by S at T in C), this makes the question of expressing or failing to express a proposition far too contingent and dependent on "the world" as opposed to considerations internal to the linguistic/semantic realm. To say that the *truth value* of a proposition can change as the world changes is one thing, but to say that *what proposition is being expressed* by a sentence (or *whether it expresses a proposition at all*) changes as the world changes is quite another. As Bradley Armour-Garb once said when criticizing a similar view, "such semantic action at a distance seems impossible" (Armour-Garb, 2001, p. 284).

Take the example where you write some obviously true sentence on the whiteboard in your classroom and, just before you say *The sentence on the board is true*, some mischievous student quietly sneaks up behind you, erases it, and replaces it with, "What the professor is saying right now is false." On an account according to which (a) the proposition expressed your sentence is to attribute a truth property to whatever sentence happens to be written on the white board, and for example (b) the result of the student's mischievousness is to render your sentence *gappy*, what is going on here is a normal, even mundane phenomenon—a competent speaker is wrong about the truth value of his utterance because of a contingent fact about external reality of which he was unaware. According to your theory, although the speaker had every reason to believe that he was expressing a true proposition, he is not expressing a proposition at all. Surely competent speakers have more access to the contents of their utterances than this would suggest.

Reply. One of the primary ways in which having a truth operator enhances our communicative abilities is that it allows us to make endorsements "blindly"—to express propositions without knowing exactly which propositions we are expressing. We are doing this when (as in the Einstein case) we use the truth operator to sign onto long strings of propositions we do not entirely remember and perhaps did not entirely understand in the first place. We do this when we use it to express long or even infinite conjunctions or disjunctions, such as "all contradictions are false" and "LEM instances are true." We do it in a more reckless way when we say things such as "everything that such-and-such says is true."

If the mischievous student had been slightly less clever, and instead of "What the professor is saying right now is false," he had written down $2 + 2 = 5$, I would have

been expressing a proposition *that I know to be false* when I said *The sentence on the board is true.* This is one of the risks involved in not having perfect access to the content of my utterances. A closely related risk is that I could simply try and fail to express proposition. The second phenomenon is far rarer than the first, but both happen for the same reason—that the truth operator allows us to say things without being entirely aware of what (if anything!) we are saying. This same kind of "semantic blank" phenomenon could arise in a non-paradoxical context if the mischievous student randomly wrote *This is that* on the board. Even an experienced fisherman can occasionally *think* she has got a fish on the other end of the line when the hook has just snagged on something, and even a competent speaker of a natural language can under these unusual circumstances *think* that he has expressed a proposition when he has fired a semantic blank.

As to the alleged problem about "semantic action at a distance," this is either everyone's problem or no one's problem. Imagine that a substitute teacher writes the following message on an elementary school chalkboard in Lansing, Michigan:

YOU HAVE A NEW SUBSTITUTE TODAY. FUN FACT: HE IS A CANADIAN.

Elementary schools being great big petri dishes of infection, that substitute catches a cold by the end of the day. No one bothers to erase the message. The real teacher is still out on a long-term absence, so the next day a new substitute takes over the class. The new sub also happens to be an immigrant from Canada, so he does not erase the message until the end of the day. Does the propositional content of the message stay "fixed" (so that "today" refers to the previous day and "he" refers to the previous substitute) or do those two sentences now express true propositions?

If they do express new propositions, we have a case of semantic action at a distance that has absolutely nothing to do with disquotationalism or the Liar Paradox. (I have picked a case where the new propositions have the same truth values as the old ones to avoid confusing the issue.) If, on the other hand, those sentences are forever stamped with their original propositional content, the same should apply to sentences that "start out" attributing truth or falsity to some ordinary sentence before those sentences are erased and replaced.

Objection. Even assuming for the sake of argument that this solution works for cases in which truth is attributed to *one* sentence, it feels less plausible for cases in which broad generalizations are made about the truth or falsity of a whole group of sentences, very few of which use the word "true." Consider Kripke's Nixon Paradox.

> Consider the ordinary statement, made by Jones:
>
> (1) Most (i.e. a majority) of Nixon's assertions about Watergate are false.
>
> Clearly, nothing is intrinsically wrong with (1), nor is it ill-formed. Ordinarily the truth value of (1) will be ascertainable through an enumeration of Nixon's Watergate-related assertions, and an assessment of each for truth or falsity. Suppose, however, that Nixon's assertions about Watergate are evenly balanced between the true and the false, except for one problematic case,

(2) Everything Jones says about Watergate is true.

Suppose, in addition, that (1) is Jones' sole assertion about Watergate, or alternatively, that all his Watergate-related assertions except perhaps (1) are true. Then it requires little expertise to show that (1) and (2) are both paradoxical: they are true if and only if they are false (Kripke, 1975, pp. 691–692).

Reply. The late Laurence Goldstein, whose contributions to the literature on paradoxes have been criminally underappreciated, has a beautiful paper in the anthology *The Law of Non-Contradiction: New Philosophical Essays* called "The Barber, Russell's Paradox, Catch-22, God, and More: A Defense of a Wittgensteinian Conception of Contradiction." I disagree with the Wittgensteinian conception of contradiction, but I heartily recommend the paper. One of Goldstein's most useful insights is that some puzzles and paradoxes look harder to crack than they are because take the form of "contradictions immersed in noise."

Were I to tell you that there is a barber, the sole survivor of a plague in a certain village, who shaves himself if and only if he does not shave himself, you would be right to reply 'No there isn't—there cannot be—it is a logical impossibility.' No individual can satisfy the condition: (Sxx ↔ ∼Sxx). [. . .] The Standard Barber Paradox speaks of a village barber who shaves all and only those male villagers of Alcala who do not shave themselves. Suppose there to be n male villagers, v1, . . . vn and the barber b. We are given (with x ranging over the male villagers)

(x)(Sbx ↔ ∼Sxx)

Expanded out, this becomes

(Sbv1 ↔∼S v1v1) & (Sbv2 ↔∼S v2v2) & ... & (Sbv*n* ↔ ∼S vnvn) & (Sbb ↔ ∼Sbb)

or, more shortly,

(Sbb ↔ ∼Sbb) & NOISE

The noise diminishes as plague ravages the village and its population dwindles. [. . .]

The Epimenidean Liar is just a noisy version of the Eubulidean. If to the Eubulidean Liar ('This statement is false') one adds the noise 'and all other statements made by me and my fellow Cretans are false', one obtains the Epimenidean Liar ('All Cretans are liars', spoken by Epimenides, the Cretan). With paradoxes, it is often the case that examining the stripped down, noiseless versions helps us see to the heart of the problem. For example, it is useful to consider the 'Surprise Examination' paradox in a reduced form in which a teacher says to the class 'There will be a surprise examination tomorrow.' (Goldstein, 2004, pp. 296–297)

Kripke's Nixon paradox is, at least as Kripke presents it, essentially a noisy version of the Card Paradox. To reduce the noise to a manageable level, imagine that Nixon only made three statements about Watergate.

(i) I had no prior knowledge of the Watergate break-in. (FALSE)
(ii) Everything Jones says about Watergate is true. (????)
(iii) The break-in took place on the night of June 17th, 1972. (TRUE)

Furthermore, let us take Kripke up on his offer of considering. . .

(iv) Most (the majority) of Nixon's assertions about Watergate are false.

. . .to be "Jones' sole assertion about Watergate." Given Kripke's assumptions, Jones' (iv) amounts to a disjunction of three possibilities—that (i) and (ii) are false, that (i) and (iii) are false, and that (ii) and (iii) are false. Put slightly differently, either (i) and (iii) are false, or both (ii) and one of (i) and (iii) are false. Or:

Either NOISE or both (ii) is false and NOISE.

Notice too that nothing would change about this example if (ii) and (iv) were rephrased to remove the misleading phrase "about Watergate," like so:

(ii)* Jones' statement (iv)* is true.
(ii)* Most of Nixon's assertions in the list of assertions {(i), (ii)*, (iii)} are false.

The phrase "about Watergate" makes us think that these statements are ultimately entirely *about Watergate*, which on the Kripkean analysis they are not. It is no different than taking the card from the Card Paradox and writing some extraneous facts about Watergate under each of the two sentences.

As I have said, all of this is true *given Kripke's treatment of the case*. It is not true given CID. If CID is true, sentences (ii) and (iv) really are *entirely* "about Watergate." Jones' (iv) expresses the odd but clearly false statement that the majority of Nixon's *two* statements about Watergate are false. (The only way for the "majority" of a collection of two statements to be false would be for 100% of them to be false. This is not the case, so (iv) is false.) Nixon's (ii) in turn expresses *exactly the same proposition*. While Kripke has to accept the counterintuitive consequence that these ordinary statements are pathological, I can accommodate the surface appearance that these are ordinary empirical statements with ordinary bivalent truth values.

Objection. You have spent several chapters engaged in "Liar reasoning"—thinking about what follows from applying various logical rules to Liar sentences, how various gappy and glutty solutions run up against revenge problems, and so on, and now you are telling us that logic is about propositions and Liar sentences do not express propositions. Are you claiming, like Wittgenstein at the end of the Tractatus, to be "kicking away the ladder?" How can we understand your earlier claims in light of your preferred solution.

Reply. There is no (male) barber who lives in a town and (succeeds in) shaving all and only the men in the town who do not shave themselves. There could not be—this is a logical impossibility. (There could be a male barber who lived in the town and posted a sign on the window of his barbershop advertising this policy, and every other man in town could play along, but it would be logically impossible for him to apply the policy to himself.) The unrestricted comprehension axiom of naïve set theory must be false for similar reasons. If it were true, the Russell Set would exist, and it cannot. However, despite the fact that these things are logically impossible, we are quite capable of (non-trivially) counterfactual reasoning about them.[11] We know that *if* the unrestricted comprehension axiom were true and there was a Russell Set, the set of all cats would be a member of it, while the set of all sets would not be a member of it, and the Russell Set both would and would not be a member of itself. If

[11] For an in-depth exploration of why there is no conflict between being logical monists about classical logic and accepting that some counterfactual conditionals with logically impossible antecedents are true and others are false, see Daniel Nolan's wonderful paper "Impossible Worlds: A Modest Approach" (Nolan, 1997). Along with the Goldstein paper mentioned above, Nolan's has a safe spot on the list of my topic five favorite papers in the philosophy of logic.

we were not capable of this kind of reasoning, we would not *know* that naïve set theory is false, and the barber cannot exist.

Similarly, when we are told that L=¬Tr<L>, what we are being told is that there is a proposition that is identical to its own negation. This is logically impossible, of course, but we can still reason about what *would* be the case if it were true—how if all LEM instances were true, L would be both true and untrue, how the principle that we should accept all propositions that we have good reason to think are true and the principle that we should reject all propositions we have a good reason to think are untrue would come into conflict with regard to this hypothetical proposition, and so on. The same applies to a case where we were told that a conditional proposition is somehow identical to its own antecedent (as in Curry's Paradox), or where we were told that a proposition is somehow identical to a complex conjunction or disjunction that *includes* its own negation (as in certain "revenge" Liars), and so on.

Objection. What about the validity-curry discussed in 5.1? Even if CID gives us the resources to deal with Standard curry sentences ("If this sentence is true, then. . ."), can it help us with Curry sentences formed by use of Beall's validity predicate (first introduced in Beall & Murzi, 2013)?

Reply. One of the fundamental assumptions of every argument made in this book is that validity is best understood as truth-preservation. As we have seen, this is an assumption that I share with most of my non-classical interlocutors. The exception is Hartry Field, who as we saw in n. 23 above gives up on validity-as-truth-preservation only because he takes it to be forced on him by Curry's paradox—and who comforts himself with the thought that, even if on his account validity is not *identical* to truth preservation, he can still maintain that validity preserves truth "when it matters" (Field, 2009, p. 266). There does not seem to be any reason to drop that assumption here. Thus, V(α, β) reduces to T(α) → T(β) and V-curry is no more of a threat to the view than T-curry.

Objection. It follows from CID that not only paradoxical sentences such as the Strengthened and Simple Liars (1) and (2) but sentences like the "Truth-Teller" (4) are semantic orphans.

(4) Sentence (4) is true.

This is implausible. It is a strike against a solution if it classifies benign, non-paradoxical sentences as pathological.

Reply. Even in a world where no one had ever formulated the Liar Paradox, the "Truth Teller Puzzle" would still be a problem. Is (4) true or false? Even assuming for the sake of argument that this is a meaningful question, the question seems impossible to answer. All the T-Schema tells us about (4) is the supremely unhelpful piece of information that it is true iff it is true. We might be tempted to declare it gappy, but as Graham Priest and Chris Mortensen point out in their paper on the subject, this will not work.

> In view of the fact that truth and falsity are both consistently assignable to (α) [a Truth-Teller], and so there are no a priori proofs of its truth and its falsity, it is tempting to suppose that it is neither true nor false. However, this runs into the following problem. Suppose that (α) is neither true nor false. Then certainly it is not true. But since it says of itself that it is

> true, it is false. This contradicts the supposition that it is neither true nor false. Hence, it is either true or false (Mortensen and Priest 1981, 381–382).

In the first chapter of *Spandrels of Truth*, JC Beall tentatively declares for the sake of uniformity that *all* "*ttruth*-ineliminable" sentences (i.e., the sentences that come out as "semantic orphans" on my account) are gluts.[12] When he gets to Curry's Paradox, this gets him into trouble, since his stance in *Spandrels* is that any suitable conditional had to be detachable, so he introduces the qualification that the "symmetrical" treatment of all "*ttruth*-ineliminable" sentences as gluts "remains, but only for the conditional-free portion of the language" (Beall, 2009, p. 34).

As we saw in 5.1, Beall's new position involves sacrificing Modus Ponens. He may think that an advantage of this view is that it allows him to retain full symmetry by extending the glut treatment of "*ttruth*-ineliminable" sentences to the portion of his language that includes conditionals. If he thinks this, he is mistaken. To see why, consider sentence (78).

(78) Sentence (78) is (a) true and (b) 500 words long.

Unless Beall is willing to abandon Conjunction Elimination along with Modus Ponens, he cannot get symmetry in his treatment of these paradoxical and non-paradoxical instances of these strange and apparently similar sentences. I can.[13]

Objection. This solution only looks like it saves classical orthodoxy because of your insistence on exiling sentences that do not express propositions from the logical realm. This evades the problem rather than solving it. In effect, by declining to attribute *either* truth or falsity to these sentences and declining to either accept or reject them, you have given us yet another gappy or paracomplete solution.

Reply. There is nothing to accept or reject about a sentence that does not express a proposition. If my line on the semantic orphans adds up to gappiness or paracompleteness, then so does everyone else's line on rocks and egg sandwiches and cups of coffee. (Do you accept a rock? Do you accept its "negation?" Is there such a thing as a rock's negation?) You can say that rocks and semantic orphans are different because rocks are not well-formed *sentences*, but by saying that you are admitting that the question explored in 8.1 about whether logic is best understood as sentential or proposition is a substantive issue. If you want to object to it, you need an argument and not just an accusation of evasiveness.

There is also the question of non-paradoxical orphans, such as the random utterance of *This is that*. Do you "accept" that sentence? The idea of doing so is

[12] In their paper on the "Truth Teller Paradox," Priest and Mortensen toy with calling Truth Tellers gluts, but do not pursue the possibility. "While it is arguable that this position can be maintained without collapse into triviality, it would tempt few" (Mortensen and Priest 1981, pp. 383-384). One wonders if they would have been bolder in pushing this possibility if the paper had been written after the publication of *In Contradiction.*

[13] Of course, Field can get it too. As far as I can tell, applying his paracomplete solution to sentences like (4) as well as sentences like (1) and (2) would not create any *new* problems for his view. All I can say about that is, "See Chap. 7 for an exploration of all the pre-existing problems!"

deeply counterintuitive, but the idea of *rejecting* it is equally counterintuitive. Applying the language of acceptance or rejection to it just looks like a mistake.

If you insist on calling a refusal to grant that certain sentences that have the linguistic character of instances of Bivalence or the LEM amount to *rejecting* Bivalence or the LEM, then you should equally insist that I am "rejecting" the LNC and the Law of Identity and *every other logical generalization*, since sentences with the linguistic character of instances of *any* logical generalization can be formed around semantic orphans like (1), (2), and (4). Similarly, if I am supposed to be paracompletishly rejecting the inference *to* LEM instances from arbitrary premises because I classify "Either (1) or not-(1)" as a semantic orphan that we cannot sensibly speak of as either entailing or being entailed by anything—any more than we can speak sensibly of a rock entailing or being entailed by anything—then my solution is just as *paraconsistent* as it is paracomplete. After all, my analysis of alleged "inferences" from contradictions *to* sentences like (1) and (2) would be the same. In fact, by parity of reasoning, if I in effect "reject" the same laws of classical logic as a gappy or paracomplete theorist would, I also "reject" *every single other logical rule* in exactly the same way. If someone insists on talking this way, I cannot stop them, but I would accuse them of using words such as "reject" in a non-standard and unhelpful way.

In any sense in which I am not a logical nihilist, I reject gaps and gluts and accept the full machinery of classical logic. We can argue about whether that position is correct, but it is far less productive to argue about whether I "really" hold it.

8.6 Comparing the Solutions: The Principle of Uniform Solution

Earlier in the book, I committed to the idea that any adequate classical solution to the Liar Paradox should (a) be at least uniform as Priest's dialetheic solution (for which he claims just such a virtue of uniformity) or a hypothetical just-as-uniform paracomplete account, and in particular (b) solve the Irrationalist's Paradox, and (c) solve Curry's Paradox. Let us take these in reverse order.

Curry sentences are on my account clearly semantic orphans. This avoids the need for an inferential Sophie's choice between Conditionalization and Modus Ponens.

Things are slightly less clear when it comes to sentences like (79).

(79) A fully rational agent in a situation of full information would not accept sentence (79).

As we saw in Chap. 5, sentences like this (along with ordinary Strengthened Liars) force Priest to accept that there are rational dilemmas about accepting and rejecting sentences. This, in turn, makes it impossible to hold onto Principle R. As I argued in depth in Chaps. 5 and 6, the demise of Principle R is a death blow to the

whole project of combining dialetheism with any plausible theory of what is going on when we (apparently rationally) use classical inferences such as Disjunctive Syllogism. There just is no substitute.

Sentences like (79) should be about as much of a problem for Field's paracomplete view, considering the outsized role of acceptance and rejection in his account. What about my view? It depends on whether you are thinking that something should be rationally accepted has anything to do with thinking that it is rational to believe that it is *true* (and conversely whether thinking that something is rationally rejectable has anything to do with thinking that it is rational to think it is *false*). It seems at least plausible that this is the right view, and that the combination of such an account with CID generates the result that (79) too is an orphan.

So far, things are looking good on the Irving Berlin-inspired promissory note from 1.6, that I would show that "anything you (non-classical paradox-solvers) can do, we (classical paradox-solvers) can do better." The CID-based solution to the Liar Paradox solves it in just the way all classical logicians solve Russell's Paradox—by denying the Existence part of the Inclosure Schema—and the same way that *everyone* (even dialetheists and their paracomplete counterparts) solve IS paradoxes such as the Barber Paradox. Even those who take the Barber to be logically possible do not think it is epistemically possible that some barber in some distant place actually succeeds in shaving all the men in the town (including himself) who do not shave themselves. Most promisingly of all, CID gives the classical logician the ability to extend this uniformity to Curry's Paradox, which is where (as we saw in 5.1) uniformity breaks down most catastrophically for any paraconsistent or paracomplete logician unwilling to sacrifice Modus Ponens.

8.7 Conclusion: Vagueness, Quantum Weirdness, and Other Loose Ends

Some people have thought a gappy or glutty approach to the odd properties of quantum particles is promising, but this does not even get the superficial facts of the case right. I am skeptical about the idea that the right solution to these puzzles is to change the logic, but even if it was, changing it by adding gappy or glutty options to the truth tables of classical logic would not get the right result. Those logicians (such as Putnam 1969 or more recently Dickson, 2001) who *are* willing to reject classical logic to make sense of the oddities of the Copenhagen Interpretation of quantum mechanics have in their sights **Distribution** rather than, say, Disjunctive Syllogism.

Distribution: $\alpha \wedge (\beta \vee \delta)$ // $(\alpha \wedge \beta) \vee (\alpha \wedge \delta)$

This logically revisionary interpretation of the quantum physics (henceforth, following Dickson, we will just call it "the quantum-logic interpretation") has

sometimes been argued for based on the two-slit experiment.[14] Unlike the non-classic theorists we have looked at earlier in this book, which tend to run into problems when it comes time to extend their logic into probability theory, the quantum logician can move in the other direction. The distribution of photons through the slits are best understood in terms of a non-classical probability calculus, which in turn calls for a revision of the underlying logic in favor of one in which Distribution fails.

As the philosopher of physics Eddy Cheng has pointed out to me in conversation, this result seems to rest on a fallacy of equivocation. Does "the probability of a given particle passing through Slit A" mean "the probability that a given particle would pass through Slit A *while Slit B was open*" or "the probability that a given particle would pass through Slit A *without Slit B being open*?" Perhaps the quantum logician could get around this objection by adding a premise to her argument specifying that we can be sure that the two possibilities have the same probability, but whether this premise would be plausible is a substantial further question.

The most recent attempt to revive the quantum-logic interpretation that I am aware of is Michael Dickson's cleverly titled article, "Quantum Logic Is Alive ∧ (It Is True ∨ It Is False)." Dickson avoids the usual objection that quantum logicians are postulating a mysterious "logical split" between the quantum and macroscopic realms by arguing that "the distributive law is wrong," full stop.

> It is not wrong 'for quantum systems' or 'in the context of physical theories' or anything of that sort. It is just wrong, in the same way that '(p or q) implies p' is wrong (Dickson, 2001, p. S275).

This of course leaves him open to a different objection. If Distribution is "just wrong," then how can we explain the fact that—unlike the inference from "p or q" to p—it just happens to not have any macroscopic counterexamples?

Dickson's answer is that Distribution only fails in cases involving "incompatible" propositions and "the propositions applicable to the macroscopic world are, or are effectively, compatible" (Dickson, 2001, p. S283). If "compatible" is read in any normal way, this is a deeply confusing claim. "Donald Trump will win the 2020 election" and "A Democrat running against Donald Trump will win the 2020 election" are incompatible propositions in the usual sense, but "The 2020 election will take place and (Donald Trump will win or a Democrat running against him will win)" certainly *seems* to entail that "Either (the 2020 election will take place and Donald Trump will win) or (the 2020 election will take place and a Democrat running against him will win)." Dickson clarifies:

> (By 'effectively compatible' I mean that the probability of finding interference effects, or violations of distributivity, is extremely small.) (Dickson, 2001, p. S283)

So, to review, Distribution is "just wrong." It does, however, hold of some cases. Which cases? Ones for which the probability that it will not hold is very low.

[14] To be clear, Putnam uses this argument. Dickson does not.

Putting the "logical split" issue to one said, there is a deeper problem with the quantum-logic interpretation. If one accepts that the standard truth tables for disjunction and conjunction are correct:

$\alpha \vee \beta$	$\alpha \wedge \beta$
T T T	T T T
T T F	T F F
F T T	F F T
F F F	F F F

. . .then how can we deny that Distribution is truth-preserving?

$\alpha \wedge (\beta \vee \gamma)$ // $(\alpha \wedge \gamma) \vee (\alpha \wedge \gamma)$	
T T T T T	T T T T T T T
T T T T F	T T T T T T F
T T F T T	T F F T T T T
T F F F F	T F F T T T F
F F T T T	F F T F F F T
F F T T F	F F T F F F F
F F F T T	F F F F F F T
F F F F F	F F F F F F F

Let us call this the Objection from the Tables. Dickson's reply is that the view that "the meanings of, and therefore the laws obeyed by, the connectives are established by appeal to truth-tables [. . .] seems to come from some principle such as the following: if each of two sentences is true under the same conditions, then the sentences must mean the same" (Dickson, 2001, p. S282). Dickson replies by arguing that it is at least "not obvious" that this principle is true, and that "if not, then 'truth under the same conditions' is not sufficient for 'sameness of meaning', and the account of the meaning of the sentential connectives in terms of truth tables loses its initial appeal." As such, "those who give up the truth-table view are not giving up anything that is obviously true" (Dickson, 2001, p. S282).

Now, I agree with Dickson that it is far from obvious that "truth under the same conditions" is is *sufficient* for "sameness of meaning." For example, the material conditional $(\alpha \rightarrow \beta)$ is true under the same circumstances as $(\neg\alpha \vee \beta)$. . .

$\alpha \rightarrow \beta$	$\neg\alpha \vee B$
T T T	FT T T
T F F	FT F F
F T T	TF T T
F T F	TF T F

. . .but I find it quite plausible that the two expressions have different meanings. However, for the Objection from the Tables to go through, all we need is the considerably more obvious principle that "truth under the same conditions" is at least *necessary* for "sameness of meaning." If the quantum logician denies that conjunctions and disjunctions are not true under the circumstances in which they come out true in the classical truth table, why assume that she is offering a heterodox

view of the logical behavior of *those connectives* rather than changing the subject and talking about new connectives?

It is important to note that a parallel Objection from the Tables made against a dialetheist or a paracompletist would be much less persuasive. To see why, think about, for example, the dialetheist's argument that there are counterexamples to Disjunctive Syllogism. Here is one version of what a dialetheist truth table for Disjunctive Syllogism can look like, with "B" representing "both true and false."

$\alpha \vee \beta$	/	$\neg\alpha$	//	β
T	T	T	FT	T
T	T	F	FT	F
T	T	B	FT	B
F	T	T	TF	T
F	F	F	TF	F
F	B	B	TF	B
B	T	T	BB	T
B	B	F	BB	F
B	B	B	BB	B

The dialetheist does not deny that disjunctions and negations are true under the circumstances in which the classical logician thinks they are true or false under the circumstances under which the classical logician thinks they are false. The "classical" lines of the truth table above just *are* the classical truth table. The lines highlighted in red represent circumstances which the dialetheist takes to be logically possible, but which were not considered in the classical table. The dispute between the dialetheist and the classical logician can thus be plausibly understood not as a debate about the meaning of negation (or disjunction) but about the whether the best theory of logic is one that admits these (alleged) logical possibilities.

There is no parallel way to get the result that Distribution has counterexamples by adding extra options to the truth tables of classical logic.[15] This leads me to find the possibility that logical revision is the right way to make sense of the quantum results dubious, since if the straightforward truth preservationist way of making sense of validity is the right account, we *should* be able to get this result by fiddling with the truth tables, and if some sort of semantic or proof-theoretic way of thinking about the validity is correct, it is mysterious how empirical investigation should show us that we were wrong about issues such as "what our logical terms mean" or "what rules we implicitly use to define connectives."

In any case, even if the quantum-logic interpretation *was* right, not only would this not bolster the case for gaps or gluts, it would also undermine it. If Distribution fails, it can be true that "[(1) expresses a proposition] and (That proposition is either true or untrue)" without it being true that "((1) expresses a proposition and it is true) or ((1) expresses a proposition and it is false)."

[15] There was a time when this possibility intrigued me, and I devoted many hours of my life to the attempt. I am happy to leave it here as an exercise to those readers with time to burn.

Moving from the cutting edge of empirical science to the puzzles of imprecise language, Edwin Mares has suggested that some terms are "overdefined" just as several previous theorists have suggested that some terms are "underdefined."[16] An underdefined predicate is one for which the rules (or raw specification of cases) governing which objects fall under it and the independent rules (or raw specification of cases) governing which objects fall under its negation will be jointly incomplete. A popular line of thought holds that this purely linguistic mechanism generates "gaps" without "the world" having to be somehow metaphysically gappy. Mares makes the more unusual suggestion that a parallel phenomenon involving the two sets of rules getting tangled up so that a predicate and its negation *both* apply to some object generates gluts without the world having to be metaphysically glutty. I take Mares to be entirely correct when he says that if we accept underdefined predicates, we should accept overdefined predicates in the bargain, but as I hope to have showed in the discussion in Chaps. 4, 5, and 6, the consequences of accepting dialetheism for our ordinary reasoning practices are grim. In his article "The Law of Non-Contradiction as a Metaphysical Principle," Tuomas Takho speculates that "most philosophers" would join him in "happily accepting" the purely "semantic" dialetheism advocated by Mares (Takho, 2009, p. 38). I hope he is wrong, since this would mean that most philosophers would happily reject Disjunctive Syllogism and the rest whenever reasoning in the languages that we all speak. Instead, I would suggest taking "overdefined predicates" as a *reductio* of "underdefined predicates." Our theory of negation should be very simple. To fall under the extension of the negation of a predicate is simply to fail to fall under the extension of the predicate itself. No other rule applies.

. . .which brings us, at last, to the thorny subject of vague predicates. Many people have thought that borderline applications of vague predicates cannot be fit into the simple view advocated above. Here is a nice intuitive example from *Saving Truth from Paradox*.

> Consider the following argument:
> (1) Bertrand Russell was old at age 3×10^{18} nanoseconds (that's about 95 years)
> (2) He wasn't old at age 0 nanoseconds
> (3) So there is a number N such that he was old at N nanoseconds and not old at k nanoseconds for any $k < N$ (Field, 2008, p. 100).

Is there really going to be a sharp cutoff point in this progression—a nanosecond N such that Russell was old at after N but was not old for all nanoseconds before N? Even if we could slow down his life and examine each nanosecond in turn, it does not seem plausible that we could ever decide on "the nanosecond at which Russell became old."

The situation is even worse if natural language is really littered, as many philosophers who work on vagueness seem to think it is, with pairs of predicates such that each is logically equivalent to the negation of the other, but neither seems to be more conceptually basic than the other. As I hinted in 6.3, I am skeptical of this

[16] See Mares (2004).

claim. The fact that the stock example is "bald" and "haired," the latter being a painfully artificial label slapped onto what any ordinary speaker would just call "clearly not bald," should at least give us pause. I also think that JC Beall's view on "overline mates," briefly developed in *Spandrels of Truth*, is a promising beginning of a plausible classical theory of one of the main sources of the phenomenon.[17] Natural language *is* littered with pairs of predicates such that each member of the pair is *entailed by* (but not quite logically equivalent to) the negation of the other. Take "being in the room" and "being outside of the room." They are clearly mutually exclusive, but they are not quite jointly exhaustive. You can stand in the doorway.

In any case, I do not claim that any of this adds up to anything like a fully satisfactory theory of the myriad conceptual headaches involved in a complete investigation of vagueness. It is a complex phenomenon, and I agree with Beall that "there's no obvious reason to treat vagueness and *ttruth*-theoretic paradox the same way," given that the two sets of headaches seem to have very different sources (Beall, 2009, p. 130) Still, both paraconsistent and paracomplete theorists *have* claimed to provide such a uniform solution. Let us turn to those now.

We got a glimpse of Field's account in 7.2. He tries to accommodate the intuition that there are cases in which it would be arbitrary to apply either a vague predicate or the negation of that predicate to a given object by placing such borderline cases in the same paracomplete negative zone as Liars and Truth Tellers. If the number of stones is N and this is a clear case of *not a pile*, it is implausible to call N+1 stones *a pile*. Field's view seems to solve this, since if the. . . .jumble?...of N+1 stones is in the negative zone, the claim that it is a pile and the claim that it is not a pile should both be rejected, but it really only "solves" the problem by relocating it. I cannot see any reason why "if we should accept the claim that N stones is not a pile, it's implausible to reject the claim that N+1 stones *is* a pile" should be any more plausible (or, to be fair, any *less* plausible) than "if the number of stones is N and this is a clear case of *not a pile*, it's implausible to call N+1 stones *a pile*."

Zach Weber's glutty view, developed in his article "A Paraconsistent Model of Vagueness," seems to solve this problem of sharp cut-off points in a clever way. It seems that there both *must* be a sharp cut-off point (since at some point well before age 95, he must have become old). It also seems like there *can't* be sharp cut-off point, since any choice of cut-off is hopelessly arbitrary. With characteristic flair, Weber concludes that, "The solution to the paradox is that it is true" (Weber, 2010, p. 1026). The idea is that if the borderline applications are *all* glutty, they are *all* sharp cut-off points—it will be true for each of them that every lower number of stones is *not a pile* and that every number of stones equal to or greater than them *is a pile*. As clever as this is, it still "solves" the problem by relocating it. At what point does it *stop* being true that the number of nanoseconds that have passed in Russell's life make him *not old*? If there are gluts, this is not the same question as when it starts to be the case that he is old.

[17] See Beall, 2009, pp. 107-108. Beall is of course a non-classical paradox-solver when it comes to the Liars, but it is important to him to *restrict* non-classicality to that domain.

Either way, as with the puzzles of quantum mechanics, postulating gaps or gluts will not help. Meanwhile, when it comes to the truth-theoretic paradoxes, a classical solution fares far better than the paracomplete and paraconsistent alternatives. It certainly avoids the considerable epistemic costs of those alternatives. As such, on balance, our best theory of logical truth includes Bivalence, the LNC, and the LEM. Put differently, we should continue to accept a logic without gaps or gluts.

References

Aberdein, A. (2008). Logic for dogs, In Hales, pp. 167–181.
Almog, P., & Wettstein (Eds.). (1989). *Themes from Kaplan*. Oxford University Press.
Arenhart, J., & Melo, E. (2017). Is dialetheism self-coherent? *South American Journal of Logic, 3*(1), 87–109.
Armour-Garb, B. (2001). Deflationism and the meaninglessness strategy. *Analysis, 61*, 280–289.
Azzouni, J. (1995). *Mathematical myths, mathematical practice: The ontology and epistemology of the exact sciences*. Cambridge University Press.
Azzouni, J. (2006). *Tracking reason: Proof, consequence, and truth*. Oxford University Press.
Beall, J. C. (2000). A neglected response to the grim result. *Analysis, 60*, 38–41.
Beall, J. C. (2005). Transparent disquotationalism. In Beall & Armour-Garb (Eds.), pp. 7–22.
Beall, J. C. (Ed.). (2007). *Revenge of the liar*. Oxford University Press.
Beall, J. C. (2009). *Spandrels of truth*. Oxford University Press.
Beall, J. C. (2013). Free of detachment: Logic, rationality and gluts. *Noûs, 49*, 410–423.
Beall, J. C. (2014a). End of inclosure. *Mind, 123*, 829–849.
Beall, J. C. (2014b). Finding tolerance without gluts. *Mind, 123*, 791–811.
Beall, J. C. (2015). Triviliazing sentences and the promise of semantic completeness. *Analysis, 75*(4), 573–584.
Beall, J. C., & Armour-Garb, B. (Eds.). (2004). *The law of non-contradiction: New philosophical essays*. Oxford University Press.
Beall, J. C., & Armour-Garb, B. (Eds.). (2005). *Deflationism and paradox*. Clarendon Press.
Beall, J. C., & Bueno, O. (2002). The simple liar without bivalence? *Analysis, 62*, 22–26.
Beall, J.C. & Murzi, J. (2013). Two Flavors of Curry's Paradox. *The Journal of Philosophy 110*(3), pp. 143–165.
Beall, J. C., & Restall, G. (2006). *Logical pluralism*. Clarendon Press.
Beall, J. C., & van Fraassen, B. C. (2003). *Possibilities and paradox: An introduction to modal and many-valued logic*. Oxford University Press.
Bellos, A. (2016). Can you solve it? The logic puzzle almost everyone gets wrong. *The Guardian*, 28 March 2016.
Berkeley, G. (1713). *Three dialogues between Hylas and Philonous*, in the version presented www.earlymoderntexts.com
Berto, F. (2006). Meaning, metaphysics, and contradiction. *American Philosophical Quarterly, 43*, 283–297.
Berto, F. (2007). *How to sell a contradiction: The logic and metaphysics of inconsistency*. College Publications.
Black, M. (1952). The identity of indiscernibles. *Mind, 61*, 153–164.

© Springer Nature Switzerland AG 2022

B. A. Burgis, *Logic Without Gaps or Gluts*, Synthese Library 458,
https://doi.org/10.1007/978-3-030-94624-1

Bueno, O., & Burgis, B. (2019). Liars with curry: Dialetheism and the prospects for a uniform solution. In *Dialeteism and its applications* (Trends in logic, Vol. 52). New York: Springer.
Bueno, O., & Colyvan, M. (2003). Paradox without satisfaction. *Analysis, 63*, 152–156.
Burgis, B. (2011, October). Paracompleteness and revenge. *Analysis, 71*(4), 688–689.
Caret, C., & Hjortland, O. (2015). *Foundations of logical consequence*. Oxford University Press.
Chihara, C. (1984). The semantic Paradoxs: Some second thoughts. *Philosophical Studies, 45*, 223–229.
Coghlan, A. (2011). Parrots join apes and Aristotle in the club of reason, *New Scientist*, 22 June 2011.
Cook, R. (2005). What's wrong with Tonk(?). *Journal of Philosophical Logic, 34*, 217–226.
Denyer, N. (1995). Priest's paraconsistent arithmetic. *Mind, 104*, 567–575.
Dickson, M. (2001). Quantum logic is alive ∧ (it is true ∨ it is false). *Philosophy of Science, 68*, S74–S287.
Evnine, S. (2008). Epistemic Dimensions of Personhood. Oxford University Press.
Field, H. (2005). Is the liar sentence both true and false? In Beall & Armour-Garb (Eds.), pp. 23–40.
Field, H. (2008). *Saving truth from paradox*. Oxford University Press.
Field, H. (2009). What is the normative role of logic? *Proceedings of the Aristotelian Society, 83*, 251–268.
Field, H. (2015). What is logical validity? In Caret & Hjortland (Eds.), pp. 33–70.
Garfield, J. (2004) To pee and not to pee? Could that be the question? (Further reflections of the dog). In Beall & Armour-Garb (Eds.), pp. 235–244.
Goldstein, L. (2001). Truth bearers and the liar: A reply to Alan Weir. *Analysis, 61*, 115–126.
Goldstein, L. (2004). The Barber, Russell's paradox, catch-22, god, and more: A defense of a Wittgensteinian conception of contradiction, In Beall & Armour-Garb (Eds.), pp. 295–313.
Grim, P. (1984). There is no set of all truths. *Analysis, 44*, 206–208.
Grim, P., & Plantinga, A. (1993). Truth, omniscience, and Cantorian arguments. *Philosophical Studies, 71*, 267–306.
Hales, S. (Ed.). (2008). *What philosophy can tell you about your dog*. Open Court.
Harman, G. (1986). *Change in view: Principles in reasoning*. MIT Press.
Hofweber, T. (2007). Validity, paradox and the ideal of deductive logic. In Beall (Ed.), pp. 145–158.
Horwich, P. (1998). *Truth* (2nd ed.). Oxford University Press.
Kaplan, D. (1989). Demonstratives. In P. Almog & Wettstein (Eds.), pp. 481–563.
Kripke, S. (1975). Outline of a theory of truth. *The Journal of Philosophy, 72*, 690–726.
Lewis, D. (1976). The paradoxes of time travel. *American Philosophical Quarterly, 13*, 145–152.
Mares, E. (2004). Semantic Dialetheism. In Beall & Armour-Garb (Eds), pp. 264–275.
Martin, R. L. (Ed.). (1984). *Recent essays on truth and the liar paradox*. Clarendon Press.
Maudlin, T. (2004). *Truth and paradox: Solving the riddles*. Oxford University Press.
Mortensen, C. (2004). Dharmakīrti and priest on change. *Philosophy East and West, 54*, 20–28.
Mortensen, C. & Priest, G. (1981). The Truth Teller Paradox. *Logique et Analyse 24*(95/96), pp. 381–388.
Musgrave, A. (1999a). Critical rationalism and logical omniscience. In Musgrave (1999b).
Musgrave, A. (1999b). *Essays on realism and rationalism*. Rodopi.
Nolan, D. (1997). Impossible worlds: A modest approach. *Notre Dame Journal of Formal Logic, 38*, 535–572.
Priest, G. (1985). Inconsistencies in motion. *American Philosophical Quarterly, 22*, 339–346.
Priest, G. (1994). The structure of the paradoxes of self-reference. *Mind, 103*, 25–34.
Priest, G. (1996). On paraconsistent arithmetics: A reply to Denyer. *Mind, 105*, 649–659.
Priest, G. (1997). Yablo's paradox. *Analysis, 57*, 236–242.
Priest, G. (1998). The import of Inclosure: Some comments on Grattan-Guiness. *Mind, 107*, 835–840.
Priest, G. (2000). On the principle of uniform solution: A reply to Smith. *Mind, 109*, 123–126.
Priest, G. (2002). *Beyond the limits of thought* (2nd ed.). Clarendon Press.

Priest, G. (2004). What's so bad about contradictions? In Beall & Armour-Garb (Eds.), pp. 23–38.
Priest, G. (2005). Spiking the field artillery. In Beall & Armour-Garb (Eds.), pp. 41–52.
Priest, G. (2006a). *Doubt truth to be a liar*. Oxford Clarendon Press.
Priest, G. (2006b). *In contradiction* (2nd ed.). Clarendon Press.
Priest, G. (2012). The sun may not, indeed, rise tomorrow: A reply to Beall. *Analysis, 72*, 739–741.
Prior, A. N. (1960). The Runabut inference-ticket. *Analysis, 21*, 38–39.
Quine, W. V. O. (1970). *Philosophy of logic*. Harvard University Press.
Ramsey. (1927). Facts and propositions. *Proceedings of the Aristotelian Society, 7*, 153–170.
Russell, B. (1903). Principles of Mathematics. Cambridge University Press.
Russell, B. (1959). *My philosophical development*. Simon and Schuster.
Sharp, K. (2018). Shrieking in the face of vengeance. *Analysis, 78*(3), 454–463.
Smith, N. J. J. (2000). The principle of uniform solution (of the paradoxes of self-reference). *Mind, 109*, 117–122.
Statman, D. (1995). Moral Dilemmas. Brill Rodopl.
Takho, T. (2009). The law of non-contradiction as a metaphysical principle. *Australasian Journal of Logic, 7*, 32–47.
Tanaka, K. (2007). Dharmakīrti and priest on an inconsistent theory of change – A comment to Mortensen. *Philosophy East and West, 57*, 244–252.
Wansing, H. (2006). Connectives stronger than Tonk. *Journal of Philosophical Logic, 35*, 653–660.
Wansing, H., & Priest, G. (2015). External curries. *Journal of Philosophical Logic, 44*, 453–471.
Weber, Z. (2010). A paraconsistent model of vagueness. *Mind, 476*, 1025–1045.
Weber, Z., Ripley, D., Priest, G., Hyde, D., & Colyvan, M. (2014). Tolerating Gluts. *Mind, 123*, 813–828.
Weir, A. (2004). There are no true contradiction. In Beall & Armour-Garb (Eds.), pp. 385–417.
Woods, J. (2019). Logical partisanhood. *Philosophical Studies, 176*, 1203–1224.
Young, G. (2015). *Revenge: Dialetheism and its expressive limitations*. Ph.D. thesis, University of Glasgow.

Index

A
Accept
  Accept(F), 75
  Accept(G), 94
Acceptability, 75, 77, 80, 83, 92, 95, 97
Acceptance, 3, 9, 12, 64, 75, 77–79, 82, 83, 92, 93, 95–101, 120, 121, 132–134, 162, 173, 174
Anti-Foundation Axiom (AFA), 23, 34, 35
Aristotle, 108
Axiom of choice, 23

B
Barber Paradox, 34, 169, 174
Bayesianism, 121
Beall, J.C., 4, 18, 33, 55, 56, 60–62, 64, 67, 120, 126, 130, 133, 134, 136, 138, 141, 142, 154
Berkeley, Bishop George, 8, 15, 25–27, 31, 44–47, 98
Berkeley's Master Argument
  as Berkeley's Paradox, 31
Berto, Francisco, 99–101, 119, 121–124
Bivalence, 9–12, 52, 135, 173, 180
Black, Max, 44–47
Burali-Forti Paradox, 20
Buridan's Ass, 74

C
Cantor, Georg, 165
Capture and Release, 3–4, 6, 12, 13, 75, 76, 94, 96, 136, 137, 142, 153–155, 159
Card Paradox, 161, 166, 169, 170
Chrysippus, 60
Classical logics, 1–5, 7, 9, 12–15, 31–48, 50, 58, 59, 72, 85–87, 93, 96, 120, 121, 125, 126, 131, 137, 138, 143, 147–180
Classical recapture, 14, 59–64, 104, 120–121
CompoundPrec, 150
  See also Prec
CompoundProp, 148, 150
  See also Prop
Conditional
  classical, 122
  and detachability, 127, 172
Conditionalization, 72, 126, 173
Conditional probability, 132, 133
Conjunction Elimination, 5, 32, 58, 86, 153, 172
Contraction, 66, 67, 69, 71, 72, 126
Contraposition, 54, 55, 57, 71, 76, 88, 91, 106, 109, 113
Curry, Haskell, 66
Curry sentences
  and conditionals, 71, 126, 173
  Lying Curry, 138
Curry's Paradox, 50, 65–101, 126, 171–174

D
Degree of belief, 9, 128, 133, 135
DeMorgan's laws, 5, 19
Denying the Antecedent, 55
Deontic logics, 32
Descartes, René, 108
Determinate truths, 140, 142, 143

© Springer Nature Switzerland AG 2022

B. A. Burgis, *Logic Without Gaps or Gluts*, Synthese Library 458,
https://doi.org/10.1007/978-3-030-94624-1

Dialetheism, 6–9, 11–13, 15, 17–29, 36, 38, 48–101, 103–124, 130, 134, 141, 149, 174, 178
Disjunction addition, 5, 50
Disjunction introduction, 153
See also Disjunction additional
Disjunctive Syllogism (DS)
local validity of, 119
Disquotation, 136, 152, 153, 156–159
Disquotationalism
CID (Content Inheritance Disquotationalism), 147, 156, 160
Distribution, 59, 82, 174–177
Double Negation, 5–7, 19, 33, 57, 72, 86, 94, 113, 127, 135

E
Einstein, Albert, 36, 99, 100, 157–159, 167
Entailments, 14, 52, 54, 55, 57, 58, 67, 76, 88, 91, 92, 94, 104, 128, 129, 134
Exclusion Principle
A to R Exclusion Principle, 86
R to A Exclusion Principle, 86
Ex contradiction quodlibet, 66, 122
Exhaustive Characterization Project (ECP), 141–143, 160
Explosion principle, 4, 5, 58
See also Ex contradiction quodlibet

F
FailToAccept, 93, 94
Falsity preservation, 54, 57–58
Field, Hartry, 5, 8, 9, 11, 15, 19, 22, 51, 53, 61, 82, 95, 124–146, 167, 171, 178
Finitism, 162
Frege, Gottlob, 99

G
Geach, Peter, 74
Gödel, Kurt, 70, 105
Gödel's first incompleteness theorem, 105
Grelling's Paradox, 145

H
Harman, Gilbert, 2, 3, 15, 84
Hegel, G.W.F.
hegeleanism, 41–43
Hypersets, 23

I
Idealism
Berkeley's, 44
Theistic, 25
Identity of indescernibles, 47
Inclosure Schema (IS)
and the element of Closure, 22, 70
and the element of Existence, 22
and the element of Transcendence, 22, 70
and schema paradoxes, 25
Indeterminism, 42
Inference rules
classical, 39, 87
as truth preserving, 50, 51
Inferentialism
disquotationalist inferentialism, 153, 154
Intersubstitutivity rules, 6, 75, 76, 94, 128, 136
Irrationalist's Paradox, 80, 81, 83–85, 87, 88, 173

K
Kant, Immanuel, 32, 74
Kaplan, David, 148, 150, 151
Kripke, Saul, 133, 142, 163, 168–170

L
Laplacean determinism, 41
See also Laplaceanism
Laplaceanism, 41, 42
Law of excluded middle (LEM), 1, 2, 4, 5, 7, 8, 11–14, 19, 20, 28, 43, 47, 48, 54, 59, 82, 83, 85, 100, 120, 123–127, 129, 130, 135, 137, 141, 143, 155, 157, 160, 167, 171, 173, 180
Law of Noncontradiction (LNC)
rationality LNC (RLNC), 6, 7
Truth Value LNC, 7
Lewis, David, 39
Liar Paradox
contingent Liar, 129
and Revenge, 10, 22, 71, 136, 171
Simple Liar, 10, 76, 171
Strengthened Liar, 10, 21, 76, 134, 162, 171, 173
Linguistic character
as opposed to propositional content, 150, 151
Logical harmony, 154, 155
Logical pluralism, 14, 118–120

Logic of Paradox (LP)
LPm, 118–120
Łukasiewicz, Jan, 100

M
Maudlin, Tim, 137–140, 155
Meanings, 2, 4, 76, 95, 112, 127, 132, 148, 150, 152–155, 160–162, 176, 177
Meinongianism
about reference failure, 148
Metalanguage, 137, 139
Modal operators
O (obligation), 32
P (permission), 32
Modus Ponens, 54, 56, 57, 66, 69, 71–73, 89, 109, 112, 126, 127, 155, 172–174
Modus Tollens, 53–57, 59, 60, 62, 88, 92, 109, 112
Monotonic logic, 119, 120
Monty Hall Problem, 37, 40, 131

N
Naïve set theory, 7, 15, 18, 22, 24, 35, 37, 105, 165, 170, 171
Negation
as exhaustive, 11
NOT, 53, 71, 82, 89, 98, 112, 124, 127, 135, 160, 177, 179
and rejection, 98
Newton, Isaac, 100
Nixon paradox, 168, 169
NoDilemmas, 94
Nominalism, 70, 165
Non-classical
axioms, 133
logics, 73, 130
Non-existence solution, 145
Meinongianism about, 148
Russelleanism about, 148
Non-monotonic logic, 119
Non-Triviality Project (NTP), 141, 160
NOT
as a kind of super-negation, 123

O
Objection From the Tables, 176, 177
Object language, 57, 137, 139
Ockham, William of, 105, 106
Ought implies can
as a principle, 74

P
Paracompleteness, 7–9, 11–13, 55, 82, 83, 95, 96, 125–146, 172
Paraconsistency
paraconsistent solution to paradox, 8
Paradoxes
Barber, 169
Burali-Forti, 22, 25
Cards, 167
Curry's, 50, 51, 63, 65–101, 126, 154, 171–174
Grelling's, 145
and the Inclosure Schema, 25, 69, 70
Irrationalist's, 83, 87
Liars, 5, 12, 19–22, 24, 25, 48, 68–72, 83, 84, 87, 121, 140, 141, 166, 174
of motion and change, 18–20, 31, 38–43, 48
of naïve set theory, 21, 31, 34–38, 70
Nixon, 168, 169
Russell's, 21
of self-reference, 20, 69, 87, 108
of vagueness, 130
Yablo's, 66, 161
Zeno's Arrow, 40, 41
Parsimony, 105, 106
Perceptual illusions, 115–118
Permissibility, 138–140
Platonism, 36
Power set theorem, 165
Prec, 150
See also CompoundProp
Precisification, 150, 151
Priest, G., 4–6, 10, 11, 13–15, 17–29, 31–80, 82–90, 92–96, 98–100, 103–110, 112–122, 126, 127, 131, 133, 134, 138, 155, 161, 162, 165, 171–173
Principle of Uniform Solution (PUS), 21, 69–71, 83, 84, 87, 173–174
Principle R, 61–99, 103, 104, 119–121, 173
Prior, A.N., 154
Probability calculus, 9, 22, 121, 128, 129, 131–133, 135, 175
Probability theory, 9, 15, 103–124, 133, 175
Prop, 148, 150
See also CompoundProp
Propositional content
as opposed to linguistic character, 150

Propositional logic, 147, 148, 150, 161
Proto-paracompleteness, 9, 135
Psychologism, 99, 100

Q
Quantum logic, 174–177
Quantum mechanics, 7, 15, 42, 158, 174, 180
Quasi-validity, 103–111
Quine, W.V.O., 35, 47, 104, 105, 153

R
Ramsey, Frank, 156, 157
Rat Dung Principle, 163, 164
Reductio Ad Absurdum (RAA)
  RAA(B), 88–92, 99, 104
  RAA(N), 88–92, 99, 104
  RAA(P), 104
Reject
  Reject(D), 136
  Reject(DD), 2–8, 11, 12, 19, 23, 27, 28, 32, 36, 39, 49, 52, 53, 55, 56, 59, 62, 68, 73, 75, 77, 79, 82, 85, 89, 91, 92, 94, 95, 97–99, 104, 106, 107, 115, 120, 124, 126–128, 131, 134–136, 139, 140, 153, 155, 165, 171–174, 178, 179
  Reject(F), 75, 94, 134, 135
  Reject(U) (R(U)), 91
Rejectability, 64, 75, 77, 79, 80, 83, 91, 92, 95, 97, 121
Rejection, 3, 9, 11, 12, 22, 25, 37, 64–101, 103–124, 127, 131–135, 173, 174
Relevance logics, 60, 61, 125
Revenge, 10, 15, 22, 66, 71, 123, 135–137, 140, 142, 155, 162, 164, 165, 170, 171
Russell, Bertrand, 7, 8, 18, 20–25, 27, 28, 34, 36, 39–43, 87, 108, 144–146, 164, 169, 170, 174, 178, 179
Russelleanism
  about change, 43
  about reference failure, 148
Russell Set, 20, 23, 24, 27, 28, 34, 36, 38, 43, 170
Russell's Paradox, 8, 18, 20–25, 36, 38, 87, 108, 144–146, 164, 169, 174
Russell's Paradox for predicates, *see* Grelling's paradox

S
Sartre, Jean-Pual, 112, 116
Semantic completeness, 67, 68
Sentential logic
  as distinct from propositional logic, 147, 150
Set theories, 8, 20, 21, 23, 24, 34, 36, 165
S-View, 2, 21, 93, 123, 137, 156, 165, 172, 179

T
Tarski, Alfred
  and hierarchies of truth operators, 136–140
Tonk
  super-tonk, 155
  tonk problem, 154, 155, 160
  and transitivity, 154, 155
Transfinite sets
  cardinality of, 37, 40
Trivialism, 68, 69, 80, 122, 126
Trolley Problem, 89
Truth
  as coherence, 151
  as correspondence, 2
  deflationism about, 2, 76, 152
  determinate truths, 140, 142, 143
  disquotationalism about, 15, 162
  fictionalism about, 156
  as a property of truth-bearers, 160, 163
  redundancy theories of, 156
Truth conditions, 51, 62, 66, 67, 71, 126, 142
Truth operators, 13, 25, 33, 76, 93, 98, 101, 128, 130, 136–140, 146, 155–159, 163, 167, 168
Truth predicate, 13, 67
Truth preservation
  as definition of validity, 58
  of inference rules, 14, 50, 51
Truth-Teller sentences, 10, 156, 171
Truth value gaps, 9–12, 19, 75, 82, 94, 95, 135
Truth value gluts
  Type-1 *vs*. Type-2 gluts, 107
T-Schema
  T-IN, 137
  T-OUT, 137

U
Unfalsity, 131
Unrestricted comprehension axiom, 7, 20, 21, 37, 170
Untruth
  as distinct from falsehood, 96
Utilitarianism, 55, 89, 90

V
Vagueness
of predicates, 178
Validity Defined Disjunctively (VDD), 58
Validity predicate, 67, 171
van Fraassen, Bas C., 163
Verificationism, 148

W
Weir, Alan, 57–58, 67, 76
Wittgenstein, Ludwig
on contradiction, 169

Y
Yablo's Paradox, 66, 161
Yablo, Stephen, 66, 161, 162

Z
Zeno, 38
Zeno's Arrow Paradox, 40, 41
Zermelo-Fraenkel set theory (ZF)
ZFC (with the Axiom of Choice), 23