

Vienna Circle Institute Library

Joseph Bentley

Logical Empiricism and Naturalism

Neurath and Carnap's Metatheory
of Science



Vienna
Circle
Institute
Library



Springer

Vienna Circle Institute Library

Volume 8

Series Editors

Esther Heinrich-Ramharter, Department of Philosophy and Institute Vienna Circle, University of Vienna, Wien, Austria

Martin Kusch, Department of Philosophy and Institute Vienna Circle, University of Vienna, Wien, Austria

Georg Schiemer, Department of Philosophy and Institute Vienna Circle, University of Vienna, Wien, Austria

Friedrich Stadler, Institute Vienna Circle, University of Vienna, and Vienna Circle Society, Vienna, Austria

Advisory Editorial Board

Martin Carrier, University of Bielefeld, Germany

Nancy Cartwright, Durham University, UK

Richard Creath, Arizona State University, USA

Massimo Ferrari, University of Torino, Italy

Michael Friedman, Stanford University, USA

Maria Carla Galavotti, University of Bologna, Italy

Peter Galison, Harvard University, USA

Malachi Hacoen, Duke University, USA

Rainer Hegselmann, University of Bayreuth, Germany

Michael Heidelberger, University of Tübingen, Germany

Don Howard, University of Notre Dame, USA

Paul Hoyningen-Huene, University of Hanover, Germany

Clemens Jabloner, Hans-Kelsen-Institut, Vienna, Austria

Anne J. Kox, University of Amsterdam, The Netherlands

James G. Lennox, University of Pittsburgh, USA

Thomas Mormann, University of Donostia/San Sebastián, Spain

Edgar Morscher, University of Salzburg, Austria

Kevin Mulligan, Université de Genève, Switzerland

Elisabeth Nemeth, University of Vienna, Austria

Julian Nida-Rümelin, University of Munich, Germany

Ilkka Niiniluoto, University of Helsinki, Finland

Otto Pfersmann, Université Paris I Panthéon – Sorbonne, France

Miklós Rédei, London School of Economics, UK

Alan Richardson, University of British Columbia, Canada

Gerhard Schurz, University of Düsseldorf, Germany

Hans Sluga, University of California at Berkeley, USA

Elliott Sober, University of Wisconsin, USA
Antonia Soulez, Université de Paris 8, France
Wolfgang Spohn, University of Konstanz, Germany
Michael Stöltzner, University of South Carolina, Columbia, USA
Thomas E. Uebel, University of Manchester, UK
Pierre Wagner, Université de Paris 1, Sorbonne, France
C. Kenneth Waters, University of Calgary, Canada
Gereon Wolters, University of Konstanz, Germany
Anton Zeilinger, Austrian Academy of Sciences, Austria

Honorary Editors

Wilhelm K. Essler, Frankfurt/M., Germany
Gerald Holton, Harvard University, USA
Allan S. Janik, Innsbruck, Austria
Andreas Kamlah, Osnabrück, Germany
Eckehart Köhler, Munich, Germany
Juha Manninen, Helsinki, Finland
Erhard Oeser, Wien, Austria
Peter Schuster, University of Vienna, Austria
Jan Šebestík, Paris, France
Karl Sigmund, University of Vienna, Austria
Christian Thiel, Erlangen, Germany
Paul Weingartner, University of Salzburg, Austria
Jan Woleński, Jagiellonian University, Poland

Editorial Work/Production

Zarah Weiss

Editorial Address

Wiener Kreis Gesellschaft
Universitätscampus, Hof 1, Eingang 1.2
Spitalgasse 2-4, A-1090 Wien, Austria
Tel.: +431/4277 46504 (international) or 01/4277 46504 (national)
Email: vcs@univie.ac.at
Homepage: <https://vcs.univie.ac.at/>

This peer-reviewed series includes monographs and edited volumes, which complement the format of the related series “Vienna Circle Institute Yearbook” and “Veröffentlichungen des Institut Wiener Kreis”, both published with Springer Nature. The books mainly deal with individual members of the Vienna Circle, the entire Schlick-Circle as a collective, its adherents and critics, as well as with related topics of Logical Empiricism and its periphery in historical and philosophical perspective. Specifically, they are based on so far unpublished primary sources and feature also forgotten and marginalized issues as significant contributions to the most recent research in these fields.

Joseph Bentley

Logical Empiricism and Naturalism

Neurath and Carnap's Metatheory of Science



Joseph Bentley
University of Manchester
Manchester, UK

ISSN 1571-3083

Vienna Circle Institute Library

ISBN 978-3-031-29327-6

ISBN 978-3-031-29328-3 (eBook)

<https://doi.org/10.1007/978-3-031-29328-3>

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

This work is subject to copyright. All rights are solely and exclusively licensed 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

Acknowledgements

The present work began life as my Doctoral thesis. I am indebted to the University of Manchester for the PhD Studentship that made my research possible, and to my supervisors Frederique Janssen-Lauret and Thomas Uebel for their support and guidance throughout.

In the transition from doctoral thesis to monograph, I benefited from a number of helpful comments and suggestions. I would like to thank Elisabeth Nemeth, John O’Neil, Friedrich Stadler and two anonymous reviewers for their insights.

Special thanks are due to Thomas Uebel who has guided and sharpened my thinking about the present topics over the past 5 years through innumerable conversations, suggestions and criticisms, and who has read and re-read incarnations and excerpts of the present work more times than I care to remember. Whilst any errors that remain are my own, the rest has benefited immeasurably from his insight.

Contents

1	Naturalism and the Vienna Circle	1
1.1	What Is Epistemological Naturalism?	1
1.1.1	What Epistemological Naturalism Opposes	2
1.1.2	Distinctions Within Epistemological Naturalism	3
1.1.3	Epistemological Commitments of Replacement and Cooperative Naturalism	4
1.2	Naturalism in the Vienna Circle	9
1.2.1	The Wings of the Vienna Circle	9
1.2.2	Schlick and the Right Wing	11
1.2.3	Neurath's Naturalism	15
1.3	Carnap and Naturalism	26
1.3.1	Carnap's Convergence with Neurath	27
1.3.2	Carnap's Approval of Neurath's Naturalism	30
1.4	The Cooperation of Neurath and Carnap: The Bipartite Metatheory	32
2	Neurath's Epistemology of Science	37
2.1	Neurath's Empiricism	37
2.1.1	Empiricism and Verificationism	37
2.1.2	Science and Pseudo-Rationalism	40
2.2	Language, Knowledge, and Science	42
2.2.1	Science as a Social Instrument	42
2.2.2	The Indispensability of Public Language	43
2.2.3	Physicalism	44
2.2.4	The Encyclopedia and Encyclopedism	47
2.3	Decisionism	49
2.3.1	Underdetermination and Decisions	49
2.3.2	Auxiliary Motives and Decision Procedures	51
2.3.3	Auxiliary Motives, Values, and Theoretical Virtues	53
2.3.4	Decisions and Theory-Choice	55

2.4	Methodological Constructivism and the Behaviouristics of Science	59
2.4.1	Methodological Constructivism	59
2.4.2	Behaviouristics of Scholars	62
2.5	Neurath's Epistemic Agent	64
3	Neurath's Conception of Protocol Statements	67
3.1	The Protocol Sentence Debates	67
3.2	What Is a Protocol Statement?	69
3.3	The Structure of Neurath's Protocol Statements	76
3.4	The Four Conditions for Testimony Acceptance	82
3.5	The Pragmatic Condition	88
3.6	Utilising Testimony: Extraction	91
3.7	Utilising Testimony: Abduction	93
3.8	Conclusion	95
4	Carnapian Explication	97
4.1	Explication and the Bipartite Metatheory	97
4.2	Carnapian Explication	98
4.2.1	Explication: Process and Goals	98
4.2.2	Criteria for Explication	102
4.2.3	Explication as Naturalism	104
4.3	Carnap's Critics	105
5	Quine, Carnap, and Analyticity	113
5.1	The Development of Quine's Thinking About Analyticity	113
5.1.1	Initial Enthusiasm	114
5.1.2	Growing Concerns	115
5.1.3	Quine's Break with Carnap	116
5.2	What Is Carnap's Notion of Analyticity?	119
5.3	Which of Quine's Arguments Pose a Problem For Carnap?	122
5.3.1	"Analyticity" Is Unclear	122
5.3.2	Natural vs Formal Languages	126
5.3.3	Quine's Observation	126
5.3.4	Factual vs Non-factual Truth	128
5.3.5	Holism, Revisability and Analyticity	132
5.3.6	Changing Beliefs vs Changing Language	136
5.3.7	The Significance of the Analytic/Synthetic Distinction	141
5.3.8	Semantics and Epistemology	145
5.4	Divergent Conceptions of Empiricism?	146
5.5	Conclusion	149
6	Challenges to the Bipartite Metatheory Interpretation	151
6.1	Analyticity	151
6.2	Ballungen and Explication	152
6.3	Conceptions of Protocol Statements	154
6.3.1	Carnap's Changing Conception	155
6.3.2	Mature Positions, Apparent Disagreement	157

6.3.3	Carnap's Formal Objection	158
6.3.4	Carnap's Context Objection	159
6.4	Physicalism and Private Languages	161
6.5	Truth and Semantics	163
6.5.1	Carnap on Truth and Semantics	164
6.5.2	Neurath on Truth	166
6.5.3	Deflationary Truth	169
6.5.4	Mormann's Defence of Neurath	171
6.6	Conceptions of Unity	174
6.6.1	Linguistic Unity of Science	175
6.6.2	Carnap and Neurath on the Unity of Science	176
6.6.3	Different Conceptions or Difference of Emphasis?	178
6.7	Conclusions	180
7	Bipartite Metatheory in Application	181
7.1	The Bipartite Dialectic	181
7.2	Decisionism, Conventionalism and Constructivism	185
7.2.1	The Pragmatic and the Formal	185
7.2.2	Methodological Constructivism	188
7.3	Self-Reflexivity and Self-Direction	190
7.3.1	Self-Reflexivity and Pseudo-Rationalism	190
7.3.2	Neurath's Anti-Totalitarianism: Mosaics and Orchestration	192
7.4	Conclusion	197
	References	199

Introduction

1 The Historiography of Logical Empiricism

By the end of the 1960s, the heyday of Logical Empiricism was over.¹ As early as 1967, Passmore declared that “Logical positivism, then, is dead, or as dead as a philosophical movement ever becomes” (Passmore 1967, 529). The Vienna Circle in particular served as an embarrassment, a failure of both style and substance, a cautionary example of philosophical dogmatism, rigidity and excess. Magee called Logical Empiricism “a ready-to-hand instrument of intellectual terrorism” (Magee 1997, 55). Rorty’s claim captured a widespread attitude: “Most of us philosophy professors now look back on logical positivism with some embarrassment, as one looks back on one’s own loutishness as a teenager” (Rorty 1998, 33).

According to this picture, the Circle was noteworthy less for its philosophical ingenuity, and more for “the aggressive and even arrogant way in which those doctrines were propounded to the world” (Hanfling 1996, 193). These doctrines were typically interpreted as a logically modernised refurbishing of Humean empiricism. According to the then-dominant narrative, the Vienna Circle appended the modern logic of Frege and Russell, and most importantly Wittgenstein’s *Tractatus* conception of the tautologousness of logic, to an orthodox foundationalist empiricist epistemology. In their naïve picture, science was treated as simply a system of statements constructed upon the firm basis of sense-data, to be analysed in abstraction from any real-world social or historical context. The most important commitment of the Vienna Circle was a rejection of metaphysics, captured by the Verification

¹ The names “Logical Empiricism” and “Logical Positivism” will be used interchangeably since no distinction can be drawn systematically without being drawn inaccurately. Typically, if a distinction is drawn, it is geographical (“Logical Positivism” being specific to Vienna, and “Logical Empiricism” including the wider European movement) or historical (with “Logical Positivism” covering the radical pre-war period, and “Logical Empiricism” the more moderate post-war period). Both distinctions, whilst not entirely unfounded, are necessarily imprecise. Both are also largely ahistorical, as there was no agreement on or adherence to these labels amongst the participants of the movement themselves. See Uebel (2013b).

Principle; that a sentence incapable of verification is literally meaningless. In the commitment to verificationism lay the virulence of their rejection of metaphysics. But it proved unsustainable and ultimately self-defeating.

Fortunately for modern readers, and the logical empiricists themselves, a flourishing of scholarship on the Vienna Circle and its individual members over the past 40 years has done much to demolish and move beyond this caricature of logical empiricism and the accompanying historical narrative of unmitigated philosophical failure. Specialist scholarship has provided detailed studies of the Vienna Circle and its members, from both a historical and philosophical perspective. The two most important trends within this scholarship (at least for the purposes of the current work) are the continental rediscovery of Neurath and Schlick beginning in the 1980s, and the anglophone reappraisal and reinterpretation of Carnap's work developing in the 1990s.² Whilst flippant dismissals of logical empiricism are still far from rare (and left unchallenged in a way that dismissals of any other philosophical tradition typically are not), the consensus on the worth and ongoing relevance of their work has dramatically shifted.³ This rediscovery has served to rescue the Vienna Circle from the dustheap of history.

This reappraisal of the Vienna Circle fundamentally upended the conventional narrative on a variety of fronts. Firstly, it has been unquestionably demonstrated that logical empiricism was not the monolithic project it was frequently portrayed as. There was no party line or definitive set of philosophical commitments that united the movement, over and above the goal of moving philosophy beyond trans-empirical "metaphysics"; as will be demonstrated in subsequent chapters, there was a huge degree of diversity, change and internal discussion over both specific doctrinal issues and the overall aims of the logical empiricist project itself.⁴ Most obviously, empiricist foundationalism and reduction to sense-data were categorically not points of universal agreement; they were topics of fierce internal debate, as discussed in Chapt. 2.⁵ Even the verification principle, often treated as the definitive logical empiricist commitment, was a matter of intense debate, not (explicitly at least) avowed by all members, and certainly not in the same form.⁶

² For Neurath and Schlick, see for example Haller (1982), Stadler (1982/1991), Zolo (1989, 1990), Uebel (1991a, 1992), and Nemeth and Stadler (1996). For Carnap, see for example Creath (1991), Stein (1992), Richardson (1997b), and Friedman (1999).

³ For a particularly egregious example as recently as 2012, an introductory text with the title *A Brief History of Analytic Philosophy* concludes its chapter on Logical Positivism: "Ridiculing the doctrines of the logical positivists... is easy and fun" (Schwartz 2012, 69).

⁴ One of the clearest proofs of this diversity is the chart of answers to a survey compiled by Rose Rand. Questions were put to Schlick, Waismann, Carnap, Neurath, Hahn and Kaufmann, and collated into a chart so their differences in opinion could be easily recognised. Reprinted in (Stadler 2015, 143–145).

⁵ One of the most important factors in this realisation has been the reinterpretation of Carnap's *Aufbau*, as neo-Kantian in inspiration (Richardson 1990; Friedman 1999) or even naturalist (Pincock 2005).

⁶ This too will be discussed in more detail in Chapt. 2.

It is only with an understanding of the dynamics of these debates within the Circle that the most important philosophical contributions of the members of the Circle can be fully appreciated.⁷ There was not just one logical empiricism, but a variety of potential logical empiricisms.

It has also been demonstrated that the traditional narrative of the Vienna Circle's failure requires a blindness to historical context. The logical empiricists were not simply unmasked by famous critics like Karl Popper, Thomas Kuhn and W.V.O. Quine. Historical factors played as much of a role in their decline as any philosophical flaws. The deaths of Hans Hahn and Moritz Schlick, in 1934 and 1936, respectively, were huge losses. Although it briefly limped on without Schlick at the helm, by 1938 the Circle was dispersed, its members emigrating, primarily to the UK and USA, in the face of the rising tide of fascism. Logical empiricism was thus deprived of a central locus and the face-to-face interactions that had been so fruitful for the Circle in its early days. Neurath, the great organiser of the Unified Science movement, died in 1945. The cohesion lost in this period was never regained. And the intellectual climate of the post-war West, North America in particular, proved far less receptive to the socialist-adjacent work of the Vienna Circle's left-wing.⁸ This coincided with logical empiricism's transition from insurgent intellectual movement to institutionalised philosophical orthodoxy, in the process losing much of its radicalism and dynamism. Of the figures who became increasingly prominent in the logical empiricism of the post-war years, many, such as Ernest Nagel, were never members of the Circle, whilst those like Herbert Feigl and Friedrich Waismann who had been participants in the inter-war years were more closely associated with the Schlick-wing than the left-wing or, like Hans Reichenbach and Carl Hempel, had a different background in Berlin. The consequence was an important difference in the thinking and goals of pre- and post-war logical empiricism.

These historical factors are also key to explaining how the received view came to dominate the image of logical empiricism in the popular philosophical imagination. In the 1950s and 1960s, the logical empiricism of the Vienna Circle was easily equated with the much more familiar, accessible and prominent logical empiricism of post-war North America. This accessibility was in part linguistic; many of the Vienna Circle's works were only available in German. This was exacerbated by the failure of surviving members of the Circle to tell and thereby shape their own story. No comprehensive history of the Vienna Circle was ever told by the participants themselves.⁹ The Vienna Circle lost a distinctive identity of its own. Instead, the two

⁷ This is demonstrated in Uebel's extensive reconstruction on and interpretation of the Circle's protocol sentence debates, and the picture of Neurath's epistemology that emerges from it (Uebel 2007d).

⁸ See Howard (2003) and Reisch (2005).

⁹ Some brief historical pictures were produced, but most lacked an overarching narrative. Kraft's account of the Vienna Circle was in line with the emerging consensus, stressing the influence of Wittgenstein and Carnap's *Aufbau* project, but minimising the cultural and political dimensions of the Circle's work (Kraft 1950/1953). Carnap's intellectual autobiography in his *Library of Living Philosophers* volume also gives some historical insights (Carnap 1963d), as do Menger's

most important conduits for what would become cemented as the traditional story of the Vienna Circle came from outside the Circle: A.J. Ayer and W.V.O. Quine.

Both Ayer and Quine had credibility as interpreters of the Vienna Circle due to their close personal and philosophical ties to the Circle and its members.¹⁰ But neither were as reliable as they appeared. The main inheritance from Ayer has been his framing of the Vienna Circle as engaged in a narrowly epistemological project of foundationalist empiricism by strongly connecting logical empiricism to the legacy of British empiricism.¹¹ No sense of logical empiricism as a broader modernist Enlightenment project extending beyond the narrow confines of academic philosophy survived in Ayer's presentation.¹² This framing of the logical empiricist project was carried on by Quine.¹³ Consequently, after its heyday, logical empiricism was in the popular imagination relegated to "restat[ing] the foundationalist epistemology of British empiricism" (Rorty 1997, 1).

But if this inaccurate historical narrative has now been dismantled and superseded by modern scholarship, what importance does it have for the contemporary reader? The answer is that despite this growing body of Vienna Circle scholarship, there is one specific field of contemporary philosophy in which the lingering influence of the once-dominant received view has persisted. In contemporary discussions of naturalism (the contested meaning of which will be discussed in detail below), the caricature of logical empiricism is frequently taken for granted. Beyond functioning as a cautionary tale, the Vienna Circle's primary role in the naturalist story was as the orthodoxy against which Quine rebelled.¹⁴ Quine's argument for and narrative of the development of his naturalism frames the logical empiricist project as the capstone of traditional foundationalist empiricism. His argument both drew on and helped to cement the received view of logical empiricism. And as the man who made naturalism a viable position within the philosophical mainstream, Quine casts a long shadow over its subsequent conceptualisation. His influence is

posthumously published memoirs (Menger 1994). Only Frank's historical introduction to his collection of essays (Frank 1949a) communicated a narrative of the Circle's trajectory, but it has been somewhat neglected.

¹⁰ Ayer in particular exercised an extraordinary influence over how English speaking audiences were introduced to the ideas of the Vienna Circle. His best-selling *Language, Truth and Logic* was treated as the anglophone bible of logical positivism. And when in 1959 many of the early writings of the Vienna Circle were first published in English, Ayer served as editor and provided a long introduction (Ayer 1959).

¹¹ On Ayer's distorted history of the Vienna Circle, see Vrahimis (2021) and Tuboly (2021).

¹² See for example the connections between the Vienna Circle (especially Neurath and Carnap) and the modernist Bauhaus project (Galison 1990, Dahms 2004). On the Vienna Circle's politically engaged philosophy of science, see (Romizi 2012; Uebel 2020b).

¹³ In the same book cited above, Schwartz says that the version of logical empiricism he presents is the "standard view... derived from Ayer [and] Quine" (Schwartz 2012, 59).

¹⁴ This framing was present from the outset. Quine's first presentation of what became Epistemology Naturalised at a lecture in Vienna on 9 September 1968 was introduced: "I will now reflect on the differences between my present point of view and the Old Viennese philosophy" (Trans in Verhaegh 2018, 160 n.37).

so extensive that his route to naturalism has been widely accepted as *the* route to naturalism.

Consequently, amongst contemporary naturalists, it is frequently taken for granted that the rejection of logical empiricism is part and parcel of the adoption of naturalism. They consider Quine's rejection of the analytic/synthetic distinction as having cleared the ground for naturalism. This is oft repeated:¹⁵

it was "Two Dogmas of Empiricism"... which more than any other work made philosophical space for naturalism as a viable alternative to positivism... Once the demarcation between the analytic and the synthetic is surrendered, the demarcation between scientific theory and philosophical analysis is extirpated (Rosenberg 1996, 2)

This has led to the assumption that rejecting the analytic/synthetic distinction is the conceptual basis of, or even equivalent to, naturalism: "Naturalism in epistemology can be characterised negatively by its eschewal of any notions of analytic or a priori truths" (Roth 2003, 273). This same framing is found in the work of many naturalist-inclined feminist epistemologists of science.¹⁶

Logical empiricism has therefore been widely accepted by naturalists to be fundamentally incompatible with their methods and goals. What this obfuscates or ignores is that Quine's naturalism is only one among many. As recent work has made increasingly clear, and I will be at pains to stress again below, the left-wing of the Vienna Circle provided an earlier alternative conception. The logical empiricist that many understand Quine to be arguing against has been revealed to be a straw man. It is therefore necessary for the arguments that relied on this false narrative to be reappraised too. And, it will be argued, the key consequence of this reappraisal is the reconciliation of (a version of) logical empiricism with naturalism. What this requires is the realisation that the issues of epistemological naturalism and the analytic/synthetic distinction are not coextensive; that the one does not necessarily determine the other. Naturalism may be incompatible with some versions of logical empiricism, but not all.

The argument that follows advocates the naturalistic interpretation of the Vienna Circle's left-wing. This interpretation is neither entirely novel nor entirely uncontroversial. Recognising Neurath as an exponent of epistemological naturalism is now common. But it is all too easy for contemporary readers to see Neurath's naturalism through the lens of Quine. Consequently, in Chap. 1, time is taken to allow Neurath's naturalism to speak for itself. The naturalistic reading of Carnap is more contentious, but also not entirely unprecedented. Kitchener for example highlights that Carnap's epistemology contains naturalistic components (Kitchener 2004, 42), while Christopher Pincock urges a reserved, naturalistic reading of the *Aufbau* (Pincock 2005). Carus has highlighted Carnap's agreement with Quine about much of his naturalised epistemology (Carus 2007a, 267), and Lutz has

¹⁵ For other examples, see Kitcher (1992, 69–71), Kornblith (1995, 240), and Nelson (1996, 312).

¹⁶ See for example Nelson (1990, 90–91), Antony (1993, 202), and Campbell (1998, 33).

also argued that Carnap's conceptual engineering ought to be seen as a form of naturalised epistemology (Lutz 2020).¹⁷ But in these cases, naturalism has been used as a lens through which to understand the Vienna Circle. What I hope to show is that the Vienna Circle also provide a potential lens through which to understand the prospects and potential of naturalism.

I argue that the version of naturalism Neurath and Carnap developed is an appealing one, with doctrines and methods that modern naturalists ought to avail themselves of. To do so means recognising not only the naturalism of Neurath and Carnap in isolation, but the potential for the naturalistic project to which they jointly contributed. In demonstrating the compatibility of Neurath and Carnap's work, it is demonstrated how formal and logical techniques can be complementary to the sort of empirically informed epistemological work that is typically associated with contemporary naturalism. The joint naturalistic project shows how the empirical study of science can be pursued in conjunction with an explicationist methodology to supply valuable methodological and conceptual insights about the sciences.

In shaping this argument and interpretation, by far the most important influence from the existing scholarship is Thomas Uebel, in particular his argument for the Bipartite Metatheory, introduced in more detail in Chap. 1.¹⁸ In brief, this interpretation argues that Neurath and Carnap's work can be combined as part of a multi-disciplinary metatheory of science to serve as a replacement for traditional armchair philosophy. Latent in this interpretation is that we must understand both Neurath and Carnap as adopting a form of epistemological naturalism. In what follows, I aim to both extend and defend this interpretation. Where the naturalistic elements of the bipartite metatheory have not previously been given centre stage, I argue that they are essential to the interpretation. I argue that the naturalist orientation is the shared foundation on which the joint project can be built. I demonstrate that not only do both Neurath and Carnap embrace naturalism, but even more significantly that the form and motivations for this naturalism are the same; a deliberate rejection of not only traditional epistemology, but the philosophical impulse that Neurath calls "pseudo-rationalism". Carnap's adoption of naturalism is facilitated most importantly by his adoption of the principle of tolerance in 1934, and this principle is itself a manifestation of the rejection of pseudo-rationalism as applied to the philosophy of logic. The rejection of logical monism in favour of a pluralistic and instrumentalist view of logic and language is a move rooted in naturalist thinking. This is what makes Carnap's subsequent formal work such a valuable, but as yet under-utilised, resource for the naturalist.

The current work, it can be seen, draws on the existing scholarship of the previous 40 years, and applies these insights to the topic of naturalism. In doing better justice to the intellectual legacies of the Vienna Circle, this scholarship has done invaluable historical and exegetical work in its own right, but we can also regard it as a first step in a constructive philosophical process. What the existing

¹⁷ Both Carus and Lutz's arguments will be discussed in greater detail in Chap. 4.

¹⁸ This is not to claim total agreement, any errors are my own additions.

scholarship proves is that many of the Vienna Circle's ideas are still of profound relevance for contemporary philosophy. The apparently negative work of disproving and dispelling is also constructive work, as it clears the intellectual space in which an account of the left-Vienna Circle's actual project can be fully developed and applied to contemporary problems. This project, here characterised as jointly Neurath and Carnap's, grounded in the rejection of pseudo-rationalism and a resulting embrace of naturalised epistemology, creates a role for the metatheory of science as an empirical and instrumentally normative project, for the study and conscious self-direction of the sciences. From beneath the rubble of the caricature of a stale, reductionist empiricism, we can rediscover and begin to reconstruct a broad, naturalistic and humanistic metatheoretical project for the creation of a self-reflexive, self-directing science, as captured in the Vienna Circle's manifesto:

We witness the spirit of the scientific world-conception penetrating in growing measure the forms of personal and public life, in education, upbringing, architecture, and the shaping of economic and social life according to rational principles. *The scientific world-conception serves life, and life receives it.* (Carnap et al. 1929/1973, 317–318)

The understanding of their joint project here pursued reveals the consistency of Neurath and Carnap's mature philosophical work with the Enlightenment ideals that animated the Vienna Circle in its earlier stages. Through the self-reflexivity that comes with the rejection of pseudo-rationalism, naturalism, explicationism and the humanist goals of the enlightenment can be rendered consistent.

2 Chapter Overview

In Chap. 1, I begin with a discussion of naturalism as it relates to the Vienna Circle. This requires differentiating the various breeds of contemporary naturalism and the theoretical commitments they entail, and orienting the members of the Vienna Circle relative to them. This means distinguishing the left- and right-wings of the Vienna Circle, with only the former having embraced epistemological naturalism. It will be shown that the naturalism of the left-wing cannot simply be equated with the naturalism of Quine or other contemporary naturalists. Uebel's Bipartite Metatheory is then introduced as a naturalistic project on which Neurath and Carnap collaborated.

From this foundation, Chaps. 2 and 3 provide an in-depth exploration of Neurath's contribution to the joint project. Chapter 2 involves an account of Neurath's philosophy of science. Whilst he himself never provided a comprehensive overview of his thought, such a picture can be reconstructed due to the numerous theoretical continuities throughout his works. Chapter 3 focuses in depth on Neurath's conceptions of protocol statements, providing an analysis of his account of the process of receiving and appraising scientific observation reports as a case study of Neurath's metatheory of science.

In Chap. 4, the focus changes to Carnap, specifically his methodology of explicationism, which I argue must be understood within the wider context of the naturalistic bipartite-metatheory, and which I defend from some common criticisms from this specifically naturalist perspective. With this grounding in Carnap's project, Chap. 5 focuses on the debate between Quine and Carnap over the possibility of coherently distinguishing the analytic from the synthetic. I demonstrate that, despite the frequency of their interactions and the extensiveness of their debate, there was a failure of communication lying behind their discussions. It will be demonstrated that Quine's arguments against Carnap fail to undermine the version of analyticity that Carnap provided, and that consequently Carnap does provide a coherent and sufficient notion of analyticity for the purposes of the bipartite metatheory.

In Chap. 6, I address a variety of potential arguments for the incompatibility of Carnap and Neurath's respective philosophies. I demonstrate that such arguments typically rely on misreading a division of labour and difference of emphasis as disagreement, and can be alleviated by careful attention to the use of language and argumentative and historical context.

Finally, in Chap. 7, I demonstrate the interplay of Carnap and Neurath's projects, and their consequences for science. Here a first attempt is made to demonstrate the potential for fruitful interaction and collaboration between formal-logical and empirical work within an instrumentalist and naturalistic framework. I do not pretend to have fulfilled the promise of such a programme, or even have fully captured its potential scope. But the intentions behind this project are elaborated, and the potential direction in which it was intended to develop are sketched.

Chapter 1

Naturalism and the Vienna Circle



Despite the growth of interest in the project in the late 1980s and early 1990s, naturalized epistemology has existed for a lot longer. Its origins are typically traced to Quine's "Epistemology Naturalized" in 1969, although Alex Rosenberg argues that Ernest Nagel also made the argument for naturalism in 1956, and again in 1961, albeit within the intellectual space that had been cleared by Quine's "Two Dogmas" (Rosenberg 1996, 2). But an even earlier form of naturalized epistemology had already been developed in the 1930s by the so-called left-wing of the Vienna Circle, most obviously by Otto Neurath. This chapter explains what is meant by epistemological naturalism, and lays the foundations for demonstrating that this label is appropriate for the philosophical project pursued by the left-wing of the Vienna Circle, including Rudolf Carnap.

1.1 What Is Epistemological Naturalism?

What naturalism is, and how to identify it, remains controversial and contested.¹ Even those who claim to espouse it frequently disagree. Naturalism is more of an attitude or approach than a specific set of doctrines. According to Larry Laudan for example, on 'the intellectual road map, naturalism is to be found roughly equidistant between pragmatism and scientism' (Laudan 1990, 44). Various philosophers have still attempted to highlight the definitive aspects of the naturalist outlook. Philip Kitcher highlights 'the reintroduction of psychology into epistemology and the suspicion of the *a priori*' (Kitcher 1992, 59). Richard Feldman finds a 'revolt against armchair philosophy' and advocacy of 'the relevance of psychology to

¹ "Naturalism" in this work refers to methodological naturalism, concerned with epistemological issues surrounding the investigation of the world. Ontological naturalism, concerned with the metaphysical nature of the components of reality, is irrelevant to current concerns.

epistemology’ (Feldman 1999, 170). Alex Rosenberg argues for a tripartite commitment to the ‘repudiation of ‘first philosophy... Scientism... [and] Darwinism’ (Rosenberg 1996, 4). For Laudan, naturalism ‘denies that the theory of knowledge is synthetic a priori’ (Laudan 1990, 44). James Maffie argues that ‘Naturalists are united by a shared commitment to the continuity of epistemology and science’ (Maffie 1990a, 281). Paul Roth similarly argues that ‘naturalism asserts a normative and methodological continuity between epistemological and scientific enquiry’ (Roth 2003, 273). But none of these definitions captures every use of the term, or really explains it. These definitions give a naturalism, rather than *naturalism*. But even if there is no definitive naturalist claim or set of necessary and sufficient criteria, there must be some shared commitments to give the label “naturalist” coherence.

What is common to all versions of naturalism, as can be seen from the quotes above, is an affinity for science, and an attempt to connect philosophy more closely to it, as well as a dissatisfaction with traditional epistemology in philosophy, especially when the lack of apparent progress in philosophy is compared with the obvious successes of the sciences. Epistemological naturalism is most readily characterised by what it is not, and what it is a rejection of. So what constitutes the “traditional” epistemology that naturalists oppose?

1.1.1 What Epistemological Naturalism Opposes

Feldman highlights four key features of traditional epistemology, to which I have added a fifth (2012, 1.1):

TE1: A Priori Methodology

Epistemology is traditionally seen as a non-empirical discipline, one based on thought not investigation.

TE2: Independence of Science

Philosophy is a discipline separate from and prior to the sciences. The claims of science are irrelevant to the philosopher and the claims of philosophy are independent of the claims of the sciences.

TE3: Normativity

Traditional epistemology takes its findings to be evaluative and even prescriptive, not merely descriptive. Epistemology tells us how we should reason, not simply how we do reason.

TE4: Response to the Sceptic

Traditional epistemology takes the sceptic’s doubts seriously and considers a plausible response to them a core goal of epistemological enquiry. The most famous example is Descartes’s cogito.

TE5: Foundationalism

The aim of traditional epistemology is a theory of justification. Most justification occurs inferentially. But if an infinite regress is to be avoided, a basis of certain, secure, non-inferential knowledge to which inferential knowledge is reducible must be discovered. For rationalists, this foundation is typically a priori knowledge. For empiricists, it is typically immediate experience. The majority of traditional epistemologists since Descartes have been foundationalists.

These criteria are obviously inter-related and mutually supporting. For example, the attempt to refute the sceptic combined with the assumption of a science-independent methodology naturally leads to an a priori methodology. Similarly, foundationalism frequently coincides with an attempt to respond to the sceptic. But it is also true that they can be taken in isolation; some can be maintained while others are rejected. As such, if naturalism amounts to rejecting certain aspects of traditional epistemology, there are various combinations open to the naturalist. A naturalist may reject any or all hallmarks of traditional epistemology, as we shall see presently. It would be superfluous to pursue a complete taxonomy of all the hypothetical versions of naturalism, but it will still be helpful for what comes later to introduce some terminological distinctions.

1.1.2 Distinctions Within Epistemological Naturalism

Methodological naturalists argue for a continuity between philosophy and science. But this continuity allows of degrees. And how this continuity is understood depends on what aspect of science the naturalist sees as most important. Science is both a methodology and a body of theory. For some naturalists, it is this body of theory that is important; they want to make the findings of empirical science accessible and relevant to the discussion of epistemological problems. For them, naturalism ‘means they have permission to draw on the concepts and results of natural science’ (Hylton 2014, 152). This position is variously called ‘modest’ (Haack 1993, 336), ‘expansionist’ (Haack 1990, 111), ‘moderate’ (Goldman 1994, 1), ‘Weak’ (Stich 1993, 100) or ‘cooperative’ (Feldman 2012, §3) naturalism. In its minimal form, this position does not involve a radical break with traditional epistemology. Rather, the ‘epistemological enterprise needs appropriate help from science’ (Goldman 1999, 3). The unique philosophical method is supplemented, not supplanted. This position aims to be a plausible, and pragmatic adjustment to standard epistemology.

In this minimal form, epistemological naturalism is open to accusations of banality. As Kitcher says, ‘How could our psychological and biological capacities and limitations *fail* to be relevant to the study of human knowledge?’ (Kitcher 1992, 58). For more radical naturalists this moderate position is still too close to traditional epistemology, and consequently shares the same failings. They argue that it is the empirical methodology of science that is important. The successes of empirical science, and the lack of progress in a priori philosophy, is proof of the superiority of

the scientific method. But the recognition of the authority of the empirical method undermines the notion of a distinctively philosophical problem. If all questions are to be answered empirically, then all the answers given will be scientific ones, not philosophical ones. Consequently, philosophy as an external discipline with a unique and separate a priori methodology disappears, and what legitimate work there is for the epistemologist becomes scientific; the epistemologist ‘reflects on science from within science’ (Hylton 2014, 151). In Laudan’s terms, the naturalized epistemologist ‘takes to heart the claim that his discipline is the theory of claims as theories or hypotheses about enquiry, subject to precisely the same strategies of adjudication that we bring to bear on the assessment of theories on the assessment of theories within science or common sense’ (Laudan 1990a, 45). This position is variously called ‘scientistic’ (Haack 1993, 336), ‘revolutionary’ (Haack 1990, 112) ‘Strong’ (Stich 1993, 100) or ‘replacement’ (Feldman 2012, §2) naturalism.²

Going forward, I will use the terms replacement and cooperative naturalism for these two positions. With this distinction in mind, where do replacement and cooperative naturalists stand on the criteria of traditional epistemology given above?

1.1.3 Epistemological Commitments of Replacement and Cooperative Naturalism

TE1: A Priori Methodology

For replacement naturalists, TE1 is typically rejected. The incorporation of epistemology into science requires the abandonment of substantive a priority, the abandonment of first philosophy. Roth captures this contrast as one between methodological monism and methodological dualism (Roth 2003, 274). Strong naturalists are methodological monists; epistemological questions are pursued by the same methods as the sciences, and the answers supplied are of the same kind. There is one method; the scientific method. By contrast, moderate naturalists are methodological dualists; their naturalism allows for the coexistence of scientific methodology alongside a separate, complementary and uniquely philosophical methodology for epistemology. Whilst philosophy is no longer seen as an exclusively a priori discipline, a priori knowledge is still considered a potentially legitimate element of epistemology. Philosophy therefore has a role, the provision of a priority, distinct from but complementary to science. Goldman for example defends the a priori, not as a different type of truth, but as a different source of justification (Goldman 1999, 1–2).

² “Replacement” is often used by critics to mean naturalism that denies the normativity of epistemology. Very few (if any) naturalists actually maintain this position, so this use of the term carries a pejorative and critical, rather than simply descriptive connotation. In this work, none of these terms will be taken as implying any position on normativity.

TE2: Independence of Science

Both the cooperative and replacement naturalist rejects TE2, arguing instead that there is a continuity between science and philosophy. But this continuity can be interpreted in a number of ways. Maffie identifies six important senses of continuity (Maffie 1990a, 282–284):

1. Contextual Continuity: epistemology occurs within the context of natural science, not prior to or outside it
2. Epistemological Continuity: epistemology, like science, is an a posteriori, empirical project
3. Methodological Continuity: epistemology is conducted using the tools and methods of science
4. Analytic Continuity: analytic continuity between the languages of science and philosophy via analysis/reduction/explication of epistemological concepts in terms of descriptive concepts
5. Metaphysical Continuity: epistemic properties are reduced to/constituted by/supervenient on descriptive properties
6. Axiological Continuity: investigation of epistemic ends and norms is also a posteriori

These are not necessarily rival conceptions, nor are they mutually incompatible. Most naturalists understand continuity to incorporate more than one of these senses. However, they are potentially maintainable in isolation too.

The crucial divide within naturalism, according to Maffie, is between what he terms limited and unlimited naturalism. What divides them is the scope of epistemological, contextual, and methodological continuity. Unlimited naturalism extends continuity to the meta-epistemological level; naturalism ‘all the way up’ (Maffie 1990a, 287). Questions concerning methodological norms, the meaning of epistemic expressions and the goals of enquiry are conducted according to scientific methods, within science. Limited naturalism, however, only allows this continuity at the level of epistemological enquiry (Maffie 1990a, 288). Evaluation and prescription of belief can be internalised by science, but meta-epistemological questions must be answered outside science by uniquely philosophical methods. The division between cooperative and replacement naturalism somewhat overlaps with the distinction between limited and unlimited naturalists. Whilst all replacement naturalists must be unlimited naturalists, some cooperative naturalists would even shy away from limited naturalism as defined by Maffie. Specifically, those who want to maintain the relevance of a priori methodology at the epistemological level are putting even greater restrictions on the degree of continuity between science and philosophy, but can do so while emphasising contextual continuity. The weakest form of cooperative naturalism simply argues for contextual continuity, whilst rejecting either methodological or epistemological continuity.

TE3: Normativity

The topic of normativity in naturalism is a controversial one. Many critics of naturalized epistemology have argued that it cannot preserve the normative element

of epistemological enquiry. Kim famously argues that naturalists, following Quine, ‘set aside the entire framework of justification-centred epistemology’ and ‘put in its place a purely descriptive, causal-nomological science of human cognition’ (Kim 1988, 388). He claims that if epistemology is reduced to science, then it becomes a solely descriptive discipline. And famously, one cannot derive an ought from an is; this is to commit the naturalistic fallacy. Naturalism provides an epistemology in name only. To call what is left “epistemology” is (intentionally or not) a linguistic sleight of hand to disguising a change of subject.

Yet many naturalists make what Uebel calls the ‘legitimation claim’; that naturalism is equipped to make prescriptive or evaluative, not merely descriptive, claims (Uebel 1995, 634). For cooperative naturalists, this is simple. Kim’s challenge does not apply to the cooperative naturalist; they retain a priori knowledge in epistemology, and therefore the supposed source of normativity. As Kim admits, he does not ‘want to quarrel with Quine about the interest or importance of the psychological study of how our sensory input causes our epistemic output’ (Kim 1988, 391). Some cooperative naturalists even use such arguments against the replacement naturalist.

What response is available to the replacement naturalist? How can empirical science arrive at normativity, which is usually seen as empirically inaccessible; we do not experience justification or warrant. A rare few bite the bullet. Traditional normative epistemology has failed, and the closest thing we can come to it is a new purely descriptive project. If this is a disappointment to traditional epistemologists, it is because they are pursuing impossible goals. They need to adjust their expectations accordingly. Maffie calls this Eliminativism (Maffie 1990a, 285). Eliminativism achieves continuity of science and philosophy, but at the expense of normativity. This is a position far more common amongst sociologists of knowledge than philosophers.

But most replacement naturalists, including Quine, maintain that naturalised epistemology is a normative endeavour.³ And they don’t commit the naturalistic fallacy. They simply reject the received wisdom that normativity can only be derived a priori, arguing that we must understand the norms, rules and proposals of epistemology as hypothetical imperatives. Normative claims are interpreted as containing an implicit “if... then” and are thereby naturalised by being placed in the context of, and determined by, human goals and purposes. Maffie maintains that ‘epistemology is normative only within the framework of instrumental reason’ (Maffie 1990b, 333). Kitcher suggests that normativity results from the aim of enquiry being ‘*significant truth*’, where significance is determined by ‘practical concerns, or... epistemic interests’ (Kitcher 1992, 102). Kornblith goes even further, arguing that we can only arrive at normativity empirically, and that a priori prescriptions are irrelevant to the actualities of giving epistemic advice (Kornblith 1995, 20–1). Quine was less specific, but outlined a broadly similar approach:

³ Despite this, Quine is still frequently interpreted as maintaining eliminativism (Kim 1988, Laudan 1990, 45, Maffie 1990a, 285, Goldman 1994, 305).

For me, epistemology is a branch of engineering. It is the technology of truth-seeking, or, in a more cautiously epistemological term, prediction... The normative here, as elsewhere in engineering, becomes descriptive when the terminal parameter is expressed (Quine 1987, 664–665)

However, science cannot be narrowly explained as aiming simply at the provision of predictions. Quine's view is more nuanced:

prediction is not the main purpose of the science game. It is what decides the game, like runs and outs in baseball. It is occasionally the purpose, and in primitive times it gave primitive science its survival value. But nowadays the overwhelming purposes of the science game are technology and understanding (Quine 1992, 20)

Prediction is not the goal of science, explicitly or implicitly, but establishes the boundaries within which science exists.

There is not space to evaluate the successes or failures of the specific versions of or arguments for the various justifications of the legitimization claim. But this summary is sufficient to provide an overview of some the available positions for naturalists with regards to normativity.

TE4: Response to the Sceptic

The naturalist response to the sceptic is often very hard to pin down. For many naturalists, their attitude towards scepticism is so dismissive that the basis for dismissal is unclear. For example:

The project of responding to skepticism... is one which naturalists regard as a dead end. Naturalists will argue that this project has a history of failure, and the manner in which the project has failed calls the very point of the project itself into question (Kornblith 1999, 166)

This sort of response seems to suggest that we should just ignore the sceptic. But to the opponent of naturalism, such comments may appear to flippantly dismiss legitimate philosophical worries, or worse still, to refuse to answer a question on the grounds that it is too hard. As Barry Stroud argues, if the naturalist cannot provide an answer to the sceptic's question, then what is needed is a 'demonstration of the incoherence or illegitimacy of that question' (Stroud 1981, 468).

Kitcher adopts the latter strategy. The demand of the sceptic is a 'synchronic reconstruction of beliefs: take the totality of things you believe, subtract [the claim to be justified] and everything that you cannot defend without assuming it, and now show that the claim is correct' (Kitcher 1992, 90). But the naturalist should automatically realise the impossibility of certain synchronic reconstructions. In the absence of significant a priori knowledge, all these reconstructions are dependent on empirical bases. Consequently, the reconstruction of principles on which all a posteriori enquiry relies (such as the principle of induction, or empiricism) will be impossible. The absence of epistemic foundations inevitably imposes certain epistemic limitations. On this understanding, the reasons for dismissing the sceptic's questions are the same arguments for naturalism in the first place.

And yet, by no means all naturalists adopt this position. As Haack highlights, the cooperative and replacement naturalists have different options open to them as

regards the sceptic (Haack 1993, 343). For a moderate naturalist who recognises the *a priori* method, there is still the open possibility of finding an *a priori* source of certainty as a response to the sceptic, or else of refuting the sceptic's argument by other *a priori* means. Kitcher's position only applies to replacement naturalists. And even then, some replacement naturalists do not always refuse to converse with the sceptic. Quine for example considers sceptical questions quite meaningful. But, he emphasises that the sceptic is reliant on scientific knowledge. He argues that the awareness of illusions as illusions is only possible against a background of scientific theory; scepticism is 'an offshoot of science' (Quine 1975, 67). This is not to accuse the sceptic of begging the question. Their doubts, he says, are legitimate. Quine simply wants to show that 'sceptical doubts are scientific doubts' (Quine 1975, 68). The response to the sceptic remains a legitimate part of epistemology, but as with the rest of epistemology, it is a scientific problem. So whilst Quine is allowing the sceptic more legitimacy than other replacement naturalists might, he is still not addressing them in the way of traditional epistemology. But crucially, Quine's sceptic is a sceptic within science.⁴ It is only this internal scepticism that the naturalist can provide answers to.

There is then, no single naturalist response to the sceptic. Whilst replacement naturalists refuse to engage with the traditional sceptic on his terms, cooperative naturalists allow for the possibility of engagement with the traditional sceptic in the style of traditional epistemology.

TE5: Foundationalism

Is there a typical naturalist position on foundationalism? Not really. But before discussing the various naturalist perspectives, a further clarification of the term "foundationalism" is required. As defined above, foundationalism seeks a basis of justification in non-inferential, certain knowledge. As a picture of foundationalism within traditional epistemology, this is true. However, more recent philosophers have separated non-inferentiality from certainty. This more recent definition of foundationalism requires the pursuit simply of non-inferential knowledge, not certain knowledge. This is a subtle, but hugely significant difference. For the sake of clarity therefore, I will refer to these two positions as traditional foundationalism (which pursues non-inferential, certain knowledge) and contemporary foundationalism (which requires only non-inferential knowledge).

The most famous example of anti-foundationalism in the naturalist tradition is Quine's. In fact, Quine's turn to epistemological naturalism is framed as a direct response to and motivated by the failure of foundationalist epistemology: 'The Cartesian quest for certainty had been the remote motivation of epistemology... but that quest was seen as a lost cause' (Quine 1969b, 74). For many subsequent naturalists, without the threat of the sceptic legitimised, they lack the motivation for traditional foundationalism. 'Philosophical foundationalisms look for answers to... questions that hold the sciences to stricter standards than the sciences as we

⁴ For more, see Uebel (2004b).

now find them hold themselves’ (Roth 2003, 282). But as always with naturalism, this position is not universal. Many also reject the notion that any such foundation is forthcoming; the two obvious candidates of a priori knowledge and immediate experience are each deeply flawed. Additionally, naturalists typically situate any discussions of justification within the context of the current state of knowledge, relative to current practices. But whether this constitutes an alternative position to foundationalism, rather than simply a rejection of it, is not always clear. Quine, with his holistic view of beliefs as horizontally organised in fields rather than hierarchically structured, is sometimes considered a coherentist for example. But, as Olsson argues, ‘mere rejection of foundationalism is not itself an alternative theory because it leaves us with no positive account of justification, save a suggesting metaphor about webs of belief’ (Olsson 2021, §). In this sense, whilst most naturalists are forthcoming with a rejection of foundationalism, they are not always so forthcoming with an alternative.

1.2 Naturalism in the Vienna Circle

The basic naturalist attitude appears universal within the Vienna Circle. Dissatisfaction with traditional metaphysical philosophy, which they dismissed as cognitively meaningless, and an enthusiasm for the sciences were the shared basis for the Circle’s formation. The Vienna Circle’s manifesto calls this attitude the ‘*scientific conception of the world*’ (Carnap et al. 1929/1973, 2). The ‘goal ahead is *unified science*’ and this means ‘rejecting overt metaphysics and the concealed variety of apriorism’ (Carnap et al. 1929/1973, 5–6). But much diversity hides behind this apparently united front. The easiest way to clarify the forms of naturalism espoused by the Circle is to go through the hallmarks of naturalistic and traditional epistemology, and the distinction between cooperative and replacement naturalism, and see to what degree the members of the circle express these positions. But first, an introduction to the Circle itself is necessary.

1.2.1 *The Wings of the Vienna Circle*

The “Vienna Circle” was a group of logicians, scientists and philosophers who participated in regular meetings which grew out of the Thursday night discussion group around Moritz Schlick, beginning in 1924.⁵ The most important members of

⁵ The scholarly debate over the status of the “first Vienna circle”, Frank, Hahn and Neurath’s coffee-house discussion group before WWI, is not important for current purposes and so no position is taken on it here. See Haller (1985/1991), Uebel (2003a), Uebel (2003b), and Limbeck-Lilienau (2018).

the Circle (both then and for current purposes) were Moritz Schlick, Rudolf Carnap, and Otto Neurath. Less well known, but of great importance within the Circle and for the current work, were Hans Hahn and Philipp Frank. Neurath proposed the name for the group with the intention that it ‘would be reminiscent of the Viennese waltz, the Vienna woods, and other things on the pleasant side of life’ (Frank 1949b, 38). Although it may have failed to carry the desired connotations, the name stuck. The Vienna Circle became known as such when it entered what Stadler calls its ‘public phase’ in 1929, with the publication of the manifesto *The Scientific World Conception: The Vienna Circle* and the group’s participation as the “Vienna Circle” at the First Conference on the Epistemology of the Exact Sciences in Prague (Stadler 2007, 17).⁶ This was in addition to the establishment of the Ernst Mach Society in 1928, and was followed in 1930 by the establishment of the journal *Erkenntnis*. Both were intended for the wider dissemination of the Vienna Circle’s ideas, and demonstrated the self-conscious move from private discussion group to incipient international intellectual movement.

For the purposes of this chapter, I am dividing the Circle into three parts; the left-wing, Carnap, and ‘the more conservative right wing’ (Carnap 1963d, 57). The right-wing is primarily represented in what follows by Schlick. The left is represented by Frank, Hahn, and most prominently, Neurath. Carnap will be treated separately. Importantly, this is not a rejection of the common left-right division of the circle; this tripartite division is purely for exegetical purposes. The left-right distinction is legitimate, and I agree that the division can be understood as a clash between the influences of Neurath and Wittgenstein, via Schlick.⁷ Additionally, and more importantly, I will argue that the left-right distinction is one between incipient naturalism and traditional empiricism. The mature Carnap also belongs within the Circle’s left-wing. So why treat Carnap separately? Because, of all the members of the Vienna Circle, Carnap’s ideas undergo the most frequent and most significant changes. It therefore requires more interpretive work to demonstrate that Carnap belongs within the left-wing; to show that he became a naturalist.

Wittgenstein’s influence dominated the right wing of the Circle around Schlick and Waismann who were closer to him personally and philosophically than other members of the Circle were.⁸ They endorsed a picture of philosophy as a clarificatory project that, despite its anti-metaphysical tendencies and enthusiasm for the sciences, cannot be considered naturalistic due to its continuation of traditional epistemological methods and problems. The members of the Circle that formed the

⁶ The manifesto was neither official nor approved by the whole Circle. It was written primarily by Neurath and Carnap with input from Feigl, Frank, Hahn and possibly Waismann. Despite it being dedicated to him, Schlick was unhappy with the content and tone of the manifesto. For details, see Uebel (2008).

⁷ See for example Rutte (1979/1991), Haller (1982b, 192), and Stadler (1982/1991, 159). Fittingly, Ayer reported that at the Circle’s meetings, Schlick sat at the head of the table with Neurath at the opposite end (Ayer 1977, 133).

⁸ By 1929, Wittgenstein’s personal contact with the Circle was limited to Schlick and Waismann. Schlick was a friend of Wittgenstein’s, and the member of the Circle Wittgenstein respected most. Waismann, worked closely with Wittgenstein for years, despite a strained personal relationship.

left-wing were, not coincidentally, ‘those members who had already emerged in the 1920s as critics of Wittgenstein: Neurath, Carnap, Hahn, and Frank (with decreasing vehemence)’ (Stadler 2015, 60).⁹ Neurath in particular had always been suspicious of Wittgenstein, ‘regard[ing] him from the start as a mystic and metaphysician... an antiscientific person through and through’ (Neurath 16/06/1945, in Cat and Tuboly 2019, 640). Neurath functioned as the opposite pole of the Circle, leading the ‘eternal quartet’ of Carnap, Frank, Hahn and himself to pursue a project of naturalised epistemology (Neurath, Quoted in Haller 1982c/1991, 193).¹⁰ Neurath is now widely recognised as espousing some form of naturalised epistemology. Any plausible reading of Frank and Hahn places them squarely in the same camp as Neurath. Frank in particular can usefully be viewed as practicing what Neurath preaches; in Frank’s historical and sociological studies of science we have the best example of, or at least the closest approximation to, Neurath’s vision for the empirical study of science which he called the ‘behaviouristics of scholars’ (Neurath 1936a/1983, 137). They also shared a vision for the Logical Empiricist and Unified Science movements. In the later period of Neurath’s life, he became increasingly worried by the direction of the Logical Empiricist movement, worries that were shared by and reinforced by Frank.¹¹

It is more controversial, however, to label Carnap a naturalist. Kitcher, a naturalist himself, places Carnap within the analytic Fregean anti-naturalist tradition (Kitcher 1992, 53–4). He sees Carnap’s rational reconstructionism as a different project, with a ‘method... [and] ordering of philosophical problems’ distinct from the naturalist’s (Kitcher 1992, 54). Similarly, Koppelberg describes Carnap and Neurath’s epistemological theories as ‘doctrines on which [they] deeply disagreed’ (Koppelberg 1990, 201). Koppelberg sees both Quine and Neurath’s naturalism as a response to ‘the failure of epistemological foundationalism, above all in Carnap’s variant’ (Koppelberg 1990, 204). Consequently, interpreting Carnap as a naturalist requires greater justification, and more interpretive work. However, before tackling this task, I will begin with two simpler ones: categorising Schlick and Neurath.

1.2.2 *Schlick and the Right Wing*

In keeping with his role as the Circle’s paternal figurehead, Schlick was the most traditional, conservative member. Although the Circle’s manifesto was dedicated to him, Schlick did not share its broad revolutionary vision. Whilst Neurath, Hahn, Frank and Carnap were socialists, Schlick was an archetypal middle-class liberal. His work rarely makes mention of unified science, the scientific conception of

⁹ Carnap was initially more sympathetic to Wittgenstein than Neurath ever was, but was disliked by Wittgenstein personally. Wittgenstein asked Schlick not to bring Carnap to meetings and then accused him of plagiarism. See (Stadler 2015, 224–228).

¹⁰ In his diaries Carnap refers to this group as ‘our quadruple’ (Quoted in Tuboly 2017, 260 n.).

¹¹ Reisch (2005, Chapter 10 & 11).

the world, or the Encyclopedia project (which as we will see below, are central to the left-wing's logical empiricism). Where Neurath described his position as anti-philosophy and pro-Unified Science, Schlick saw the purpose of the logical empiricist movement as the reorientation of philosophy. The left-wing of the Circle were deeply influenced by Ernst Mach, whom Frank called 'the real master of the Vienna Circle' (Frank 1938/1949, 79), and French conventionalists like Pierre Duhem and Henri Poincaré.¹² But Schlick's biggest intellectual debt, after 1925, is to Wittgenstein.¹³ Like Wittgenstein, Schlick saw philosophy not as a body of knowledge but as an activity. He argued for a turning point within philosophy, not away from it (Schlick 1930a/1979). Whilst Schlick was staunchly against traditional metaphysics and school philosophy he maintained that anti-metaphysics was a longstanding position within philosophy, a tradition beginning with Socrates and within which he saw himself and the logical empiricist movement (Schlick 1936/1979, 492). He falls on the radical edge of traditional epistemology, but stops well short of naturalism.

1.2.2.1 A Priority and Independence of Science

In many ways, Schlick is a traditional empiricist. Like the rest of the Circle he accepted a version of Hume's fork, dividing formal analytic a priori knowledge from substantive synthetic a posteriori knowledge: 'all propositions are either synthetic *a posteriori* or tautologous' (Schlick 1930b/1979, 166). As an empiricist, he denied that there can be factual a priori knowledge. But, as an empiricist, Schlick faced the longstanding problem of accounting for the apparent a priority of knowledge in maths and logic. The solution to this classic empiricist dilemma was derived from Wittgenstein; mathematical and logical knowledge is tautological (Schlick 1930b/1979). Analytic knowledge lacks empirical content, because it concerns merely the rules for the transformation of symbols. An analytic statement bears no factual content, it 'is true by virtue of its form alone' (Schlick 1930b/1979, 162). All members of the Circle, as we will see below, accepted something like this position. This may suggest that Schlick has some naturalist tendencies, but really this denial of a priori knowledge of the world is a modification of traditional empiricist epistemology (the adoption of Wittgenstein's view of logic is what made the Vienna circle *logical* empiricists). Whilst it may be shared by many naturalists, by itself it does not warrant applying the label to Schlick.

Schlick is not however an entirely traditional philosopher. He denies that there can be any meaningful philosophical statements. Science is the system of empirical

¹² For the influence of Mach and the French conventionalists on Frank, Hahn and Neurath, see Haller (1985/1991) and Stadler (2015, Chapter 1). On Carnap, (Carus 2007a, Chapter 4).

¹³ Schlick called Wittgenstein's *Tractatus* 'the most significant philosophical work of our day' (Schlick 1928/1979, 136). He also outlined his vision for the future of philosophy along Wittgensteinian lines (Schlick 1930c/1979).

knowledge, and since all knowledge is empirical, ‘*the system of knowledge*; there is no additional domain of “philosophical” truths’, since any non-empirical philosophical knowledge would have to be synthetic a priori, which Schlick rejects as meaningless metaphysics (Schlick 1930a/1979, 157). But this does not amount to a rejection of philosophy as a discipline. Schlick promoted a revision of what we understand philosophy to be; ‘philosophy is not in fact a science or system of knowledge at all, but an *activity*’ (Schlick 1929/1979, 142). Here Wittgenstein’s influence on Schlick is most obvious. Schlick still sees a unique role for philosophy: ‘It is the proper business of philosophy to seek for and clarify the *meanings* of claims and questions’ (Schlick 1932a/1979, 263).

For Schlick then, the independence of philosophy from science is sharply maintained. He is very clear that ‘Philosophy is most certainly not a science’ (Schlick 1932b/1979, 368). Science is the realm of truths, philosophy the realm of meanings. ‘Philosophy elucidates propositions, science verifies them’ (Schlick 1930a/1979, 157). Whilst philosophy may therefore play a role in understanding the language of science, the tasks and subject matter of the two disciplines are entirely distinct. Neither discipline can contribute to or supplant the tasks of the other. Has philosophy therefore been demoted from the supra-scientific, most “fundamental” discipline to a more supplementary role? Even this Schlick denies:

[Philosophy is] still something so great and significant that it may continue to be honoured henceforth, as in former days, as the queen of the sciences... the philosophic activity of giving significance is thus the alpha and omega of science (Schlick 1930a/1979, 157)

For Schlick then, the purpose of logical empiricism is a reorientation of philosophy onto its proper course. This does not mean the integration of philosophy into the sciences. He is definitely not a replacement naturalist. In fact, with the strict separation of philosophy from science, he cannot even be considered a cooperative naturalist. Far from it; he explicitly denies the methodological relevance of the sciences to philosophy, and the incorporation of philosophy into the sciences, despite it being philosophy’s task to assist our understanding of the sciences.

1.2.2.2 Foundationalism and Scepticism

Unlike most naturalists, Schlick recognises the demands of the sceptic, and sees formulating an answer to them as a crucial task for the philosopher. Schlick treats the sceptical threat much as Descartes does; ‘we must utilize stretches of the Cartesian road, so far as they are sound and passable’ (Schlick 1934/1979, 380).¹⁴ And like traditional epistemologists, Schlick concludes that to answer the sceptic we must ‘seek for an unshakeable foundation, immune from all doubt and forming the firm basis on which the tottering edifice of knowledge is reared’ (Schlick 1934/1979, 370). Only a certain foundation of non-inferential, universally justified

¹⁴ On the importance of the response to sceptic for Schlick, see Popkin (1992) and Uebel (2020a).

knowledge is sufficient to alleviate the sceptic's doubts. Schlick is quite clearly, in the terms introduced above, a traditional-foundationalist. Schlick also expresses a very traditional understanding of the basis of epistemology: 'All great attempts at establishing a theory of knowing arise... from the wish for absolute certainty' (Schlick 1934/1979, 370). As for the issue of normativity, Schlick is so comfortable operating in the traditional normative realm of epistemology that he never even considers the topic worth discussing.

Although Schlick formulates the problem of the sceptic and the need for foundations in a traditional way, his solution is unconventional. Whilst traditional epistemologists searched for certainty at the outset of enquiry, from which reliable predictions can be generated, Schlick finds it in the conclusion of enquiry, in 'affirmations' (Schlick 1934/1979, 381). Affirmations are statements about the content of immediate experience: 'There is yellow here now' (Schlick 1935b/1979, 409).¹⁵ They act as the vehicle for the verification or falsification of hypotheses, being 'unshakeable points of contact between knowledge and reality' (Schlick 1934/1979, 387). They are 'the only synthetic propositions *which are not hypotheses*' and therefore the only synthetic propositions that are non-inferentially justified (Schlick 1934/1979, 387). Like analytic statements, they possess 'absolute validity' (Schlick 1934/1979, 385). Also like analytic statements, the apprehension of the meaning and truth of affirmations is simultaneous. They are therefore immune from doubt and can function as the certain foundations for human knowledge.

What makes them unique as 'fixed points' on which knowledge is established, is that:

In no sense do they lie at the basis of science, but knowledge, as it were, flickers out to them, reaching each one for a moment only, and at once consuming it. . . These moments of fulfilment and combustion are of the essence. From them comes all the light of knowledge. And it is this light for whose source the philosopher is actually asking, when he seeks the foundation of all knowledge (Schlick 1934/1979, 387)¹⁶

¹⁵ Whether affirmations can be considered genuine statements, or ought best be understood as experiences or psychological states, has been a matter of interpretive controversy. Schlick says that affirmations are not part of the language of science: 'They do not occur within science itself, and can neither be derived from scientific propositions, nor the latter from them' (Schlick 1935b/1979, 407). Scheffler therefore calls affirmations 'a special class of statements that are not themselves within science but are nevertheless essential to its function and, in particular, its confirmation' (Scheffler 1982, 104). But as we will see below, Schlick often discusses them in psychological and experiential terms. As a result, Neurath derisively claims that affirmations 'can sometimes be treated as statements, sometimes as non-statements' (Neurath 1936d/1983, 159). Hempel similarly concluded that Schlick is 'is faltering between two incompatible views of "affirmations."' (Hempel 1983a/2000, 186). Modern interpreters have also puzzled over exactly how to conceptualise affirmations. See for example Holdcroft (1983); Hung (1985); Oberdan (1996). For a recent charitable reading of Schlick, see (Uebel 2020a). As is suggested in the parallels Schlick draws between affirmations and analytic truths, I think we should take Schlick seriously and interpret his affirmations as an epistemologically and psychologically unique kind of statement.

¹⁶ The psychological character captured in this quote is what makes the translation "affirmation" so apt. In the moment of empirical verification it is not simply that specific beliefs or theories are

The certainty that philosophy demands is attained only momentarily, in these fleeting moments of confirmation and certainty. Whereas most foundationalists look for the source of certainty prior to and as a basis for science, Schlick locates it as the terminus of a dynamic process of enquiry. Certainty is not something we can find and keep, but something we must constantly strive after and re-acquaint ourselves with. Whilst still arguably a form of foundationalism, Schlick inverts the typical expectations of this foundation, locating them as the end-goal rather than the starting point. Schlick is clearly not a completely orthodox foundationalist. But nor is he a naturalist, who rejects TE4 and TE5. He does not try to situate knowledge in the concrete actualities of our current state of knowledge. Nor does he dismiss the sceptic's doubt as insignificant. He is an esoteric foundationalist empiricist, who accepts the traditional problems and solutions of epistemology as legitimate, but approaches them in a unique way.

In sum, Schlick's philosophy is strongly empiricist and anti-metaphysical, but he is ultimately still a proponent of traditional epistemology. He sharply separates the disciplines, goals and methods of science and philosophy, and considers foundational certainty the only reply to the sceptic (this being a crucial philosophical task). Schlick might come closer to naturalism than many philosophers, but he cannot be considered one himself. This is in stark contrast to his main opponent within the Vienna Circle; Neurath.

1.2.3 Neurath's Naturalism

Neurath is an epistemological naturalist. This is widely accepted in contemporary literature.¹⁷ It is not contentious to claim that Hahn and Frank are too.¹⁸ As Creath has argued, Quine's failure to mention Neurath and his caricatured portrait of the Vienna Circle delayed this realisation for some time (Creath 2007b, 335). But naturalism is now clearly understood as a central commitment of the Unified Science movement and its scientific world conception, of which Hahn, Frank and Neurath were the most enthusiastic proponents.

The name "scientific world conception" is intended both as a confession of faith and as a delimitation of a subject: It is to *confess our faith* in the methods of the exact sciences... And it is to *delimit our subject* from philosophy in the usual sense: as a theory about the

confirmed via acquaintance with reality. These moments represent the fruition and vindication of the entire human epistemic endeavour.

¹⁷ See Hempel (1983a/2000), Koppelberg (1990), Haller (1982c/1991), Uebel (1991b), Friedman (2001), Glock (2008), and Cat (2019).

¹⁸ On Frank, see Uebel (2000, 2011); Reisch (2005, Chapter 10), Tuboly (2017) and (Reisch and Tuboly 2021). There is a significant dearth of scholarship on Hahn. Kraft places Hahn in the 'radical wing' led by Neurath (Kraft (1950/1953): 12). Uebel (2005) draws interesting comparisons between Hahn's logical thought and Carnap's principle of tolerance.

world claiming to stand next to the special sciences as their equal or even above them as their superior (Hahn 1930a, 20, trans. amended)

But which sort of naturalism they are advocating is not immediately clear; cooperative or replacement? Hahn wrote almost exclusively on mathematics, and unfortunately died very early in the Vienna Circle's lifespan. Frank's work is largely historical and sociological. To an extent, his methodological commitments must be extrapolated from his practice. Neurath therefore provides the most detailed expressions of naturalism, but unfortunately, he was not always clear or careful in his use of language, and understanding his precise position involves some exegetical work.

One easy way of understanding Neurath is as a forerunner of Quine. Koppelberg in particular has emphasised the affinities of Neurath's and Quine's naturalistic projects (Koppelberg 1989, 1990). Even Quine was 'impressed by the extent of agreement between Neurath and [himself]' (Quine 1990a, 212). But we must be careful to read Neurath on his own terms, and not to read him only in terms of Quine, because this masks some important differences between them, some subtle and some less so, that result in important differences between their individual forms of naturalism. The best place to start is to see where the left-wing stood with regards to the claims of traditional epistemology, and contemporary naturalized epistemology.

1.2.3.1 Neurath's Rejection of Traditional Epistemology

Neurath, Frank and Hahn are unequivocal in their rejection of a priori methodology as a path to knowledge of the world (TE1). Like Schlick, they too distinguish the synthetic a posteriori from the analytic a priori. As empiricists, they believed that 'experience is the only source capable of furnishing us with knowledge of the world, knowledge of facts, knowledge that has content' (Hahn 1929/1980, 39). Like Schlick, they saw the a priori statements of maths and logic as analytic, tautologous, and lacking substantive content. From a pragmatic standpoint, formal truths are essentially empty; the addition of an analytic statement to our body of beliefs makes no impact on our expectations. As Neurath put it, "tautologies" do not add anything. . . . The addition of this condition to a command or statement does not make any difference, it is always fulfilled' (Neurath 1931b/1983, 56). So, whilst Neurath, Hahn and Frank do allow room for the analytic a priori, they deny that these subjects give us the substantive knowledge of traditional philosophy.

From Neurath, we get very little explicit argumentation against a priori methodology. He expresses his support for the 'empiricist, anti-apriorist standpoint' of the Unified Science movement (Neurath 1937a/1983, 189). He also obliquely refers to Hume's fork, referring to the 'investigation of tautological combinations on the one hand and empirically testable statements on the other' (Neurath 1937d/1987, 135). But he seems to take apriorism as transparently flawed and already refuted, having been cast out alongside metaphysics as the metaphysician's primary methodology. The justification of this rejection is assumed in all of his mature work, certainly

from the period of the Vienna Circle and after, which consistently emphasise empirical methodology and empirical problems, that the defeat of metaphysics and its rejection as meaningless for Neurath is equivalent to the total empiricisation of enquiry. This is true of Frank's mature work too.

Neurath, Frank and Hahn therefore cannot be neatly categorised as to their attitudes to the *a priori*. In fact, they would appear to fall outside Roth's distinction between methodological monism and dualism. If "dualism" is understood simply as allowing for the use of two methodologies, then they are dualists, allowing both formal-analytic and empirical-synthetic forms of knowledge. But if "dualism" means a separation of science and non-science in virtue of their differing methodologies, they are not dualists. Here, we have the first and one of the most obvious divergences between Neurath and Quine. For Quine, the analytic/synthetic distinction is incoherent. He argues that there is no non-circular definition of analyticity, and that once we take into account the realities of confirmation holism and the universality of revisability, we cannot maintain the distinction as epistemologically significant. Consequently, as empiricists, the very idea of the *a priori* must be abandoned. Neurath, Frank and Hahn all maintain analyticity. And yet, they also accept those positions that would later act as Quine's basis for rejecting this distinction, without themselves drawing the same conclusion. The left-wing of the circle were holists, conventionalists, and fallibilists. How is it possible then, that they disagree with Quine on a priority? How is it that they can maintain the analytic/synthetic distinction, when Quine rejects it? The question is crucial, and I will return to it below when discussing Carnap's view of a priority.

A related issue this raises is how the left-wing relates to contemporary naturalists. As we will see below, the left-wing are replacement naturalists. But how can they argue for and justify their own naturalism without rejecting a priority outright? For contemporary replacement naturalists, inspired by Quine, the rejection of *a priori* methods is typically an important part of the argument for naturalism generally, and replacement naturalism in particular. What argumentative path is left open to the left-wing of the circle? Are the arguments of contemporary naturalists available to them, given their position on analytic knowledge?

1.2.3.2 Unified Science and Philosophy

Hahn, Frank and Neurath unequivocally reject the independence of philosophy from science (TE2):

The Vienna Circle does not recognize philosophy as a discipline with its own sentences that would occupy a place above the sentences of science as the court of highest appeal; it does not recognize two or more "modes of being" with correspondingly different "methods" (Neurath 1932b/1987, 1)

In fact, Neurath says his position would more accurately be called an 'anti-philosophy' (Neurath 1931a/1983, 48). But does this lead to cooperative or replacement naturalism? Haack diagnoses an ambiguity in the way that "science" is used

in discussions of methodological naturalism, that she thinks helps to distinguish between replacement and cooperative naturalists. She finds two uses of the term science, which she (unhelpfully) calls SCIENCE and *science*. The narrower of the two, *science*, refers 'to the natural sciences exclusively', while SCIENCE is 'broader, referring to our empirical beliefs generally (and thus including common sense, the social sciences, history and, according to Quine, mathematics, logic, and philosophy as well as science' (Haack 1993, 339). Haack argues that how we interpret "science" determines how we interpret the integration of science and philosophy. Incorporating philosophy into *science* gives replacement naturalism; epistemology is internalised within natural science. Incorporating philosophy into SCIENCE simply renders philosophy a part of the wider web of empirical belief. Does this shed any light on the naturalism of the Vienna circle?

Not directly. In fact, the left-wing of the Vienna Circle show Haack's way of dividing naturalism to be insufficient for our purposes.¹⁹ For the left-wing, there is no distinction between *science* and SCIENCE, even in principle:

Where, then, as a matter of principle, should we draw the dividing line between physics and history, sociology, and psychology? All these disciplines are completely interwoven, they are all in principle pursued according to the same method... for there is... only one science, namely *unified science*. (Hahn 1933/1987, 45)

Attempts to sharply delineate between hard and soft sciences, or natural and social sciences are rejected as unprincipled.²⁰ There is no way to distinguish between individual sciences in an absolute or methodological sense. The left-wing endorses the unity of science, arguing that the separation of the different sciences is a purely pragmatic matter, a division of labour for the purposes of utility. Unity of science means that the individual sciences are not of entirely different kinds, and do not possess entirely unique methodologies. And as for the hierarchy implicit in separating the "hard" core of the natural sciences from the "soft" social sciences, the unity of science is understood in terms of horizontal integration. The sciences more closely resemble a '*mosaic pattern*' than a rigidly ordered taxonomy (Neurath 1937b/1983, 204). Haack's division therefore cannot be the basis for distinguishing replacement and cooperative naturalists, at least not in this case. Reality enforces the actuality of the unity of science on us, when we consider questions and predictions about the actual world:

In order to formulate the individual prediction: "This forest fire will soon be extinguished", we combine biological statements (concerning tress, etc.), chemical statements (concerning fire, etc.), sociological statements (concerning fire service, etc.) and statements of other disciplines (Neurath 1936a/1983, 132)

The languages of the individual sciences are all translatable into the common language of universal science, demonstrating their unity.

¹⁹ These purposes are of course, not Haack's. Her introduction of this distinction is a means for a deeper understanding of Quine's naturalism, not the Vienna Circle's.

²⁰ See Neurath (1939/1983).

The left-wing of the circle is advocating something that Haack's distinction fails to fully accommodate. Unified science does not simply assimilate philosophy into either SCIENCE or *science*. It replaces philosophy, whilst assimilating certain aspects of what was previously epistemology and incorporating these elements into a new second-order scientific enquiry; a scientific metatheory, a science of science, within unified science more broadly. This new metatheory cannot be assimilated exactly to any other pre-existing first-order scientific discipline. What it requires is the study of science conducted according to the methods and standards of science itself. This is the discipline that Neurath refers to as the 'behaviouristics of scholars' (Neurath 1936a/1983, 137). Exactly what this means will become clear below.

For the left-wing of the Vienna Circle, unified science is the only legitimate subject of purely cognitive enquiry. If a statement, theory or question does not fall within the remit of unified science, then it is cognitively meaningless. 'Unified science comprises all the sentences of the particular sciences, but also all sentences about sentences, and in short, *all legitimate sentences*.' (Neurath 1932b/1987, 1–2). Only unified science is meaningful. There is no room for philosophy as a separate subject. 'Philosophy as an independent system of doctrines is obsolete. What cannot be regarded as unified science must be accepted as poetry or fiction' (Neurath 1931a/1983, 49). But this should not be confused with Schlick's position:

Some representatives of the "Vienna Circle"... still occasionally use the term "philosophy". They want this term to signify "philosophising", the "activity of clarifying concepts". This concession to the traditional linguistic usage, though understandable for various reasons, easily gives rise to misunderstandings... But the objection to the term "philosophising" is not only terminological; it is impossible to separate the "clarification of concepts" from the "pursuit of science" to which it belongs (Neurath 1931c/1983, 58–9)

This response to the Wittgensteinian wing of the circle demonstrates a far greater break with traditional philosophy than Schlick's. The clarification of concepts is a meaningful enquiry, and consequently falls within unified science.²¹ There is no special task left for philosophy as a separate subject.

There is a final significant consequence of the replacement of philosophy with unified science, and it is a methodological one. This is also where the rejection of TE1 becomes most significant. Neurath not only rejects the possibility of a cognitive discipline beyond unified science, but also the possibility of a methodology different to that of unified science. The unity of science combines with the rejection of a priori methods (with the exception of the formal subjects, math and logic) to give the conclusion that all meaningful substantive enquiry must follow the empirical methodology of the sciences; 'the same critical and constructive methods [apply] in all areas of research, argumentation, and analysis' (Neurath 1937d/1987, 132). What becomes of epistemology?

From all this it becomes clear that within a consistent [unified science] there can be no 'theory of knowledge', at least not in the traditional form... Some problems of the theory

²¹ An enquiry that, as we will see in Chapter 4, was pursued within the metatheory of science by Carnap.

of knowledge will perhaps be transformable into empirical questions so that they can find a place within unified science (Neurath 1931c/1983, 67)

The left-wing is very much in agreement with Laudan's position, that we must take seriously the sense in which legitimate epistemology is the scientific theory of knowledge (Laudan 1990, 45). Epistemology becomes the study of scientific knowledge from within science, a second-order metatheory of science itself.

Crucially, this does not mean that carryovers from epistemology are subsumed by one of the special sciences. Questions of justification are not somehow to be addressed by physicists and geologists. Rather, it means that those remnants of epistemology, transformed into the second-order metatheoretical scientific study of science itself, becomes a part of unified science in just the same sense as all other empirical disciplines. Despite being the metatheoretical study of the sciences, the methods and tools available to the metatheorist are the exact same tools available to all scientists within unified science; the empirical and formal methodology of the sciences. 'In the Unity of Science Movement one aims to talk as scientifically about science as one talks about plants, animals, or humans in the special sciences themselves' (Neurath 2011, 22–23). The metatheory of science becomes one scientific discipline among many within unified science; it just happens to be a second-order discipline with science as its subject matter. But as with physics, chemistry, and biology, the metatheory as a subdiscipline is separated in terms of its subject matter for the purposes of practicality. It is not differentiated by a method of a fundamentally different kind.

The preceding considerations should make it absolutely clear that the Left-Vienna Circle's project of Unified Science is a programme for naturalized epistemology. For Kornblith, it 'is the investigation of knowledge as a phenomenon in the world which distinguishes naturalism from other approaches to knowledge' (Kornblith 1995, 245). It should be clear that unified science embodies this approach. As for the form of this naturalism, it should also be clear that in Neurath's physicalism we find a strong form of unlimited replacement naturalism. This is perhaps the strongest possible form of replacement naturalism; science is the only substantive subject, and so far as philosophy can be absorbed into it, it is no longer considered philosophy. And yet, there is still a place for what was once epistemology within unified science, but as an empirical, metatheoretical, scientific subject; the theory of knowledge.

1.2.3.3 Foundationalism and Scepticism

The Circle's left-wing rejected the traditional epistemological task of refuting the sceptic via the pursuit of certainty as doomed to failure. For Neurath, the very notion of absolute certainty is anathema to the scientific world view; 'Whoever wants to create a world-view or a scientific system must operate with doubtful premises' (Neurath 1913/1983, 3). For Neurath, underdetermination is a fact. The evidence never uniquely satisfies one theory, there are always competing alternatives. And

confirmation is holistic. It is only ever groups of statements that are subjected to the test of experience. Strictly speaking, '[t]he whole of science is basically always under discussion' (Neurath 1935a/1983, 118). Consequently, we are never forced to reject any specific statement. That it is always possible for a system of beliefs to be altered to accommodate a statement or for the statement to be rejected is so fundamental to Neurath's philosophy that it has come to be called the "Neurath Principle" (Haller 1982a/1991, 121). Therefore, the scientist must accept that all and any statements are open to revision. 'There is no "*noli me tangere*" for any statement' (Neurath 1932a/1983, 95). This means there is also no possibility of isolating a class of privileged certain sentences, to act as a fixed foundation. It is this conjunction of holism, fallibilism and underdetermination that generates Neurath's most famous image:

There is no *tabula rasa*. We are like sailors who have to rebuild their ship on the open sea, without ever being able to dismantle it in the dry-dock and reconstruct it from the best components (Neurath 1932a/1983, 92)

Here Neurath anticipates many of the epistemological positions that are typically attributed to Quine and have come to dominate the naturalist tradition, rejecting traditional foundationalist accounts of justification. As a result, Neurath has often been interpreted as a coherentist about justification (Cat 2019, §3.3). This seems consistent with what he says:

Each new statement is confronted with the totality of existing statements that have already been harmonised with each other. *A statement is called correct if it can be incorporated in this totality...* The definition of 'correct' and 'incorrect' as proposed here abandons the definition that is usually accepted (Neurath 1931c/1983, 66)

In other words, "correct" means accepted, not true. This is a coherentist approach to justification, not a coherentist theory of truth. The two must not be confused. What Neurath is advocating is that coherence across the body of currently accepted statements is the basis of justification.

Importantly though, we must not attribute a naive coherentist account of justification to Neurath, according to which any internal inconsistency within a set of beliefs is sufficient grounds for their rejection as unjustified. He argues that sentences in isolation are neither valid or invalid. Rather, validity is only applicable to statements 'in connection with masses of statements' (Neurath 1936a/1983, 161). Neurath maintains that while positive coherence enhances the legitimacy of a body of beliefs, the presence of contradictions is not immediate grounds for rejection. Because of his attention to the reality of scientific practice, Neurath makes clear that a plausible coherentism must allow of at least some degree of internal inconsistency within a justified set of beliefs, or else justified beliefs become a practical impossibility. Neurath recognises that even our best science is not free of contradictions (Neurath 1936a/1983, 160). In an anticipation of Kuhn's later account of anomalies, he argues that contradictions are inevitable. He writes:

The question which contradictions can just be tolerated, which not, how one behaves altogether in the development of the whole of science, is a question of behaviouristics, of history of science, of behaviouristics of scholars. But the discussion of contradictions,

the discussion of the question, which groups of statements are logically of equal content, belongs to the sphere of logic (Neurath 1936d/1983, 169)

It therefore becomes clear that if Neurath is advocating a version of coherentism, it is an empirically determined form. The normative elements are not derived *a priori* (more on which below). In fact, here we have something of a reversal of the naturalistic fallacy. *A priori* logic can tell us what sets of statements are contradictory, but they cannot prescribe what degree of inconsistency within a body of statements can be tolerated. This is a question of theory acceptance. “Acceptance”, however, is not a logical term’ (Neurath 1936d/1983, 170). It is an empirical question, to be addressed by the scientific theory of knowledge. What comes into view here will occupy us greatly below: the division of labour in epistemology between the logical (what a contradiction is) and the empirical (what degree of contradiction can be allowed). For the time being, we must simply caution against reading Neurath as a naïve coherentist.

But where does this leave Neurath as regards the sceptic? Clearly, he doesn’t think foundationalism is the correct response. But what is? As we saw, there are two common naturalist responses to the sceptic. One is to refuse to play the sceptics game, the other is to interpret the sceptic’s questions as internal to science, and not traditionally philosophical. Arguably, these are two sides of the same naturalistic coin and not mutually incompatible. As I suggested earlier, the argument for the rejection of sceptical doubts from a naturalist perspective is the argument for the naturalist perspective. This is, I think, also applicable (or at the least available) to Neurath. If, as I have argued, Neurath’s argument for naturalized epistemology is premised on fallibilism, holism and underdetermination, this same argument is a potential reply to the sceptic. If Neurath is right that the scientist is always onboard the ship, then the demands of the sceptic are clearly a demand for the impossible. As mentioned above, the sceptic demands synchronic reconstruction. This is a demand to step outside science, to justify science from an external Archimedean point. But such a step is impossible. We can only ever remain within science:

There is no judge in a chair who decides who is nearer to the truth. There is no way of ‘impartiality’ or ‘scientific objectivity’, there is no point outside our life, from which we may finally decide what is ‘impartial’ or ‘scientifically objective’ - we do not see such a point. (Neurath 1946b/1983, 243)

There is no external means of justification. Neurath, like many naturalists, can therefore argue that demanding the impossible is an illegitimate demand.

But Neurath doesn’t obviously seem to adopt this tactic. Rather, he says that ‘the possibility of science becomes apparent in science itself’ (Neurath 1931c/1983, 61). Frank makes the similar claim that the ‘fact that no special science can... “defend its own principles” does not lead to the conclusion that the system of all sciences cannot do so’ (Frank 1951, 30). Initially, their position is unclear. Are they arguing that unified science itself can answer the sceptic in the traditional sense? That unified science can provide the philosophical answer the sceptic desires? No, this would be a misinterpretation of clumsy language. I think they should be interpreted along similar lines to Quine. On this interpretation, they are arguing that, as far as questions about “the possibility of knowledge” are meaningful, they are

answerable empirically. This would presumably involve interpreting the question of “possibility” in a scientific rather than philosophical sense; for example, in terms of the social practices of pursuing knowledge, and the psychological faculties involved. This would be an interpretation like Kitcher or Kornblith’s descriptions of the study of knowledge. Science itself can account for the possibility of science.

I think Neurath can be plausibly interpreted as expressing both positions in different works. But this is unproblematic, as these two replies to the sceptic are compatible. Neurath can argue from the naturalist position that the sceptic’s demands are impossible, if they are understood in their traditional sense. But he can still allow that, if interpreted as scientific questions, then the sceptic’s doubts are answerable by science itself and do not require external validation.

1.2.3.4 Neurath’s Affirmation of Naturalism

The argument above has the immediate consequence that the only basis for understanding justification is close attention to the actual practices of scientists through empirical study, in the form of sociology, behaviouristics and history of science. A basic methodological maxim follows: fidelity to scientific practice is of paramount importance. Being able to explain and account for the actualities of real-world scientific practice is essential for the methodological naturalist. We can call this the maxim of the primacy of scientific practice.²² It is clear in Neurath’s advice for tackling questions of epistemology: ‘We best start from the operation of science and look at its procedure’ (Neurath 1936d/1983, 159). This maxim is familiar to contemporary naturalists: ‘philosophical problems about knowledge can be satisfactorily addressed only by considering the ways in which historical and contemporary figures actually undertake their projects of inquiry’ (Kitcher 1992, 69). Evidently Neurath embraced it too.

Beyond merely affirming this maxim, Neurath consistently adhered to it in his work. For example, in his reply to Popper, Neurath argues that the unconditional adoption of the methodology of falsification is unsustainable.²³ Strict falsificationism would make the practice of actual science impossible; it cannot account for the holistic nature of theory testing, or the fact of underdetermination (here Neurath’s arguments strongly prefigure Kuhn’s “normal science” arguments against Popper). Neurath urges the empirical study of how scientists actually do or have responded to cases of falsification. ‘The unconditional preference for falsification cannot be successfully maintained in the framework of a theory of research’ (Neurath 1935b/1983, 124). Again, returning to the ship metaphor, rebuilding the ship on the open sea means operating within the confines of the actual, current state of knowledge. The metaphor of the ship specifically locates the practice of science

²² Uebel similarly highlights ‘primacy of practice’ as a commitment of naturalized epistemology which Neurath adheres to Uebel (1991b, 624).

²³ Another good example of Neurath’s commitment to this maxim are his studies of the history of optics (Neurath 1915/1973, 1916/1983). For a detailed discussion, see Zemlén (2019).

in the concrete context of the now, our contemporary body of knowledge and methodology. Neurath's metaphor contains the primacy maxim within it implicitly, especially when combined with his proposal to naturalize epistemology.

The adherence to scientific practice is even more obvious in Frank's work. As suggested above, this is why we can view Frank as implementing the proposals made by Neurath. Frank's works are largely historical or sociological, and typically concern the relationship between science and the wider world. But, rather than polemical, Frank's essays tend to be diagnostic; they explain for example how metaphysical views have arisen, as a result of historical, sociological or purely intellectual reasons. He discusses how the development of quantum mechanics was used to argue for freedom of the will (Frank 1936a/1987, 116–9) or the effect of the 'chance philosophy' we passively acquire through our upbringing, which is influenced by various social, intellectual and institutional factors, and the effect these beliefs can have on the way they present scientific theories and their consequences (Frank 1962, xix). As much as these works are of interest to the historian, they also have significant normative value. Frank intends that we learn from the lessons of the past. Frank traces and explains what facilitated the development of metaphysical theories so that we can prevent the same mistakes from developing within the methodology of unified science.

It follows from the maxim of the primacy of scientific practice that, as a matter of fact, epistemology must account for the impact and import of extra-cognitive factors on the process of scientific reasoning. They cannot be simply ruled out or ignored as non-ideal. Frank's historical work, which frequently anticipates Kuhnian or post-Kuhnian themes, consistently emphasises the significance of non-scientific, external factors on the decision-making of scientists.²⁴ Specifically, he highlights the way underdetermination can necessitate or facilitate extra-cognitive factors in theory choice.

To understand that the observed facts do not determine scientific doctrines unambiguously is the most important prerequisite for understanding the role played by sociological factors in the acceptance of scientific doctrines. Many people have claimed that there cannot be any influence of state or church on scientific doctrines because no authority can "change observed facts"... But an "authority" can require the abandonment or acceptance of scientific theories without requiring a "change of the facts". (Frank 1951, 19)

Frank and Neurath both recognise the underdetermination of theory by evidence, although they do not use the phrase, a consequence of their early study of the French conventionalists, especially Duhem (Neurath 1931c/1983, 66; Frank 1951, 19). Neither saw this as an insurmountable problem so much as a simple fact about the practice of science. Frank emphasises that empirical data always underdetermines scientific theory, that 'it is not possible to derive a scientific doctrine from observed facts' and that historically external factors determine which theory is accepted (Frank 1951, 19). He gives the example of the competing

²⁴ Alexander Bird has argued that Kuhn's account of scientific change is primarily internalist, and that the attributions to him of an emphasis on the role of external factors on scientific change are misinterpretations. See Bird (2012).

Copernican and Aristotelian models of celestial mechanics, where the acceptance of one theory over the other was ultimately determined by theological and political commitments (Frank 1951, 19). The same point is emphasised by Neurath; ‘a logically tenable multiplicity is reduced by life’ (Neurath 1935a/1983, 117). The details of Neurath and Frank’s account of decision-making and rationality in light of underdetermination will be explored in detail in the next chapter.

With Neurath’s naturalism fleshed out in this manner, we are in a position to attempt to explain the status of normativity in his naturalism. As mentioned above, the possibility of normativity in a strong naturalized epistemology is contentious. I will argue that Neurath does make the legitimization claim. In fact, normativity is essential to the Unified Science movement, as understood by the left-wing of the circle. But before addressing this claim, we must discover the source of Neurath’s normativity.

The problem with explaining Neurath’s position is that we only get glimpses of it. Neurath never spells it out; we have to tease it out of his work as a whole. The role of justification for Neurath is not the same as that sought by traditional epistemologists. He only needs justification that is sufficient to allow the successful practice of science, not an answer to the traditional sceptic. Consequently, he only needs an account of normativity that will allow for prescriptions of the best course for the pursuit of scientific goals.

As argued above, Neurath takes coherence to be the basis of justification. But in light of the reality of underdetermination, this still places insufficient restraints on the prescriptions for what scientists should do. There may well be multiple equally consistent theories to choose between. Coherentism alone therefore cannot solve this problem. But we have also noted that Neurath considers decision-making an important aspect of theory choice within science, both as a matter of historical fact, and as forced on us as a consequence of underdetermination. Decision-making therefore, must also play a key role in the choice of hypotheses and theories. Decisionism is adopted for reasons that bring instrumental normativity to Neurath’s theory. In this, he is to varying degrees close to a number of other replacement naturalists. To be sure, given his suspicion of the metaphysics associated with it, “truth” cannot be the cognitive aim to which Neurath tethers his methodological norms as it is for many contemporary naturalists. But a plausible alternative that fulfils the same role without the perceived metaphysical baggage is empirical adequacy, understood as the ability to make testably veridical predictions. Unlike most contemporary naturalists, Neurath cannot ground this instrumentality in a correspondence-theoretical realism, but in this respect he is no worse off than Laudan. Like him, but unlike contemporary naturalists like Kornblith or Kitcher, Neurath can ground the normative force of his methodological prescriptions on the testable empirical successes for controlling and predicting which following them promises.

Neurath’s stance on theory acceptance shows that, as an unlimited naturalist, he takes normative matters to be decidable only by empirical, not a priori, means. ‘[I]f we reject the notion of a philosophical system which is to legislate for the sciences’ then the only option is for science to legislate for itself (Neurath 1936d/1983,

176–7). By tethering it to testably successful predictions, the normative element of epistemology is itself absorbed into unified science. As we saw, Neurath argues that only through historical and sociological studies can we learn the degree of internal contradiction compatible with a theory being accepted. What this means, although as always Neurath is not clear about it, is that normative lessons can be derived from these empirical studies. These studies teach us how effective the acceptance of certain norms and methods were. In order to avoid the naturalistic fallacy, all naturalistic norms are hypothetical and conditional. The purpose of these studies is that they allow us to revise both the methods of the sciences and the norms of science in light of them. If the goals of science are only adopted by convention, then they are permanently revisable. They are always capable of revision and improvement. We will return to these considerations in the next chapter.

We are now in a position to clarify the form taken by the theory of knowledge within unified science. Considerations so far show that there are significant similarities between the project of unified science and the familiar form of contemporary replacement naturalism. In the naturalized epistemology of unified science, epistemology is replaced with the empirical theory of knowledge. Knowledge is not a concept to be analysed a priori, but a natural phenomenon in the world, embedded in a social and historical context. And as with all natural phenomena, the method for understanding it is the empirical methodology of science.²⁵ In other words, the primacy of practice is an inevitable consequence of taking the theory of knowledge seriously within a replacement naturalist epistemology. If we want an epistemology within the purely empirical discipline of unified science, it must be composed of behaviouristics, sociology, psychology, and history. We must study the actual processes and methods of the accumulation of knowledge as physical phenomena in the world, open to the exact same scientific study as any other physical phenomena. Frank called this the ‘pragmatics of science’, while Neurath preferred the ‘behaviouristics of scholars’ (Frank 1957/2004, 360, Neurath 1936d/1983, 160). In sum, Neurath and Frank (and Hahn) emerge from our investigation as bona fide epistemological naturalists, albeit also with distinct features not shared with contemporary naturalists. Let us now turn to a first consideration of Carnap.

1.3 Carnap and Naturalism

There are two key reasons that Carnap’s naturalism has been widely overlooked. The first is the historical narrative originating with Quine, which portrays Carnap as the foundationalist empiricist par excellence whose profound failure necessitates the transition to naturalized epistemology.²⁶

²⁵ Here, I mean natural simply in the sense of being open to empirical enquiry; Neurath would use the term physical.

²⁶ Verhaegh has recently argued against this reading of *Epistemology Naturalized* as making an ‘argument from despair’ (Verhaegh 2018, 16). He argues instead that Quine’s argument is that ‘the

In this narrative, Carnap is often equated to the dogmatic caricature of Logical Empiricism discussed in the Introduction. This narrative exacerbates the second issue; Carnap's self-presentation. The radically naturalist ideas Carnap endorses are often masked by both the calm and diplomatic tone of his writing, but even more significantly by his fluctuating use of terminology. By dissecting his changing use of language, the development of this radical naturalism can be laid bare.

1.3.1 Carnap's Convergence with Neurath

Throughout his life, but especially during the 1930s, Carnap's ideas and the terminology in which he expressed them were in a state of flux. It may very well be right to maintain that the Carnap of 1929 is not a naturalist.²⁷ However, the Carnap of 1936 is. The period between 1929 and 1934, from *Aufbau* to semantics, should be viewed as one of constant leftward movement for Carnap as he was increasingly won over by the Neurathian naturalistic perspective. Carnap is never explicit about this transition. Perhaps he did not even recognise it himself. Interestingly though, Neurath seems to have seen it. In a letter to Carnap in 1932, he calls Carnap a 'centrist tending towards the left wing' (Quoted in Uebel 2007a, 161). But by 1933 Neurath includes Carnap in the eternal quartet, suggesting he recognised Carnap's leftward trajectory over the previous year.

In the early period of the Vienna Circle, Carnap's work sounds very close in spirit to Schlick. He describes philosophy as 'consisting in the *logical analysis of the statements and concepts of empirical science*' (Carnap 1930/1959, 133). This sounds like anti-metaphysical empiricism, but within the bounds of traditional philosophy rather than Neurath's naturalized epistemology. This is also consistent with interpretations of Carnap, especially in the *Aufbau*, as an empiricist foundationalist, aiming to demonstrate the possibility of knowledge on the foundation of immediate experience. But within a few years, Carnap's terminology is very different: 'Philosophy deals with science only from the *logical* viewpoint. *Philosophy is the logic of science*, i.e., the logical analysis of the concepts, propositions, proofs, theories of science' (Carnap 1934c, 6). If we have in mind the interpretation of the early-Carnap given above, this quote can easily be seen as advocating the same thing. But the difference is not merely terminological, as becomes apparent from what Carnap says next:

epistemologist's quest for foundations was *misguided from the beginning*' (Verhaegh 2018, 30). It seems however that Verhaegh's new interpretation is consistent with and complementary to the "argument from despair" interpretation. What better reason to conclude a project was misguided from the beginning than recognition of its failure?

²⁷ As mentioned above, Pincock (2005) has suggested a naturalist reading of the *Aufbau*. But whether the *Aufbau* is interpreted as propounding a naturalist, phenomenalist, or Neo-Kantian project is not of significant consequence for the current argument. My argument focuses primarily on Carnap's post-1934 position, at which point he had moved beyond the *Aufbau* project.

What one used to call epistemology or theory of knowledge is a mixture of applied logic and psychology (and sometimes even metaphysics); insofar as this theory is logic it is included in what we call logic of science; insofar, however, as it is psychology, it does not belong to philosophy, but to empirical science (Carnap 1934c, 6)

Here Carnap seemingly endorses the left-wing's project, the incorporation of substantive philosophical claims into the empirical sciences, whilst still maintaining that philosophy exists as a separate, formal discipline concerned only with logic. Carnap may seem to be advocating a strange mixture of naturalism and formal a priori philosophy. But this again is a misinterpretation. That same year, Carnap clarifies that one should not be confused by the 'connotations' of the term "philosophy" as 'a field different from the "ordinary" sciences' (Carnap 1934a/1987, 46). 'He who wishes to investigate the questions of the logic of science must, therefore, renounce the proud claims of a philosophy that sits enthroned above the special sciences' (Carnap 1934d/2001, 332). By this point, Carnap's terminology hides his affinity with Neurath, and Carnap himself seems to have soon become aware of this. As his work progresses, his terminology moves ever further from the language of traditional philosophy, and towards Neurath's.

Only a year later, Carnap makes his position much clearer: 'Apart from the questions of the individual sciences, only the questions of the logical analysis of science... are left as genuine scientific questions. We shall call this complex of questions the *logic of science*' (Carnap 1935, 279). The logic of science is no longer equivalent to but replaces '*the inextricable tangle of problems which is known as philosophy*' (Carnap 1935, 279). He now considers the retention of the term "philosophy" too misleading; the term is 'heavily burdened' and might distract from what the logic of science really consists in (Carnap 1935, 280). By 1938, Carnap has given up on "philosophy" altogether, and talks instead of the 'theory of science' (Carnap 1938/1991, 393). He argues that the 'task of analysing science may be approached from many angles' (Carnap 1938/1991, 393). One approach involves 'investigations of scientific activity... [through] history, psychology, and methodology of science... Theory of science in this sense will be dealt with at various other places in the *Encyclopedia*' (Carnap 1938/1991, 393). This is a direct reference to the project of naturalized epistemology being pursued by Neurath and Frank.

Carnap goes on to make the distinction between science as an activity and science as a body of knowledge (Carnap 1938/1991, 393). The formal investigation of the linguistic expressions of science constitutes the logic of science. He still considers this a purely formal subject. What has become clear is that he now sees the theory of science as having two parts; a formal one which logically analyses the language of the body of knowledge that is science, and an empirical project which studies the activity of science in practice. At this point, logic of science and Neurath's empirical study of science have become two separate parts of a bipartite metatheory of science, to use Uebel's term (Uebel 2007d, 435; 2015). Both parts fall within unified science; Carnap's logic of science occupies the analytic half of Hume's fork, whilst Neurath's pragmatics of science occupies the empirical half. But both parts of the metatheory are part of unified science; there is no place left for philosophy.

As noted, a detailed analysis and exegesis of the early Carnap is beyond our current scope, so for the sake of argument, I will take him as having been far closer in spirit to Schlick than Neurath was, Neurath having remained largely unchanged in his beliefs or attitude. Neurath was a naturalist from the outset, his key themes and theses are apparent from very early in his work. He simply develops them more fully over time. Carnap, however, changes significantly. And this change amounts to a transition towards epistemological naturalism, under Neurath's influence. Whether Neurath simply teased out attitudes and ideas that Carnap already somewhat agreed with, or whether he won him over through polemic, Carnap's intellectual development in the 1930s is a movement into at least broad agreement with Neurath which leaves room for disagreement over significant details. This convergence of Carnap and Neurath results in the bipartite metatheory, in which both recognise the works of the other as equally legitimate, equally important and mutually complementary elements of the metatheoretical study of science, from within unified science. Carnap's formal work and Neurath's empirical works are not separate projects. As with the delineation of the other disciplines within unified science, between chemistry and physics for instance, this is a division of labour, not a disagreement. The relationship between Carnap and Neurath is summed up nicely by Carnap; 'He who wishes to investigate the questions of the logic of science [Carnap]. . . must realize that he is working in exactly the same field [metatheory of science] as the scientific specialist [Neurath], only with a somewhat different emphasis' (Carnap 1935, 332).

Perhaps, one may argue, Carnap's changing use of language is nothing more than that, and we should not read too much into it? To infer naturalism simply from choice of words would, after all, be a leap. It is also, as Creath notes, surprising 'how little Carnap gives us in the way of sustained discussion' of empirical methodology (Creath 1992, 156). There is not the wealth of discussions present in the work of Neurath. But in his limited discussion of empirical methods, Carnap expresses significant agreement with the left-wing of the circle on a number of key issues that, as we saw above, lead to naturalized epistemology. Unlike Schlick, Carnap understands and endorses unified science in exactly the same broad sense as the rest of the left-wing: 'the thesis defended in this paper [is] that science is a unity, that all empirical statements can be expressed in a single language, all states of affairs are of one kind and are known by the same method' (Carnap 1932c/1995, 32). He even uses an argument for the unity of scientific language premised on the combination of knowledge from different disciplines required to make predictions about the construction of automobiles, which is identical in spirit to Neurath's forest fire argument.²⁸

²⁸ 'The important fact is that very often a prediction cannot be based on our knowledge of only one branch of science. For instance, the construction of automobiles will be influenced by a prediction of the presumable number of sales. This number depends upon the satisfaction of the buyers and the economic situation. Hence we have to combine knowledge about the function of the motor, the effect of gases and vibration on the human organism, the ability of persons to learn a certain technique, their willingness to spend so much money for so much service, the development of the

And crucially, Carnap agrees with Neurath about the three epistemological claims on which Neurath's naturalism rests. Carnap is a confirmational holist; '*the test applies, at bottom, not to a single hypothesis but to the whole system of physics as a system of hypotheses*' (Carnap 1934d/2001, 318). He accepts the reality of underdetermination, and that the adoption of any scientific theory is therefore conventional and decision guided; 'That hypotheses... nevertheless contain a conventional element is due to the fact that the system of hypotheses is never univocally determined by the empirical material, however rich it may be' (Carnap 1934d/2001, 320). And Carnap is a fallibilist; 'all rules are laid down with the reservation that they may be altered as soon as it seems expedient' (Carnap 1934d/2001, 320). And as a result, we can see that Carnap too accepts the primacy of scientific practice as a consequence:

whoever desires to investigate [the language of science] must accordingly take into consideration the language which is used in practice in the special sciences... [Logic of science] can only be useful and productive in practice if it has regard to the available empirical findings of scientific investigation (Carnap 1935, 332)

One may still wonder about Carnap's logic of science. Is it not *a priori*? And how does it fit with Neurath's empirical metatheory of science? These questions will be discussed further in subsequent chapters. For now, I must simply indicate briefly that Carnap agrees with Neurath's rejection of the *apriority* of epistemology because he agrees with Neurath's reasons for and means of naturalising epistemology.

1.3.2 Carnap's Approval of Neurath's Naturalism

Carnap is a Humean empiricist like the rest of the Circle. He explicitly invokes Hume's fork in his delineation of the domain of meaningful scientific statements (Carnap 1932c/1995, 61). '*Synthetic sentences are the genuine statements about reality*' (Carnap 1934d/2001, 179). So for him, formal studies don't have content, and don't provide substantive knowledge. But these contentless analytic statements are still legitimately scientific. The logic of science is '*the formal structure theory of the language of science... The logical syntax of the language of science*' (Carnap 1934c, 9). This is concerned only with the '*rules of that language*' (Carnap 1934c, 9). And as he later says, 'syntactical investigation... is indeed a purely mathematical task' (Carnap 1935, 332). Consequently, the sentences of the logic of science 'lie inside the boundary drawn by Hume' (Carnap 1934a/1987, 48). So despite logic of science being an *a priori* undertaking, its analyticity allows it to be consistent with empiricism, and therefore included within unified science. The *apriorism* Carnap opposes claims there to be substantive, synthetic *a priori* knowledge (Carnap

general economic situation, etc... For very many decisions, both in individual and in social life, we need such a prediction based upon a combined knowledge of concrete facts and general laws belonging to different branches of science' (Carnap 1938/1991, 404).

1930/1959, 143). Rejecting this apriorism does not mean rejecting the possibility of a priori knowledge outright. But like Neurath, the mature naturalist Carnap takes the battle with the metaphysician as won. He doesn't spend time arguing for the naturalist position explicitly, or arguing against traditional apriorist philosophy. His aim is not 'to defend the principle of empiricism against apriorism', the principle is simply taken for granted (Carnap 1936–37, 2). But it is at least clear that, in line with contemporary naturalism, and in the same naturalist vein as the rest of the left-wing of the Circle, Carnap rejects the possibility of substantive non-formal a priori knowledge, whether supposedly philosophical or scientific.

Carnap's response to the sceptic agrees with Neurath's (and Quine's) in effect, although different in articulation. Carnap argues that there is an ambiguity in our use of "real", and our questioning of "existence". How we respond to these questions depends on how we interpret these terms, whether they are considered as external or internal questions. Carnap distinguishes between internal questions, those we ask from within a given language-system, and external questions, which concern the sort of language-system to adopt. To interpret sceptical questions as internal means taking a certain linguistic framework, a specified set of rules for the use of our terms, and asking whether these objects are real given the framework's rule for the use of the terms. As an internal question, questions of reality are answered by empirical investigation, according to rules of confirmation and disconfirmation. 'The concept of reality occurring in these internal questions is an empirical, scientific, non-metaphysical concept' (Carnap 1950, 22). As Carnap argues, when taken as internal questions these questions of existence are relatively straightforward. These are the questions of existence that are relevant to the sciences.

From this, he distinguishes the external question which is

raised neither by the man in the street nor by scientists, but only by philosophers... And it cannot be solved because it is framed in the wrong way. To be real in the scientific sense means to be an element of the framework; hence this concept cannot be meaningfully applied to the framework itself (Carnap 1950, 22–23)

The external question is not a solvable one, there is no answer. In as much as we can question a linguistic framework, there is no correct choice, only a practical decision. Whether we choose to adopt a certain linguistic framework is a matter of convention justified on the grounds of utility. But the skeptic requests a theoretical justification of the choice of our linguistic framework, which is simply impossible. We can give only pragmatic reasons for such choices. This response to the sceptic is compatible with the responses provided by Neurath above.

Carnap's relationship to foundationalism is a hotly debated issue, at least as far as his *Aufbau* is concerned.²⁹ What must be stressed therefore is that the Carnap I am

²⁹ The *Aufbau* adopts a position of methodological solipsism starting from an 'autopsychological basis', attempting to reconstruct knowledge on the basis of 'the given' (more specifically, remembrances of relations of similarity) (Carnap 1928a/2005, 101–102). Consequently, it was for a long time interpreted as providing a logically sophisticated foundationalist phenomenism, reducing knowledge to the given in pursuit of epistemic security. This interpretation was for

concerned with is the Carnap that emerged in late 1932 in “On Protocol Sentences”: a philosopher who rejected his previous approach to epistemology that ran the risk of being regarded as foundationalist (whether he was or was not), a philosopher who endorsed physicalism unreservedly, stressed the principle of logical tolerance that facilitated a plurality of logics and linguistic frameworks, and who took philosophy to concern the provision of “explications”, offering solutions to vexing philosophical disputes through reconceptualizations of the basic terms and not, therefore, of providing a deeper, supra-scientific insight into the nature of things. This Carnap argued that ‘The thing language in the customary form works indeed with a high degree of efficiency for most purposes of everyday life.... This fact makes it advisable to accept the thing language’ (Carnap 1950, 24). The acceptance of a certain language is advisable, on the assumption of certain pragmatic purposes that we desire of the language. This is the same pragmatic instrumentalism as we found in Neurath. Carnap previously highlighted that logic of science can provide either an assertion, analysis of how languages are, or a proposal, a suggestion of how possible languages could be (Carnap 1934c, 15). These proposals were conditional on certain pragmatic concerns. Carnap’s switch from logical syntax to semantics effected no change in this respect. We will return to this in a later chapter.

1.4 The Cooperation of Neurath and Carnap: The Bipartite Metatheory

What the foregoing discussion demonstrates is the convergence of Neurath and Carnap, for the same reasons, on a fundamentally naturalist reorientation of philosophy; the replacement of substantive a priori philosophising with the primarily empirical methods of unified science, the acceptance of the epistemological consequences of underdetermination and confirmation holism, and the rejection of the questions of scepticism and the search for foundational certainty as traditionally understood. This is indicative of a broad and fundamental basis of agreement between Neurath and Carnap. This interpretation supports Thomas Uebel’s ‘bipartite metatheory conception’ (Uebel 2015, 24).

example propagated by Quine (Quine 1951a/1980, 39, 1969a, 74) and Ayer (1936b/2001) Carnap himself even seems to endorse this interpretation (Carnap 1963d, 50). More recent interpretations of Carnap’s project have emphasized its Neo-Kantian inheritance, interpreting the *Aufbau* as an attempt to construct scientific objectivity from a basis of experiential subjectivity (Richardson 1997b; Friedman 1999). Friedman directly challenged Carnap’s autobiographical account of the *Aufbau*, arguing that it mistakes the Vienna Circle’s reception and use of the *Aufbau* for the original intentions behind it (Friedman 1992). See also Friedman (2007); Carus (2007a, Chapter 2–4), Carus (2007b), and Leitgeb and Carus (2020, Supplement A). It has also been shown that the influence of Russell’s *Our Knowledge of the External World* on Carnap is compatible with this Kantian reading of Carnap. See Carus (2007a, Chapter 5) and Pincock (2007).

Uebel argues that the bipartite metatheory is a project of the left-wing of the Vienna Circle, but the most prominent figures are Neurath and Carnap. Uebel's contention is that Neurath and Carnap's mature projects, whether they recognised it or not, ought to be understood as two halves of a broader, shared project. For both, as we have seen, philosophy is replaced by scientific metatheory as a part of unified science. Uebel agrees: 'philosophy formed part of unified science because it was assimilated to science as its metatheory, as a second-order inquiry of itself a scientific nature employing only formal-logical or empirical resources' (Uebel 2015, 24). What Uebel emphasises is that within the metatheory of science there are two equally legitimate methodological approaches, capable of 'fruitful cooperation' (Uebel 2010, 198). The differing methodologies are a result of divergent subject matter. One is a formal-logical investigation of the language of science. This is Carnap's speciality, and Uebel retains Carnap's name for it: the logic of science. The other methodological approach is an empirical study of the practice of science. This is Neurath's area of interest, but Uebel opts for Frank's terminology, calling it the 'pragmatics of science' (Uebel 2015, 24).³⁰ Neurath's preferred name was the more cumbersome 'behaviouristics of scholars' (Neurath 1936d/1983, 160).

We can see the embryo of this bipartite conception in the way Carnap and Neurath describe their respective projects:

If we regard the mass of statements as the result of experiments, travels, certain other behaviour, then we move in the fields of the significant *behaviouristics of scholars, history of science, sociology of science*... If we state: "On the basis of experiments a scholar has replaced an earlier published statement by a new one", this is no statement of the logic of science; but the logic of science can compare this statement with other statements as to its logical content; in the same way it can compare the scholar's earlier statement with his later statement. (Neurath 1936a/1983, 134–5)

These investigations of scientific activity may be called history, psychology, sociology, and methodology of science. The subject matter of such studies is science as a body of actions carried out by certain persons under certain circumstances. . . We come to a theory of science in another, sense if we study not the actions of scientists but their results, namely, science as a body of ordered knowledge. . . We mean by 'results' certain linguistic expressions, viz., the statements asserted by scientists. The task of the theory of science in this sense will be to analyze such statements (Carnap 1938/1991, 393)

Both recognise the difference between a formal study of scientific language and an empirical study of scientific practice, and recognise both approaches as legitimate and valuable.

This metatheoretical self-study of science is assigned both negative and positive tasks. The former are essentially defensive and reactive; cautioning against metaphysics and unmasking nonsense. The most famous example of such anti-metaphysical de-mystification is Carnap's deconstruction of Heidegger's use of the term "Nothing" (Carnap 1932a/1959). But it also involves the study of scientific practice and history, to identify cases of metaphysical thinking being smuggled

³⁰ I will adopt the same terminology from here on.

into, or via, scientific thought. Frank for instance analysed attempts to utilise misunderstandings of quantum mechanics in support of metaphysical doctrines as one example of a broader trend of misusing science ‘in the service of the spiritualistic conception of the universe’ (Frank 1936b/1941, 154). These types of diagnostic analyses of scientific history are one of Frank’s great contributions to the scientific metatheory.³¹

On the positive side, the metatheory of science is intended to provide valuable insight into the norms and patterns of scientific practice, and to facilitate the development of new approaches to science. On the formal side, this involves the development and exploration of possible logics and languages, investigations of ‘abstract relations of evidential support’ (Uebel 2012, 118).³² As a conventionalist, there is no one correct form Carnap’s logic must take. By exploring the possible options, we can make pragmatic comparisons over their relevant usefulness. In his later work, as will be discussed in detail in chapter 4, Carnap’s emphasis shifted to the process of explication as a means for supplying clearer and more fruitful concepts to be deployed in the sciences. Carnap saw himself as essentially a creator of formal tools for the scientist. Sometimes these tools are developed in response to a specific problem, sometimes the creation of a tool generates the recognition of as-yet unrecognised problems.

The pragmatics of science, in its positive guise, is intended not only to supply tools, but to supply insights and understanding as the basis for instrumental methodological reflection. At the most basic level, reflection on which methods have been successful and unsuccessful before is valuable for considering which methods to adopt in the future. But this is pragmatics in its most simplistic form. Such studies typically take the form of history and sociology, and focus on science as a concrete practice, occurring in the world in a specific socio-historical context. Zilsel’s study of the origin of modern science for example emphasises the role of social historical and economic factors on the emergence of the modern scientific method (Zilsel 1942). As will be explored in more detail in the next chapter, such studies of the embeddedness of science within a real-world context have facilitated the realisation that science is not a dispassionate and objective search for truth. The (as we will see, inappropriately named) “post-positivist” turn toward socio-historical understanding of science, and the impact this has had on the current understanding of science, most obviously as concerns the role of extra-scientific factors in the scientific process, are a realisation and vindication of the empirical metatheoretical approach that Neurath and Frank advocated.

Having established the convergence of Neurath and Carnap towards a naturalistic metatheory of science, and their decisive break with the traditional philosophical orientation, we will now turn to each part of the bipartite metatheory in turn. We

³¹ For other examples, see Frank (1934/1941), Frank (1935/1941), and Frank (1941b/1949). Frank’s recently published unfinished manuscript contains a chapter on ‘The Sociology of Metaphysical Interpretations’ (Frank 2021, Chapter 4).

³² See also Uebel (2013a).

will begin with Neurath's philosophy of science, to flesh out both the form his naturalised epistemology takes, and the form that the pragmatics of science takes. From there, we can move to the most developed example of Neurath's idea of scientific metatheory, his theory of protocol statements.

Chapter 2

Neurath's Epistemology of Science



In the previous chapter I argued that Neurath is correctly categorised as an epistemological naturalist, but noted that his naturalism does not always conform to expectations of typical contemporary naturalists. To specify the distinctive aspects of his naturalism, and the distinctive epistemology of science that resulted from it, is the goal of the current chapter. As in the previous chapter, Frank is utilised as an important companion to Neurath's philosophical approach. Neurath never made any attempt to provide a systematic explanation of his philosophy of science. His writings are disparate and brief, and often the most insightful and intriguing ideas are implied, assumed, or simply stated and moved beyond. It is not always clear that a consistent, cohesive philosophy undergirds Neurath's work. The purpose of this chapter is to tease out the systematic account of Neurath's epistemology of science that he never provided himself.

2.1 Neurath's Empiricism

2.1.1 *Empiricism and Verificationism*

Neurath consistently and proudly referred to himself as an empiricist and would chastise both opponents and allies for failure to live up to his exacting standards. His empiricism and rejection of metaphysics were two sides of the same coin, much as they were for the other members of the Vienna Circle. But Neurath applied these criteria stringently even by the standards of his fellow logical empiricists. During meetings of the Vienna Circle, his interruptions to accuse other members of metaphysics were so frequent that Hahn suggested he simply say "M" for brevity, to which Neurath suggested it would be quicker still to say "Non-M" when they were not engaging in metaphysics (Neurath, 16/06/1945, in Tuboly and Cat 2019, 640). Despite this stridence, he appears to share something like the common

understanding of empiricism. 'In harmony with consistent empiricism, one tries again and again to refer back to 'experience'' (Neurath 1934/1983, 101). Theories must be sensitive to, grounded in and testable by experience. But beyond this (as is typical of Neurath) we are never given a definition, or analysis of how he understands "empiricism".

A useful initial orientation can be gained from Bas Van Fraassen's account of empiricism. Van Fraassen describes the empiricist tradition in philosophy as one of 'recurrent rebellions against a certain systematizing and theorizing tendency in philosophy... against the metaphysicians' (Van Fraassen 2002, 36). He categorises empiricism according to how successfully this attack is maintained. Van Fraassen considers a naïve empiricism one whose defining commitment is a statement that is either true or false. Russell for example 'defined [empiricism] as the assertion "all synthetic knowledge is based on experience"' (Russell 1948/2009, 437). Van Fraassen rejects naïve empiricism as unable to maintain the attacks on metaphysics; no factual thesis can simultaneously act as the basis for the rejection of metaphysics and itself resist the empiricist critique (Van Fraassen 2002, 45–46). Instead, empiricism should be understood as a stance; a set of attitudes, approaches, inclinations and commitments. This may be expressed through beliefs, or even presuppose certain beliefs, but cannot be captured by or equated with a specific claim.

Van Fraassen's account nicely captures Neurath's empiricism.¹ Neurath describes himself as an 'advocate of the empiricist attitude' (Neurath 1936c/1983, 146). As early as the manifesto, the Vienna Circle describes itself as 'characterised not so much by theses of its own, but rather by its basic attitude' (Carnap et al. 1929/1973, 305–6). Neurath's description of the scientific world-conception also shows clear affinities; 'the interconnection of empirical facts, with systematic testing by experiment, the joining of the individual into the texture of all sequences of events, and the uniform logical treatment of all trains of thought' (Neurath 1930/1983, 42). The most important aspect of the empiricist stance is, Van Fraassen claims, a positive attitude to science; specifically, a positive attitude to the methods of science as 'a paradigm of rational enquiry' (Van Fraassen 2002, 63). The lesson the empiricist draws from science is 'how to give up our beliefs... without succumbing to despair, cynicism, or debilitating relativism' (Van Fraassen 2002, 63). This, as we will see, is the insight of Van Fraassen's that most significantly resonates with Neurath's empiricism.

An aspect of Neurath's empiricism that one might expect to have been discussed by now is verificationism. After all, verificationism is typically treated as the definitive philosophical commitment of the logical empiricist. But understanding Neurath's empiricism requires rejecting the caricatured picture of a monolithic logical empiricism built around verificationism. Its undeniable that verificationism was a key commitment of most members of the Vienna Circle in the early 1930s,

¹ This is no coincidence. Van Fraassen sees the logical empiricists as sharing his conception (Van Fraassen 2002, 49).

and was taken up by many adherents and enthusiasts abroad. But there was a variety of verificationisms within the Vienna Circle.² Uebel has made a key distinction between two different (and not necessarily exclusive) types of verificationism. The first, takes verificationism as a theory of meaning for empirical propositions: a proposition's meaning is the method of its verification. The second takes the possibility of verification as a criterion for meaningfulness: if a statement can't be verified it has no meaning. Uebel calls these 'meaning-constitutive [V-TOM] and significance-criterial [V-CRIT]' respectively (Uebel 2019, 2). Meaning-constitutive verificationism is the stronger claim; whilst V-TOM implies V-CRIT, the opposite is not the case. Uebel acknowledges that the members of the Circle did not always clearly recognise this distinction, and frequently combined the two (Uebel 2019, 4).

Meaning-constitutive verificationism originated with Wittgenstein, and was taken up by those in the Circle closest to him, Schlick and Waismann:³

The sense of a proposition is the way it is verified. Sense itself is a method of verification; that method is not a means, not a vehicle (Waismann 1930/1979, 244).

Significance-criterial verificationism originates in Carnap's "*Scheinprobleme*", and is taken up by Hahn and Frank:⁴

If an (ostensible) statement does not express a (conceivable) state of affairs, then it has no meaning; it is only apparently a statement (Carnap 1928b/2005, 325)

Uebel consequently calls Wittgenstein and Carnap the 'two founts of Vienna Circle verificationisms' (Uebel 2019, 4). Whilst it would be an oversimplification to equate these differing conceptions of verificationism with the split between the left and right wings of the Circle, there is an unmistakable overlap of alignments on the two topics.

Given the understanding of the left-right split within the Circle from the last chapter, one would expect Neurath to fall on Carnap's side. This is the case. Neurath never adopted a meaning-constitutive notion of verification. Not only was he suspicious of anything emanating from Wittgenstein, but V-TOM is incompatible with Neurath's epistemological commitments. As will be discussed further below, Neurath was a confirmation-holist:⁵

The correctness of each statement is related to that of all the others. It is absolutely impossible to formulate a single statement about the world without making tacit use at the same time of countless others (Neurath 1913/1983, 3)

The meaning-constitutive notion of verification is incompatible with this, resting as it does on the discernibility of the specific meaning of and means of testing a single sentence in isolation.

² See Hempel (1950/1959).

³ See also Schlick (1932a/1979).

⁴ See Frank (1935/1941, 149), Frank (1936c/1949, 205), and Hahn (1933/1987, 40–41).

⁵ He may also be a holist about meaning. See Neurath (1921/1973, 198–99).

Neurath adopts a form of V-CRIT, though he rarely discusses it as such. Like Carnap, Neurath came to reject talk of verification and falsification for their 'absolutistic flavor', preferring to talk in terms of confirmation and disconfirmation (or "shaking") (Neurath 1937a/1983, 180). Neurath never gives the verification principle an explicit role in his work. But this does not preclude him from tacitly adopting a version of verificationism: 'A statement which cannot be controlled is a *thesis devoid of sense*' (Neurath 1931a/1983, 48). From this, we can discern the form of Neurath's verificationism. Neurath refers to metaphysical claims as 'isolated statements', statements that 'cannot be joined with our scientific statements to enable the creation of further scientific statements' (Neurath 1937e/1983, 208 n.). Although he does not refer to it as such, this implies a criterion of significance for statements; that a significant statement must be capable of integration into our web of beliefs to allow the derivation of further scientific statements.⁶ This is a criterion of a fundamentally pragmatist sort. Compare this with William James' characterisation of the principle of pragmatism:

what sensations we are to expect from it, and what reactions we must prepare. Our conception of these effects, whether immediate or remote, is then for us the whole of our conception of the object, so far as that conception has positive significance at all (James 1907/2000, 25)

This is however a minimum condition for cognitive meaning. What is needed in addition is the possibility that the resultant collection of statements allow the derivation of observable consequences, expressible in the form of protocol statements. It is testability via protocols - however indirectly - that is the heart of Neurath's empiricism. '*These crude sentences with their many indeterminacies are for us the begin-all and the end-all of all science.*' (Neurath 1937d/1987, 141).

Crucially then, Neurath's verificationism is not the verificationism of Schlick or Wittgenstein. The picture of Neurath's empiricism that we are left with is of an attitude, not a claim, and one that is in some ways much closer to the pragmatist tradition than orthodox (British) empiricism.

2.1.2 Science and Pseudo-Rationalism

Whilst we will see below how Neurath understands science, his attitude towards it is also crucial to his broader philosophy. Whilst Neurath admires the sciences, he refuses to idealise them. This may be a consequence of the maxim of the primacy

⁶Carnap gives a more technical explication of 'empirical meaningfulness', but his informal characterisation has clear affinities to Neurath's: 'that a certain assumption involving [the term] makes a difference for the prediction of an observable event. More specifically, there must be a certain sentence S_M about M [the term in question] such that we can infer with its help a sentence S_O in L_O ' (Carnap 1956a, 49). For more detail, and an extensive defence of the formal acceptability of this 1956 criterion, see Justus (2014).

of scientific practice. Despite his advocacy for scientific methods, Neurath never takes this method to be set in stone, nor does he attempt to portray science as an enterprise of purely objective methods, completely divorced from social, historical and material context or the personalities and behaviours of scientific practitioners. Science for Neurath is not the perfect form of inquiry, it is not a dispassionate search for objective and eternal truths, deduced according to a universal methodology. Such pictures of science are examples of the great bogeyman that haunts Neurath; pseudo-rationalism.

Pseudo-rationalism is not a doctrine or a theory, but an attitude or impulse that Neurath sees as a cause of much of traditional philosophy's excesses, and one which can easily and unconsciously be acceded to.⁷ The pseudo-rationalist impulse is manifest primarily as a desire for absolutes and finality where there can be none. Neurath diagnoses Leibniz's pseudo-rationalism as his pursuit of '*the* system of science and *the* logical key for it' (Neurath 1938/1971, 16). Pseudo-rationalism seeks *the* method, *the* answer, *the* *ultimate* cause, *the* *true* form, even in cases in which such answers are beyond our capacities as limited epistemic agents. As applied to science, the rejection of pseudo-rationalism means rejecting the notion of a universal and objective scientific method. Specifically, Neurath argues that we must abandon the desire for a 'univocal calculus' for decision-making (Neurath 1946d, 80).⁸

Science is permeated by unclarity and underdetermination, guided inevitably by pragmatic decisions rather than logical deductions. But, Neurath maintains, it is still the best we have. And, crucially, if we recognise the constructive role practitioners of science play in the creation of the norms, methods and values of science, it can be made better. This leads to what I am calling methodological constructivism, which I argue is the crucial feature of Neurath's thought and the unified science project that he advocated. This constructivism, concerning not merely the theories but the procedures of science, is eminently suited to complement Carnap's approach to philosophy as language engineering, considered in detail in Chaps. 5 and 6. To approach a first characterisation we must briefly review Neurath's views on language and science, from which his constructivism emerges.

⁷ Hahn's diagnosis of 'world-denying philosophy' captures much the same thing (Hahn 1929/1980, 7).

⁸ The rejection of pseudo-rationalism also expresses itself directly in Neurath's social and economic thinking. Neurath's arguments in the socialist calculation debate rejected Von Mises assumption that rational economic decision making requires a singular universal measure of value to allow for determination of an optimum distribution of goods. Neurath instead argued that the multiplicity of measures of value cannot be reduced in order to reach a singular correct outcome via a decision-calculus. There is no singularly rational optimal choice. See Neurath (1912/1973), Nemeth (1982b/1991), O'Neill (2007), and Uebel (2007e).

2.2 Language, Knowledge, and Science

2.2.1 *Science as a Social Instrument*

Making predictions 'is what all of science is about' (Neurath 1931b/1983, 53). Successful science provides us with predictions, which allows manipulation and anticipation of the world around us. Scientific practice involves 'making well-arranged descriptions, finding correlations, and preparing predictions' (Neurath 1944, 1). No attempt is made by Neurath to define science in terms of its truth, (or approximation to the truth) or the successful reference of scientific terms to entities in the world. Truth and correspondence are not the criteria by which Neurath understands science's success. The statements of unified science are simply 'tools for successful prediction, that is for life' (Neurath 1931c/1983, 62). Neurath sees science's value as primarily instrumental, as a means to an end; science is a collective undertaking aimed at improving the conditions of life.

People use the results of scientific work in everyday life much more than formerly, whether they travel by rail, undergo an operation in hospital, eat fruit from planned cultivation or use a fountain pen (Neurath 1937a/1983, 190)

Science is successful insofar as it fulfils the goals it pursues. Scientific metatheory therefore is successful in so far as it increases the predictive and explanatory capacity of the sciences by providing or improving the tools at the scientist's disposal. In the case of Neurath's empirical metatheory, this involves providing science with self-understanding, by which self-improvement is made possible.

What is importantly absent from Neurath's vision of science is anything approximating a correspondence-theoretical scientific realism. By this I mean the sort of scientific realism which demonstrates a strong commitment to representationalism, for which the goal of science is a veridical picture of reality. Scientific theories are true (or approximately true, or truth-like) and the entities they posit refer to elements, often unobservable, of the external world. The relationship of correspondence to or representation of the external by the theoretical-linguistic is central. Juha Saatsi expresses this when he describes the success of science as a matter of 'how well our theories latch onto unobservable reality' (Saatsi 2019, 613). Stathis Psillos claims that the realist 'takes scientific theories at face-value, seeing them as truth-conditioned descriptions of their intended domain, both observable and unobservable' (Psillos 1999, xvii). Ilkka Niiniluoto expresses this realism in its most succinct form, arguing that truth, being 'a semantical relation between language and reality', 'is an essential aim of science' (Niiniluoto 1999, 10). These forms of scientific realism are as much ontological as epistemological theses.

Whilst imagery and metaphor are not always precise, they can be illuminating. Where correspondence realists typically capture science as picturing or mirroring, Neurath opts for a more practical metaphor:

Imagine craftsmen who are building a settlement, with a chest of drawers full of instruments, only part of which are well arranged and the usage of which is only partly known by them; imagine that from behind new instruments are continually put in the drawers, that some instruments are modified by unknown people and that the craftsmen learn to use some

of the old instruments in a way hitherto unknown, and now imagine further that the plans of our craftsmen dealing with the building up of settlements are changing too. That resembles to some extent the situation of our scientists (Neurath 1941/1983, 215)

The obvious element of this metaphor is its instrumentalism. Less obvious but equally important however is that Neurath's metaphor takes sciences as imbedded in the concrete practices of human beings in the physical world. Like everything else, science lies within the 'earthly plane' (Neurath 1928/1973, 295). Compare this to the picture of science as pursuing a verisimilitudinous picture of reality; science is abstracted from the people who do science, to a theory as it somehow confronts reality. For Neurath science is not abstractly theoretical, but a human endeavour undertaken under specific conditions and with certain tools.

2.2.2 *The Indispensability of Public Language*

Science is also a public enterprise. In fact for Neurath, any intellectual undertaking is inherently social; 'it is not a single individual who can really think new notions through to the end, but only whole groups or generations. Thinking, too, is a collective occurrence' (Neurath 1928/1973, 293). With science as a social project the scientific language, as a form of public discourse, must be publicly accessible. But Neurath goes further, arguing that inter-subjective public accessibility is necessary for any responsible and potentially reflective language. If one wants to use language, 'he must make use of the "*inter-subjective*" language' (Neurath 1932a/1983, 96). A statement is either publicly accessible, or it is unacceptable for use in science (remember that Neurath's notion of science involves essentially all legitimate knowledge).

This position is grounded in Neurath's private language argument, made publicly in two papers from 1931 and 1932 (Neurath 1931b/1983; 1932a), which grew out of discussions over physicalism within the Circle.⁹ The argument begins from the basis that for language to be used meaningfully and coherently requires a guarantor of constancy of use. Without a guarantee that words and concepts are being used consistently, the communicative function of language is insecure. Neurath argues that this guarantee cannot be provided by a person in isolation.¹⁰ A single person's

⁹ Uebel shows that Neurath first presented a version of the argument to members of the Circle in a private 'Discussion About Physicalism' in March 1931 (Uebel 2007d, 222). Uebel also suggests the influence of the pragmatist James Dewey, Neurath's mentor Ferdinand Tönnies, and Marx's argument in *The German Ideology* on Neurath's argument (Uebel 2007d, 413–417).

¹⁰ Wittgenstein's more famous private language argument has certain similarities to Neurath's (Cat 2019, §2). There is however little evidence of direct influence in either direction. This interpretive issue is complicated by the controversy caused by Wittgenstein's claim to be the originator of "physicalism" (against Carnap, who himself credits Neurath). For a disentangling of these controversies, and an account of the differing physicalisms and private language arguments of Carnap, Neurath, and Wittgenstein, see Uebel (1995).

mental faculties are not sufficient to guarantee the consistency of a private language, because there is no means of checking whether they are right or wrong. We have no privileged epistemic relationship to our own past statements, but rely simply on memory:

Let us assume a man who “has lost his memory” and “his eyesight”, and at the same time learns afresh to read and write. His own notes of earlier times, which he can read with the help of special apparatus, will for him be those of “another” person as much as the notes of any contemporary (Neurath 1932a/1983, 96)

And memory is faulty. A person might have absolute conviction that their usage of terminology is constant, but has no internal means to verify or falsify this. In isolation we cannot even reliably check, let alone guarantee, this constancy. So constancy of use can only be guaranteed through the natural mediation that comes with participation in a linguistic community. The requirements for self-communication are therefore exactly the same as those for communication in general; inter-subjectivity. ‘The Robinson of yesterday and the Robinson of today stand in precisely the same relation in which Robinson stands to Friday’ (Neurath 1932a/1983, 96). Inter-subjectivity facilitates constancy of use because the linguistic community provides the basis for checking and regulating this continuity. As a tool of communication therefore, language cannot be private. One may wish to grant, perhaps, that these private statements had some significance in the moment.¹¹ But this would not be the significance possessed by a language. And even for the same person at a later date, this content would be lost; it would not be communicable. It certainly could not play any role in science, as communicability is required for testability.

2.2.3 *Physicalism*

The language that guarantees inter-subjectivity, and thereby facilitates communication and the possibility of science, must be physicalist. Importantly, Neurath must not be understood as advocating physicalism as the term is used in contemporary philosophy. His thesis is not an ontological one. Rather, by “physicalist” Neurath means a language of ‘statements contain[ing] references to the spatio-temporal order’ (Neurath 1931b/1983, 53). He also intended physicalism to bring about a demystification of language, in part as a response to the influence of Wittgenstein (Neurath 1931b/1983, 53). The Vienna Circle were part of what is sometimes

¹¹ Neurath would probably not make this concession. As can be seen from Rose Rand’s notes, Neurath took these kinds of inexpressible experiences to be highly problematic, bordering on incoherent: ‘Can we even say such a thing? Is it possible even to express such a fact by a sentence? If not then in such cases we take the position of solipsism and cannot communicate. . . If it is held to be permissible to speak of such states, then this means that inexpressible experiences as such. . . have a place in science’ (Neurath, recorded in Rose Rand’s notes, in Stadler 2015, 90–91). It was these worries that lay beneath Neurath’s criticisms of Schlick’s affirmations.

called by historians of philosophy the “linguistic turn”. And while Neurath was an enthusiast of this reorientation, he was also keen to ensure that this focus on language did not involve a reification and externalisation of language. As Carnap recalled, ‘Neurath emphasized from the beginning that language phenomena are events *within* the world, not something that refers to the world from outside’ (Carnap 1963d, 29). Language is a part of the world, and a bifurcation between language and reality is not only illegitimate but the root of metaphysical mystification. Rather than requiring a confrontation between language and reality, Neurath emphasises that ‘language itself is a physical formation whose structure, as physical arrangement... can be discussed by means of the very same language’ (Neurath 1931b/1983, 53). Against Wittgenstein’s *Tractatus*, Neurath maintains that language can talk about language. This in itself is not controversial. But Neurath also made the further claim that ‘*statements are always compared with statements*’ (Neurath 1931b/1983, 53). Language is never confronted with the world, it is already a part of it. Consequently, ‘it is impossible to turn back behind or before language’ (Neurath 1931b/1983, 54).

So far though this is abstract. How did Neurath envision this physicalist language in practice? What would such a language look like? Neurath referred to the physicalist language he envisioned for unified science as the ‘universal jargon’ (Neurath 1932a/1983, 98) or ‘universal slang’ (Neurath 1937c/1983, 180). This universal jargon is not an entirely new artificial language like Esperanto, but a ‘purified everyday language’ (Neurath 1937c/1983, 180).¹² But this emphatically does not mean across-the-board replacement of everyday language with technical terminology. Our everyday language is for the most part both functional and physicalist, describing occurrences in time and space. The universal jargon results simply from the purging of meaningless terms and the addition of some technical scientific ones. We ‘must be prepared to speak cautiously whenever we think it necessary, but we should not adopt a pedantic attitude throughout’ (Neurath 1941/1983, 214). The universal jargon is not eliminative; a more precise scientific means of describing events can coexist with a more familiar everyday description. The addition of “H₂O” to our vocabulary does not preclude continuing to refer more colloquially to “water”. Neurath makes no complex demands or requirements on the way we speak or the way language functions. He simply desires a more strictly empiricist and scientifically literate version of the everyday language in which we already communicate.

But the universal jargon is not static. Neurath rejects the possibility of a final, completed language of clear and precisely defined terms as a pseudo-rationalist fantasy. He argues that the very notion of a final language is delusional, that all language must be recognised as provisional, and that all languages are inevitably permeated with a degree of indeterminacy. This indeterminacy comes primarily from *Ballungen*, the ‘[i]mprecise “verbal clusters”’ that organically develop within natural languages (Neurath 1932a/1983, 92). These *Ballungen*, possessing a multi-

¹² Strictly speaking there will be multiple Universal Jargons, corresponding to natural languages, capable of inter-translation (Neurath 1946a/1983, 234).

plicity of meanings and uses, mean we cannot be assured of shared usage. This is perpetuated by our limited cognitive capacities, and consequent inability to survey all our knowledge at once: 'our experience in this is like that of a miner who at some spot of the mine raises his lamp and spreads light, while all the rest lies in total darkness' (Neurath 1921/1973, 198). The emergence of *Ballungen* requires perpetual vigilance, because the growth of science alters and adds to the scientific vocabulary alongside the organic development of natural language. 'Universal Jargon will always be in the making, just as our life and our sciences' (Neurath 1941/1983, 214).

'The establishment of cross-connections [between special sciences] is in close relationship with the question of unity of terminology, with the creation of a "universal jargon"' (Neurath 1936c/1983, 155). This interconnectedness facilitates practical collaboration. For some, the fragmentation into increasing numbers of specialisations has appeared to pose a challenge to the unity of science thesis. Frank compares this worry to the Biblical story of the tower of Babel; with each sub-discipline of the sciences using a different technical vocabulary, collaboration on the project of unified science is made impossible (Frank 1947, 162). As Frank argues, the increasingly specialist nature of science is true, but this is all the more reason to emphasise the importance of unity, and train scientists to recognise the capacity for cooperation and communication such that they can 'transcend the limitations of the traditional departments' (Frank 1950, 59). Neurath argues that the creation and maintenance of a universal jargon is a key aspect of this, as it provides a medium for communication in which the findings of different sub-disciplines can be shared, and in which collaboration between specialties is facilitated. Neurath does not propose that the unity facilitated by universal jargon is new. He thinks that linguistic unity is already the basis for communication, and is 'common to human beings, past and present, all over the world' (Neurath 1946a/1983, 233). This unity is in evidence in apparently mundane examples of the same terms being used in different scientific fields. Neurath gives the example of the word "man", which is equally applicable in biology, mechanics or sociology (Neurath 1936d/1983, 159–60). And this linguistic unity functions as a basis for practical unity, as we saw in the previous chapter in Neurath's example of the forest fire.

It is the possibility of communication in this common language of unified science to facilitate action that is the true criteria for unity. And how is this unity achieved? Only 'historically, by special decisions or by life on a common social and technical basis' (Neurath 1935a/1983, 115). Whilst it is probable (perhaps all but inevitable) that the practical necessities of life will impose a certain degree of unity on our language, Neurath recognises the possibility that it will not do so fully. And in this circumstance, we must decide to impose this unity on our language, to provide the most useful and successful form of science we can. The universal jargon is simply an attempt to actively further the pre-existing unity. Here we see the first glimpses of Neurath's radical voluntarism.

2.2.4 *The Encyclopedia and Encyclopedism*

Linguistic unity forms the most basic cross-connections between the different branches of unified science. But Neurath intended to go beyond mere theorising, and attempted to begin strengthening these cross-connections in practice. Alongside the annual International Congress for the Unity of Science, the *International Encyclopedia of Unified Science* was a concrete attempt at a practical encapsulation of the unity of science through a series of publications on a wide variety of cutting-edge topics in the sciences. The goal was to bring together thinkers from different disciplines across the world in a spirit of cohesion and cooperation.¹³ The *Encyclopedia* was to serve as a forum for discussion and as a means to integrate ideas, findings and perspectives from different fields.

Whereas the other encyclopedias give a retrospective synthesis, so to speak, this new work will have to show above all in which direction new ways open themselves, where the problems lead, and where, from the point of view of a unified science, unsuspected possibilities can be discovered (Neurath 1936b/1983, 140)

The encyclopedia was designed to be progressive and forward-looking.¹⁴ Neurath's enthusiasm for Encyclopedias would ultimately be reflected in his theorising. He adopted the encyclopedia as the model for science itself.¹⁵

A systematic model of the sciences takes science as constructed from certain atomic premises upon which the edifice of science is built, composed of clear and precise statements, according to fixed methods and standards. Successive theories are understood as gradual approximations towards the eventual goal of a completed science, each new discovery and theory gradually building toward this edifice. A perfect example of this type of thinking is provided by Popper:

All work in science is work directed towards the growth of objective knowledge. We are workers who are adding to the growth of objective knowledge as masons work on a cathedral (Popper 1979, 121)

Neurath emphatically rejects such views as pseudo-rationalist:

we do *not* arrive at “one” system of science that could take the place of the “real world” so to speak; everything remains ambiguous and in many ways uncertain. “*The*” system is the great scientific lie. Not even as an anticipated goal is it a useful guiding thought (Neurath 1935a/1983, 116)

For Neurath, science, as both activity and body of knowledge, has no terminus. There is no completed body of doctrine which renders further work unnecessary. Nor are there certain, guaranteed starting points from which all science starts. There is just a continual, indefinite process of development, revision and refinement in

¹³ On the connections to Diderot's *Encyclopédie* and the Enlightenment, see Dahms (1996) and Tega (1996).

¹⁴ For a sense of the intended scope and scale of the *Encyclopedia*, see Neurath (1938/1971) and Morris (1960).

¹⁵ On the many facets of the *Encyclopedia* metaphor, see Pombo (2011).

which all aspects of science, doctrine and method, are perpetually open to doubt, revision and replacement. In the previous chapter, I introduced Neurath's famous metaphor of the boat as indicative of his anti-foundationalist epistemological stance. But the metaphor is deployed again elsewhere as a picture of science more generally:

Our actual situation is as if we were on board ship on an open sea and were required to change various parts of the ship during the voyage. We cannot find an absolute immutable basis for science (Neurath 1937c/1983, 180–181)

In contrast to Popper's picture of science, which he dubs a '*model-system*', Neurath takes the encyclopedia as both the model and the term for our body of scientific knowledge (Neurath 1935a/1983, 122). An encyclopedia, for Neurath, is an assemblage of 'masses of statements whose connection is only partly systematic' (Neurath 1935a/1983, 122). His encyclopedias include everything from fully developed theories, to stray observation statements; 'the totality of scientific matter now at our disposal' (Neurath 1936c/1983, 146). There are no atomic statements upon which theories are constructed that act as the solid foundation of our systems of knowledge. Even the protocol sentences, distinguished by their role in testing, are open to revision. Rather, we start 'with a full lump of irregularities and indistinctness, as our daily speech offers it' (Neurath 1944/1971, 18). There need not, in fact cannot, be definitive systematic connections between all the statements of an encyclopedia. This is what distinguishes encyclopedism from the conception of science-as-system: 'the anticipated completeness of *the* system is opposed to the stressed incompleteness of an encyclopedia' (Neurath 1938/1971, 21). We simply do not have the cognitive capacity to survey our body of knowledge as a totality to verify such consistency. As Mormann highlights, Neurath allows that certain subsections of science are systematizable and free of contradictions. The encyclopedia therefore allows for some local systematization, but not global systematization (Mormann 1996, 91). But these are islands of clarity. Encyclopedias are not even free of contradiction. Although it would be preferable for all our scientific theories to harmonise, Neurath argues that the history of science demonstrates 'how frequent it is... that two incompatible theories are in use at the same time' (Neurath 1936c/1983, 156). And importantly, in so far as systematisation is possible, it is not as the anticipation of a final system, but an organic developed systematisation; 'No system from above, but systematisation from below' (Neurath 1936c/1983, 153).

An encyclopedia is not fixed, but is constantly added to. These encyclopedias are under constant revision and open to replacement. And, again as a consequence of underdetermination, there is no one encyclopedia. There are always, whether actual or potential, alternative encyclopedias. Although no element of an encyclopedia is immune to revision, Neurath argues that there are elements that exhibit more stability, 'those not very precise [statements] of everyday life (Neurath 1936c/1983, 149). Such terms as 'water' and 'tree' are sufficiently general to survive through various encyclopedias (Neurath 1936c/1983, 150). But this helps to facilitate communication and understanding between different encyclopedias, whether separated in time or geography. The Jargon therefore acts as a common touchstone that allows for communication even amidst a diversity of projects.

But how do we choose between competing encyclopedias?

Since we do not possess a special theory of the importance of these possibilities for the progress of science or life in general or since we cannot assign a rank to these two encyclopedias from the point of view of this theory, we must function as a 'touchstone' ourselves and decide for the one or the other (Neurath 1936c/1983, 157)

We must make decisions.

2.3 Decisionism

2.3.1 Underdetermination and Decisions

Perhaps the greatest departure from the pseudo-rationalist picture of science that Neurath's makes is the centrality of decision-making to his account of science. This position has been called decisionism. It appears in Neurath's earliest works, and retains a key role in his picture of scientific practice. Frequently, in practicing science, 'to make progress at all... we have to choose between several equally possible groups of content statements and do this on the basis of a decision' (Neurath 1934/1987, 103). And such decisions can be made 'on the basis of *extra*-logical factors' (Neurath 1934/1983, 106). This decisionism is a natural consequence of holism and underdetermination. Neurath is a holist about testing, accepting that no sentence is ever tested in isolation, only ever en masse. '*The whole of science is basically always under discussion*' (Neurath 1935a/1983, 118). When a prediction fails, no single statement is uniquely designated as having been falsified. And, he argues, underdetermination is more pervasive than Duhem or Poincaré realised:

Poincaré, Duhem and others have adequately shown that even if we have agreed on the protocol statements, there is a not limited number of equally applicable, possible systems of hypotheses. We have extended this tenet of the indeterminacy of systems of hypotheses to all statements, including protocol statements that are alterable in principle (Neurath 1934/1983, 105)

In cases of inconsistency between theory and observation, it is not only statements of theory but also observation reports that may be revised. No statements are uniquely designated by falsification, but nor are any statements immune from the possibility of revision in light of it. There is then no uniquely sign-posted correct response to an instance of falsification, but a vast array of potential alterations that can be made to the system of statements (both theory and observation). 'The question is to find a pattern of statements we are finally to "accept" tentatively in our arguing' (Neurath, 25/09/1943, in Cat and Tuboly 2019, 596).

How we respond to falsification is consequently a matter of choice. This expansion of the underdetermination thesis has come to be known as the Neurath principle: 'Given the non-agreement between a sentence and the whole system there exist always two possibilities to effect agreement: either alter the sentence to be integrated or alter the system' (Haller 1982a/1991, 121). Confirmation faces a similar issue. Like falsification, confirmation is not uniquely designative. If a

prediction is successful, it is an entire body of theory that is confirmed. However, there will always be the possibility that multiple encyclopedias are confirmed by the same evidence. And how we select between them is, again, a matter of choice. The only way past the impasse created by underdetermination and holism is through pragmatic, non-theoretically determined decisions.¹⁶ 'I tried to convey the insight that we need a kind of "decision" wherever we have to make a "choice," even when we are trying a scientific theory' (Neurath 1946c, 526–527).

This pervasive underdetermination explains why Neurath refrains from speaking in terms of verification and falsification, instead speaking of the confirmation and shaking of theories (Neurath 1935b/1983, 123). Verification and falsification are conclusive.¹⁷ For a falsificationist like Popper, a single instance of a false prediction deduced from a theory requires the rejection of the theory. Anticipating Kuhn's later (and more famous) response to Popper, Neurath rejects this idealised methodology as completely impractical, and at odds with the records of actual scientific practice. As Neurath pithily put it in correspondence with Carnap, Popper 'kills the bird in the egg' (Neurath, 1/04/1944, in Cat and Tuboly 2019, 618). '[I]n spite of all the warnings by Duhem', Popper's falsificationism places the fiction of the *experimentum crucis* at the centre of scientific methodology, making it incompatible with the practical realities of underdetermination in testing (Neurath 1935b/1983, 124).¹⁸

By contrast, confirmation and shaking are never absolute. Even when a statement is remarkably well confirmed, it is not necessary that we accept it. We must still decide to accept it. The same applies to shaking. Any single negative result is never sufficiently decisive to require the rejection of a theory. Nor is there a certain number of such results that would require it. Rather, it is a decision as to what degree of shaking by evidence can be tolerated. Often a theory's utility or a lack of alternatives leads to a shaken theory being retained. Neurath gives the example of Newton's gravitation hypothesis, which continued to be used 'in spite of the fact that for about a hundred and fifty years scholars have felt again and again that there were contradictory and ambiguous elements' (Neurath 1944/1971, 26). And since neither confirmation or shaking are decisive, it is possible for there to be numerous 'statements about which a decision has not yet been made' and therefore are considered neither true nor false (Neurath 1936a/1983, 135). They are neither advocated, nor specifically excluded.

¹⁶ Kitcher distinguishes between permanent and transient underdetermination; the latter being a contingent consequence of currently insufficient evidence, the former being cases in which no amount of evidence will ever be sufficient to prevent underdetermination (Kitcher 2001, 31). But this distinction is irrelevant for Neurath. So long as there is underdetermination in the here and now, we are stuck in a decision position without a calculus. As Neurath rightly argues, no decision, delaying until we know all the facts, is a decision to do nothing.

¹⁷ As Cat has emphasised, verificationism and falsificationism require the use of precise terms, and are therefore rendered impracticable by the ubiquity of Ballungen in natural language (Cat 1995, 239–240).

¹⁸ Defenders of Popper may object that what Neurath rejects is only naïve falsificationism.

2.3.2 *Auxiliary Motives and Decision Procedures*

If we accept that Neurath is right so far, we may well ask what guides these decisions? Are they simply arbitrary? Or does Neurath provide some extra-scientific guide to decision making? The first guiding force for these decisions is necessity. One of Neurath's metaphors drives this message home. He compares choosing between competing encyclopedias to choosing between potential layouts for a railway network (Neurath 1934/1983, 106). We can never spell out conclusively all the potential outcomes of every possible version of such a complicated system. To require a categorically determined "best" conclusion on these terms would leave us like Buridan's ass, with no possibility of deciding between the available alternatives, and therefore unable to act. 'The wish to found action on perfect insight means to nip it in the bud' (Neurath 1921/1973, 158). At a certain point, a decision needs to be made. But this still doesn't tell us how we are to make these decisions.

Whilst Neurath refuses to give necessary and sufficient criteria for these situations, he also emphasises that these decisions are not arbitrary. Unfortunately, Neurath never gives a detailed account of what factors are relevant to these decisions. The apparent solution to such decision-situations are what Neurath calls 'auxiliary motives', which he introduces in one of his earliest papers (Neurath, 1913). An auxiliary motive acts as 'an aid to the vacillating', that allows a decision to be made where the aims in question are not definitive by themselves (Neurath 1913/1983, 4). Surely these are exactly the sort of vacillatory situations Neurath was referring to?

An auxiliary motive 'has nothing to do with the concrete aims in question' (Neurath 1913/1983, 4). They do not provide additional reasons or information; they are not auxiliary hypotheses. Rather, they are intended to allow us to bypass the process of reasoning when in an inextricable Buridan-decision situation; 'ambiguity of approach... can be overcome practically, in the last resort, only by the unambiguity of action, that is of decision' (Neurath 1930/1983, 45). What allows a way out of these situations is not cognition, but conation; will, not reason (Cartwright et al. 1996, 132). Rather than providing a means to win an argument, the auxiliary motive is there simply to put an end to the internal debate. Where enquiry typically aims at the right or the best outcome, an auxiliary motive is simply a means to achieve an outcome. Unlike the pseudorationalist, who presumes to reach the correct decision on the basis of admittedly insufficient evidence, the best thinker 'will think as well as he can, but not refrain from admitting that his insight is too weak, and quietly allow himself to decide by lot' (Neurath 1913/1983, 10).

The purest examples of auxiliary motives are drawing lots, tossing a coin, or rolling a dice. The rationality of adopting an auxiliary motive 'derives from the utility of the practical decision to adopt it' (Cartwright et al. 1996, 135). It is a piece of instrumental rationality. The auxiliary motive in the modern world fulfils the role played by instinct and tradition in the past, but where they were based on 'traditional uniformity', use of the auxiliary motive in the aftermath of rationalism requires 'conscious cooperation' (Neurath 1913/1983, 10). Use of the auxiliary motive is

self-conscious and deliberate. But this also opens it up to being communal, for groups to co-operate and decide together to adopt certain auxiliary motives. This is especially true for science which, as we have seen, is a necessarily communal enterprise.

It can be tempting to see auxiliary motives as operative in all cases of decision making in science according to Neurath. After all, Neurath insists that decisions are pervasive and he has supplied a decision-making tool. Stöltzner for instance frequently seems to imply this, deliberately or not, in his description of Neurath's auxiliary motives (Stöltzner 1996, 113). Okruhlik also seems to interpret Neurath in this way, but she also rightly argues that this conclusion does not follow from underdetermination alone. 'Even if the empirical evidence underdetermines certain decisions, there may nonetheless be good (non-evidential, non-empirical) reasons for preferring one decision to another.' (Okruhlik 2004, 60). Whilst I think it is possible to interpret Neurath in this way, I also think it is wrong. This confusion stems (typically) from Neurath's failure to provide a systematic account of the use of auxiliary motives.

We have only Neurath's 1913 article which provides any real detail on the idea of an auxiliary motive. But the purpose of that article was not to elaborate on the methodological role of decision-making in science. The article is concerned with breaking down a false dichotomy between theoretical and practical rationality. The article argues for three specific points. First that, contrary to Descartes, the necessity of decisions extends to the realm of the theoretical, not just the practical; 'It was a fundamental error of Descartes that he believed that only in the practical field could he not dispense with provisional rules' (Neurath 1913/1983, 3). After all, the 'differences between thinking and action are only of degree' (Neurath 1913/1983, 2). Secondly, Neurath wanted to show that in a Burridan's ass situation, any decision is better than indecision, since "'non-action" is also an action', albeit the worst possible one (Neurath 1913/1983, 2). In such a position, doing anything is better than doing nothing; when lost in the woods, walking in a random direction is better than standing still and arguing. Thirdly, Neurath wants to show that genuine rationalism, in opposition to pseudo-rationalism, involves recognising the limits of human reason. The assumption that all choices are conclusively determined by principled reasons, rather than pragmatic decisions, is pseudorationalism. 'Rationalism sees its chief triumph in the clear recognition of the limits of actual insight' (Neurath 1913/1983, 8). We are at our most rational when we recognise, and therefore do not attempt to operate beyond, our limitations, be these a result of our own cognitive limitations or the limitations imposed by circumstance.

Crucially, Neurath is not arguing that every decision situation is like that of Descartes's wanderers. The situation of the scientist is not typically equivalent to being lost in the woods. Neurath is not claiming that auxiliary motives are always deployed when scientific practice requires a decision to be made. They may only be required extremely infrequently. But the fact that they may sometimes be rationally required is the realisation Neurath is trying to bring about. They are the last resort for an honest rationalist making a decision with insufficient information. 'In the end all our thinking depends on such inadequacies. We must advance, even

without certainty! The only question is whether we are aware of it or not' (Neurath 1921/1973, 159). This passage may be misread as saying that ultimately all thinking relies on the use of auxiliary motives. But the "inadequacies" Neurath refers to are underdetermination and the impossibility of finding secure epistemic foundations, and the consequent impossibility of having definitive outcomes uniquely designated. What Neurath means is that we are onboard the boat; the rational thing to do is recognise and accept this situation and the consequent limitations. His emphasis on auxiliary motives is not to suggest that they are ubiquitous in scientific practice, but to demonstrate that the use of auxiliary motives can be rational, and enforce the realisation that an honest rationalism, involving self-awareness and a consequent recognition of one's cognitive limitations, is preferable to pseudorationalism: 'insight becomes awareness of its own limits' (Neurath 1921/1973, 159). Rationality involves recognising where reason alone is insufficient and honestly embracing this limitation.

The wanderers in the forest have 'no indication at all as to which direction to follow' (Neurath 1913/1983, 10). But this is not typical of scientific decision situations. In many cases, there are indications:

Of course we do not toss a coin (though even that is sometimes better than a pseudo-scientific fairy tale, which tells of *one solution*) but we listen to various 'instances,' and in the end we have to reach a 'decision' not based on a calculus (Neurath 1946a/1983, 235)

Here, I think, my point is clearly demonstrated. In the practice of (to use Kuhn's term) normal science, flipping a coin is not needed. The use of auxiliary motives is a limiting case, not the standard case. They are the last resort for a rational decision. But they are not necessary when we have available to us other factors on which to base our decision. What then are the factors that we have to listen to and weigh when making decisions? These could be many things, but most frequently in science they will be theoretical virtues. And it is the weighing of theoretical virtues in the decision procedure of typical science that prevents the pervasiveness of decisions from rendering science arbitrary. We are now able to return to Okruhlik's criticism. Can Neurath make sense of there being non-determined decisions, not made simply on the basis of auxiliary motives, which can be rationally reached through the weighing of arguments against one another? In the next section, I will argue that he can. But, first we need quickly clarify the relationship between auxiliary motives and the other values relevant to decision making in science.

2.3.3 *Auxiliary Motives, Values, and Theoretical Virtues*

In the secondary literature on Neurath there is, with good reason, much confusion and ambiguity over exactly what constitutes an auxiliary motive. Whilst the previous section has hopefully clarified what an auxiliary motive is, it is still worth briefly clarifying what differentiates them from the values relevant to scientific decision-making. We must then distinguish three things: internal values (or scientific virtues), external values, and "pure" auxiliary motives.

The distinction between internal and external values is comparatively clear in the secondary literature, primarily because it has already been discussed by philosophers of science with regard to questions of scientific theory choice. Longino for instance distinguishes constitutive values, those established according to the goals of science, from contextual values, those independent of the goals of science like social and political values and personal preferences (Longino 1987, 54). For terminological convenience, from now on I will refer to internal values as scientific virtues, and restrict the term “values” to those external to science, such as the moral, political or aesthetic. The theoretical virtues are those which we typically think of scientific theories as evaluated in terms of. There exist various ways of differentiating and categorising these virtues.¹⁹ But they typically involve evidential accuracy (how well a theory accommodates evidence), internal and external consistency (logical consistency both within a theory's own statements, and with the statements of the rest of the encyclopedia), simplicity (fewer theoretical elements), unity (explains and predicts more with fewer theoretical elements) and fruitfulness (generating novel predictions) (Keas 2018, 2761–2762). What differentiates theoretical virtues from external values like political or ethical beliefs or personal taste is that the theoretical virtues ‘are the traits of a theory that show it is probably true or worth accepting’ (Keas 2018, 2761). Theoretical virtues are indicative of or conducive to the success of science qua science. This would not be “truth” for Neurath, but our capacity to predict and control the world around us. External values however are independent of the likelihood of success of a scientific theory.

The confusion in the secondary literature is between external values and auxiliary motives. Okruhlik is explicit about her confusion over Neurath's terminology, questioning whether Neurath intended external values to be understood as a type of auxiliary motive, or separate from it (Okruhlik 2004, 63). Elsewhere, we see a great variety of examples of auxiliary motives from a variety of philosophers. These include ‘human happiness’ (Stuchlik 2011, 192), ‘political agenda’ (Howard 2019, 53), ‘creative burst[s]’ and ‘inductive reasoning’ as examples (Zemplén 2019, 223). Uebel also refers to simplicity, unity and economy as examples of auxiliary motives (Cartwright et al. 1996, 135). Biddle interprets the auxiliary motive as ‘any factor that fills the gap between “insight” and decision; some of these factors might properly be characterized as values, while others clearly are not’ (Biddle 2013, 131). If we allow such broad conceptualisations, every decision in science is guided by auxiliary motives. After all, underdetermination is pervasive, so there will always be gaps between insight and decision. But if decision-making in science and the use of auxiliary motives are coextensive, then the concept ceases to be quite so interesting or useful.

To clear up this confusion, I suggest that external non-scientific values ought not to be taken as auxiliary motives. The term “auxiliary motives” ought to be restricted simply and exclusively to pure decision-making devices. They have no

¹⁹ Keas for instance lists twelve theoretical virtues, of four types (Keas 2018).

other goal than resolving a decision-situation. Perfect examples are coin tosses, drawing straws, or playing eeny-meeny-miny-moe. Their sole function is to select one of a number of options. These sorts of examples are familiar from childhood games, and this is a useful parallel. Such methods are often used to choose a person to start a game, who is “it” in tag for example. And they are used because they seem fair or objective, principally because they are arbitrary. There are no other factors at play. These procedures simply make a selection, and that’s all they do.

What the secondary literature is confused by is the use of external values in scientific decision making, in cases of either theory-choice or theory-change. These values provide reasons for accepting or rejecting a theory. Unlike scientific virtues, these reasons are independent of the goals of science, but they are reasons all the same. Their role is cognitive, not conative. They provide an extra reason by which to win the debate. They are auxiliary hypotheses, not auxiliary motives. Both scientific virtues and external values are types of value, it is their relevance to the function of science that differentiates them. However, auxiliary motives aren’t values at all, they are simply decision-making tools. The only purpose of an auxiliary motive is to resolve a decision-situation, and they do not do so by providing reasons. This is why Neurath refers to them as auxiliary motives, not auxiliary hypotheses. As we saw above, they are not intended to determine a decision through reason but put an end to discussion through an act of will. The way that an auxiliary motive imposes a decision through an act of volition rather than reason is what differentiates it.

Before moving beyond the issue of auxiliary motives, one final caveat must be added to the above. I have so far argued that Neurath does not use the term as broadly as he has commonly been interpreted. However, it is possible that I could be proved wrong on this front by further textual evidence. Even in that circumstance however, I maintain that he should have used the term as I have interpreted it, and that we should too. The above interpretation allows the auxiliary motive to retain the unique role of a purely conative means to escape indecision. Whilst the role of both external values and theoretical virtues involves conation, it is only the pure decision-procedure of the auxiliary motive that is exclusively conative. It therefore functions according to different standards. Using “auxiliary motive” in the wider sense dilutes its significance.

2.3.4 Decisions and Theory-Choice

We can now return to the question raised above. In the absence of a singular methodology of science and a definitive decision calculus, how does Neurath understand scientific decision making without rendering it arbitrary? As we have now established, the role of auxiliary motives in science is more minimal than often claimed. However, it is worth re-emphasising, that if and when recourse to an auxiliary motive is required, it is not arbitrary but pragmatically justified. But if I am right that the use of auxiliary motives in science is not pervasive, then how are decisions made in the practice of normal science?

Frank's discussion of theory-choice begins, like Neurath, from a recognition of the reality of underdetermination. He argues that "agreement with observed facts" never singles out one individual theory' (Frank 1957/2004, 355). And on the other hand, '[i]t has never happened that all the conclusions drawn from a theory have agreed with the observable facts' (Frank 1954, 139). What we typically have with rival theories are two theories that each accommodate and explain a certain amount of observable evidence. So empirical sufficiency alone is an inadequate determinant of theory choice. The obvious next step is to utilise theoretical virtues to decide between theories. But, Frank argues, it is very unlikely that any single theory will maximise all theoretical virtues (Frank 1954, 144). So in cases in which one theory does not clearly have an advantage, which are the norm not the exception, we have to weigh these virtues against each other.

But it is possible for competing theories to be both better and worse than the alternative according to different theoretical virtues. If one theory is simpler but the other is more fruitful then 'one has to make a choice of a theory by a compromise between... criteria' (Frank 1954, 139). And there is no calculus for comparison in such situations. Say the wave-theory is simpler and the particle-theory is more fruitful, we have no decisive criteria to decide whether the wave-theory is more simple than the particle-theory is fruitful. The process of theory choice frequently comes down to a weighing up of the relative merits of different options. A further complication is added to this picture by a later argument made by Kuhn. Kuhn argues for the difficulty of unambiguous determinations in theory choice for two main reasons. Firstly, the theoretical virtues themselves are sufficiently vague to allow disagreement. He gives the example of the competing Oxygen and Phlogiston theories, which can each be interpreted as the more accurate theory, depending on the decision about 'the area in which accuracy was more significant'; accounting for changes in weight or the similarity of the properties of metals respectively (Kuhn 1977, 104). Not only is there difficulty in weighing theories relative to multiple theoretical virtues, but even relative to single theoretical virtues.

But it is not just the scientific virtues that are relevant to theory choice; external factors can play a role too. Neurath anticipated much of the discussion in contemporary post-positivist and feminist philosophy of science in the recognition 'that decisive transformations of scientific praxis are not only determined by the intense thought of a generation of scholars, but in addition by what happens in the life of society, of which scholars form a part' (Neurath 2011, 16). This was again explored in more detail by Frank, who gives various examples of historical cases in which theory choice was influenced by broader societal circumstances, such as the rejection of Copernicus and Galileo's heliocentric theories for their contradiction of Aristotelianism and the problems this posed for Thomist arguments for God, or the rejection of Einsteinian relativity in the USSR for contradicting the dialectical materialist orthodoxy of the state's philosophy (Frank 1957/2004, 354–5). Frank even went further, arguing that 'it is difficult to draw a clear dividing line between strictly scientific and sociological criteria' (Frank 1957/2004, 354).

But for Frank and Neurath, this conclusion was not so much a dramatic revelation as a natural consequence of the left-wing's conception of science. The role of values

in theory-choice is simply a consequence of recognising the instrumental function of science, and the hypothetical nature of all methodological norms. ‘The question is merely which theory is more practical in obtaining a certain purpose’ (Frank 1955, 279). This instrumental view of science also means abandoning the notion of absolute, Gods-eye-view objectivity:

There is no judge in a chair who decides who is nearer the truth. There is no way of ‘impartiality’ or ‘scientific objectivity’, there is no point outside our life, from which we may finally decide what is ‘impartial’ or ‘scientifically objective’ (Neurath 1946b/1983, 243)

However, this leaves room for human-scale objectivity, grounded in intersubjective agreement.²⁰ We can only retain the view of science as totally independent of moral and political values when we mistakenly understand it as ‘a collection of facts or as a picture of objective reality’ (Frank 1954, 143). When we abandon the idealised image of science as a correspondence-realist project, the embeddedness of scientific practice in a broader historical context becomes obvious.

Scientists and scientifically minded people in general have often been inclined to say that these “nonscientific” influences upon the acceptance of scientific theories are something which “should not” happen; but since they do happen, it is necessary to understand their status (Frank 1957/2004, 355)

Here we see how naturalism, in this case the maxim of the primacy of scientific practice, stands precisely at odds with the pseudo-rationalist. The genuinely rational response is to take reality, even its less desirable features, as it is.

But as Kuhn later emphasised in response to the strong programme in the sociology of science, recognising the role of external values in science, and even of auxiliary motives, does not necessitate the conclusion that it is power and interest alone that determine which theory succeeds. Kuhn rightly dismisses this as ‘absurd’ (Kuhn 1992/2000, 110). Neurath and Frank would have done so too. As Kuhn argues, the fact that values are inconclusive, that they do not unequivocally dictate specific choices, is not a failure (Kuhn 1977/1998, 111). There are still better and worse, rational and irrational decisions. When faced with two theories, one of which exhibits all the theoretical virtues to a greater degree, there is an obvious rational choice. This decision is not theoretically necessitated by any calculus, but it is the rational decision all the same. This is the important difference for a Neurathian in this position. Even if a certain decision seems obvious, is the clear rational choice, even if it would be bizarre and counter-intuitive not to make the decision, it is still not necessitated by a theoretical decision-calculus. There is no logically deduced, theoretical requirement to make such a decision. These decisions are, even if rational, choices. Their acceptance is not of the same sort as accepting $[A = C]$, on the basis of $[A = B]$ and $[B = C]$.

all we can ever do is choose between several sentential wholes, where certain elements may tell in favour of one whole and others in favour of another, but without our being

²⁰ For details of such a conception of objectivity, see Douglas (2007).

able to see a way of setting up a kind of "index number" which would make it possible for us to arrange the views expressed in a linear order and then to decide, on the basis of calculation, which view was to be preferred. But this scepticism need not have a crippling or a slackening effect, for in the end it leads us back to our coarse everyday experience (Neurath 1937d/1987, 141)

This picture also has the advantage that it accommodates the possibility of rational disagreement between scientists, even scientists pursuing the same goals and evaluating according to the same values. This is a feature of actual science, and one that is not easily explained by those who endorse a scientific decision-calculus.

Hempel makes a distinction between methodological rationalists and pragmatists that can be usefully applied here. The former believe that scientific theory choice must be governed by hard-and-fast rules, but for the latter maxims are sufficient (Hempel 1983b, 87). Neurath and Frank are methodological pragmatists. What Hempel describes is specifically highlighted by Neurath as a case of pseudo-rationalism: 'The danger of pseudorationalism also appears where the replacement of the decision of the practice of science. . . is believed possible through the calculus of the logic of science' (Neurath 1936a/1983, 136). What this ultimately results in is a demystification of the idea of rationality. Whilst rules are unambiguous and determinate, norms allow for variations and exceptions. But in being non-definitive they are not rendered worthless. Kuhn makes an interesting comparison to everyday maxims that we learn as children like "Look before you leap" and "Too many cooks spoil the broth" (Kuhn 1977/1998, 110). That these maxims do not concretely determine actions in any given circumstances is obvious. But clearly this does not invalidate them. That such phrases persist as pieces of folk wisdom suggests an intuitive understanding that rationality does not necessitate a decision-calculus, and that only in science do our instincts suggest so.

In their discussion of the historicity and value-ladenness of scientific decision-making Neurath and, perhaps to an even greater extent, Frank anticipated much of the post-positivist and contemporary work on the topic, including that of Kuhn and Feyerabend, and the naturalism of those like Laudan.²¹ Despite the framing (discussed in the introduction) of such work as a critical response to logical empiricism, it might be more accurate to view this tendency in the philosophy of science in continuity with the work of Neurath and Frank. Many of their insights have been embraced (though not always knowingly so) and often further developed by contemporary historians and philosophers of science. In some ways, their ideas have even been absorbed into the mainstream of philosophy of science. As Heather Douglas highlights, the relevance of values to science has been conceded to in a significant way. It is nearly universally acknowledged that values guide what she calls 'external' parts of science. For instance, values play a key role in determining the projects scientists pursue (developing sustainable energy sources), the methods used (prohibition on human testing for example) and the use of technology (Douglas

²¹ There may have been some direct influence. Robert Butts recalls a lecture given by Frank, and attended by Kuhn, where Frank described what is now known as Kuhn-loss. See (Butts 1999, 12).

2007, 122). Some contemporary feminist philosophy of science has even begun to recognise their affinity and continuity with some of the logical empiricists.²²

2.4 Methodological Constructivism and the Behaviouristics of Science

2.4.1 *Methodological Constructivism*

We have seen the extent to which decision-making permeates scientific practice according to Neurath, and he was well aware of the discomfort this voluntaristic picture of science leaves some philosophers with:

For those who, in a language alien to us, speak of the idea of the true system of the world, this basic encyclopedia must seem a miserable resignation, a scepticism; whereas we see in it the expression of an activist that we equally meet elsewhere. Starting from the situation in which we live and act we march on as well as we can. And we do not think that we can replace acts by dreams (Neurath 1936c/1983, 157–8)

He acknowledges the pseudo-rationalist impulse that, in the absence of transcendental approval, only unambiguous logical deduction seems a sufficient justification for choosing one alternative over others (Neurath 1935b/1983, 127). But, Neurath argues, he is simply being realistic. The pursuit of knowledge is always undertaken within specific historical circumstances, and working from a certain body of pre-established knowledge and operating with a historically given language. We ‘cannot start from a *tabula rasa*. . . We have to make do with words and concepts that we find when our reflections begin’ (Neurath 1921/1973, 198). We necessarily begin onboard the boat and must always recognise the contingency of our entire body of knowledge. But this realism, this rejection of pseudo-rationalism, need not lead to pessimism. As the above quote shows, Neurath is hopeful about the scope of the activist epistemology it entails.

His position has been referred to as ‘voluntarism’ (Cartwright et al. 1996, 94). Also as ‘radical conventionalism’ (Cartwright et al. 1996, 132). Both are accurate so far as they go. It is true that Neurath’s decisionism means allowing acts of will to bridge the gap between underdetermination and the necessity of choice. But I think neither label on their own does justice to the significance that Neurath places on this position within the wider Enlightenment project of unified science. It is the active role in creating through decision-making, and the hope that Neurath sees in it, that is at the core of what I am calling his methodological constructivism. Methodological constructivism is a direct consequence of Neurath’s simultaneous embrace of naturalised epistemology and rejection of pseudo-rationalism (two sides of the same coin), as is clear from the quote above. The embrace of naturalism

²² Okruhlik (2004) focuses on Neurath, whilst (Yap 2010) discusses Carnap’s affinities with contemporary feminism.

places one in a position that parallels the existentialist. The traditional existentialist's dilemma is that in the absence of pre-determined God-given ethical norms, we must create our values for ourselves. However, this absolute freedom brings with it absolute responsibility, since we are entirely responsible for the values and norms that we create. A parallel epistemological and methodological dilemma arises for naturalists like Neurath. In rejecting the pseudo-rationalist myth of universal and pre-determined norms, standards and methods for science, the scientist is made entirely responsible for creating them. The scientist is therefore free to create science as they choose to, but bears the responsibility for what they create.

Rather than fall into nihilism Neurath sees this absence of absolutes as liberating, both requiring and vindicating the deliberate action of human beings. If we are responsible for creation, that's all the more reason to get on with it! Whilst Neurath is not naively optimistic, neither does he see reason for despair. The recognition of humanity's active role in constructing our science ought not to be scary and diminishing but empowering, widening the scope for deliberate and constructive action.²³

Neurath is hopeful about humankind's capacity to embrace these circumstances. The manifesto of the Vienna Circle ends on a distinctly zealous note: '*The scientific world-conception serves life, and life receives it*' (Carnap et al. 1929/1973, 318). This is the role of the scientific practitioner as 'touchstone' (Neurath 1936c/1983, 157).

There is however an important disanalogy between the existentialist and the methodological constructivist: the former's dilemma is necessarily an individual one. 'We are left alone, without excuse' (Sartre 1946/2007, 38). But as we have seen, for Neurath science is an inherently social process. So the scientific community as a whole bears this burden. What this calls for is not despair but organisation and cooperation. Neurath saw the institution of the *Encyclopedia of Unified Science* as a first step-toward conscious self-direction by the scientific community:

if we reject the notion of a philosophical system which is to legislate for the science, what is the maximum coordination of the sciences which remains possible? The only answer that can be given for the time being is: *An Encyclopedia of the Sciences* (Neurath 1937c/1983, 176–77)

It is methodological constructivism that motivated the development of the Unified Science movement as an international intellectual project, rather than a narrowly academic discussion group. As Nemeth puts it:

the philosophical point of the movement... was systematically to create the occasions at which scientists were enabled to perceive themselves as active agents, where they become aware that not all decisions that go into the acceptance of theories can be made on the basis of logical or empirical criteria... Rather, their self-reflection must seek to lay bare the manifold of norms by which the cognitive practice of science is connected to the social, cultural, and political world of their day (Nemeth 2007, 302)

²³ I think what I describe here is akin to what Nemeth describes as Neurath's 'will to hope' (Nemeth 1982b/1991).

At this point, however, it may seem that Neurath's position amounts to little more than a hopeful and pragmatic attitude to the voluntaristic consequences of underdetermination. Neurath may simply seem like a more exuberant Duhem. What separates Neurath, and what is definitive of his methodological constructivism, is that it is not just the content of science for which we act as a touchstone, but the methodology too.

As *scientific* people, we are prepared to check all our tenets by observation statements, but also – far removed from every absolutism – to alter the principles on which the checking is based, when this seems necessary (Neurath 1935a/1983, 115)

The absolutism Neurath refers to here is both epistemological and methodological: 'we not only deny that there could be general methods of 'induction' for the factual sciences, but also that there could be general methods of 'testing' (Neurath 1935b/1983, 123). That is to deny that there is a universal methodology of science. Neurath rejects Popper's falsificationism, not simply because it is the wrong methodology for science, but because the notion of *the* method of science is wrongheaded in and of itself (Neurath 1935b/1983, 123).²⁴ Just as they must decide on matters of scientific theory, scientists must decide on matters of scientific method. Scientists must constantly make new decisions about which norms and methods will provide the most fruitful form of scientific practice. In his later correspondence with Carnap, Neurath refers to this tolerant position, that there is more than one system, theory and method of science, as 'pluralist empiricism', and excoriates Popper for his 'anti-pluralis[m]' (Neurath, 25/09/1943, in Cat and Tuboly 2019, 594). Pluralism here means recognising that the methodology of science is in the exact same state of constant revision as the statements of science.

What this shows, and what has so often been underestimated in the secondary literature, is that Neurath's boat concerns not just epistemology, but the methodology of science too, perhaps even primarily the latter. The image of the boat serves not only as a naturalistic rejection of foundationalist epistemology, it is also a naturalistic rejection of the pseudorationalist vision of the singular and definitive methodology of science. Scientific methodology is in exactly the same position as scientific knowledge; a state of permanent flux and indefinite development, in which nothing is immune to revision, and which the practitioners of science play an active role in creating. But as with scientific theory, we cannot start from a *tabula rasa*. We begin *in media res*, and go from there, modifying the techniques as inherited. As Frank puts it, 'scientific theory is, in a way, a tool that produces other tools' (Frank 1957/2004, 356). This vision of the perpetual, deliberate reconstruction of scientific methodology is clear in Neurath's final use of his boat metaphor:

A new ship grows out of the old one, step by step - and while they are still building, the sailors may already be thinking of a new structure, and they will not always agree with one another. The whole business will go on in a way we cannot even anticipate today (Neurath 1944, 47)

²⁴ Although he thinks the former too. Neurath says Popper has 'No feeling for scientific research' (Neurath, 22/12/1942, in Tuboly and Cat 2019, 566).

But this should come as no real surprise to us. After all, the metatheory of science is itself a part of unified science.²⁵ Clearly, this falls under the scope of scientific metatheory, behaviouristics of scholars. Neurath saw it to be the role of scientists to evaluate and refine the norms guiding scientific practice. Just as Sartre concludes that we, as moral agents, are responsible for determining our own ethical rules and norms, so Neurath concludes that we as epistemic agents are responsible for determining our own methodological rules and norms. As in the case of theory-choice, decision-making is central. But to make metatheoretical decisions, metatheoretical information is needed.

2.4.2 *Behaviouristics of Scholars*

The information required for making methodological decision belongs to the research domain that Neurath calls the behaviouristics of scholars. The behaviouristics of scholars, as we saw, is the empirical component of the Bipartite metatheory. Its primary task is 'to represent the *empirical procedure* in concrete detail', through the history and sociology of science (Neurath 1937d/1987, 136). And Neurath repeatedly emphasises that a key concern of the field is the process of acceptance: 'it has to be stressed that the term "accept" belongs to behaviouristics' (Neurath 1936d/1983, 160). The process of decision-making is a (perhaps the) key topic for the behaviouristics of scholars.²⁶

The question which contradictions can just be tolerated, which not, how one behaves altogether in the development of the whole of science, is a question of behaviouristics, of history of science, of behaviouristics of scholars. But the discussion of contradictions, the discussion of the question, which groups of statements are logically of equal content, belongs to the sphere of logic (Neurath 1936d/1983, 169)

Carnap might tell us what sets of sentences are consistent or inconsistent with one another. That is a logical question. What behaviouristics can tell us is when sets of statements exhibiting certain levels of inconsistency have tended to be accepted or rejected, and how successful these decisions were relative to their purposes. This is an historical matter.

The role of the behaviouristics of scholars is to observe and document the actual practices of science, and its practitioners are therefore in the position to judge the relative successes and failures of the adoption of different methods, norms and rules. For example, it would be within the scope of the discipline to study the

²⁵ The position Neurath arrives at is one with notable similarities to Laudan's normative naturalism, according to which the normative rules of epistemology are hypothetical imperatives, contingent upon the ends of scientific enquiry and sensitive to empirical research of the frequency with which certain epistemic means lead to certain epistemic ends (Laudan 1990, 46).

²⁶ Frank similarly emphasises that 'discussing the "acceptance" of theories as an activity of the scientist' takes us out of the realm of the logic of science, and into the realm of the pragmatics (Frank 1957/2004, 348).

different outcomes of accepting two different methods of testing, or two different notions of induction. By studying their use, the behaviourist may for instance determine that one method of testing is more efficacious for one purpose than another. In that case, when a decision between the methods of testing was required, the behaviourist would essentially supply the data that allowed the decision to be made more simply. Given certain purposes, the behaviourists' data would suggest which method to adopt. Similarly, these studies may find correlations between the adoption of certain theoretical virtues and success in specific disciplines. For instance, nuclear physics may consistently see success when fruitfulness was prioritised, where geology may benefit most when unity is prioritised. Whilst such information would never necessitate the adoption of certain methods or virtues, it would surely prove useful in discussions within the scientific community. There is then a reflexivity in the relationship between the behaviouristics of scholars and the decision process, with each decision the community takes acting as a case for further study by the behaviouristics of science.

A brief example of these sort of considerations is found in one of Neurath's (until recently unpublished) papers. Historical study demonstrates that the requirements on evidence, when a hypothesis or theory is deemed well confirmed, will vary between disciplines: 'as far as possible one makes highly accurate prognoses, for example in determining an eclipse of the sun, then again in geology or history one is satisfied with much less accurate forecasts' (Neurath 2011, 25). But they also vary relative to specific purposes. As was famously argued by Rudner, how well-confirmed we require a hypothesis to be is itself dependent on the 'seriousness of a mistake' (Rudner 1953, 3). The requirements for deciding a new drug is fit for human consumption are (and should be) higher than for the hypothesis that a new trainer design will provide better grip for runners. Rudner himself raises the question of how well-confirmed the members of the Manhattan project required their hypotheses to be before detonating the first atomic bomb (Rudner 1953, 3). Presumably (hopefully), pretty high. Such decisions cannot be made on *a priori* grounds. They depend on an understanding of the methods and goals of these sub-disciplines, which must be studied empirically. In-depth research on what standards were adopted in specific cases and how successful or appropriate they were are not of mere historical interest, but should be used to inform considerations of future practice.

What this means is that the output of the behaviouristics of scholars is not merely descriptive, but also prescriptive. Keeping in mind the instrumental role of science, the behaviouristics of scholars can provide guidance on which methods, norms and standards are most appropriate as a means to a specific end. It also facilitates even broader consideration about the goals themselves being pursued; whether the means we have are sufficient for example. The methodology of science is not simply arrived at, it does not merely evolve of its own accord, but is engineered with specific aims in mind. The behaviouristics of scholars provides a body of information from which these engineering decisions can be made more clearly, more transparently, and by better informed participants. And each new decision is a new case to be studied and learned from. Just as science is an indefinite endeavour, reflection on science is too.

2.5 Neurath's Epistemic Agent

With this understanding of methodological constructivism established, we are now in a position to make clear the role of Neurath's epistemic agent in unified science. Neurath's epistemic agent is not a lonely Cartesian intellectual wanderer, divorced from time, place and social circumstance and concerned only with ideas, but a real person born into concrete historical, geographical and intellectual circumstances in possession of a pre-existing body of scientific knowledge and an established methodology. The epistemic agent is always onboard Neurath's boat. They can never start from a *tabula rasa*, but must operate within the messy reality of the current stage of intellectual development.

Because for Neurath the production of knowledge is a social endeavour, epistemic agency must also be necessarily social. 'Unified science... is not the work of individuals, but of a generation' (Neurath 1931c/1983, 58). Neurath continually reminds us of the need for the cooperation of the various branches of the sciences, and enthuses over the potential of Unified Science as a movement when they do so. The construction of unified science must itself be '*collective work*' (Neurath 1932a/1983, 99). And these beliefs were demonstrated in practice in Neurath's constant attempt to build a real-world collaborative project of unified science, through conferences, journals and (of course) an encyclopedia. Science cannot be conducted in isolation. We are not only insufficiently intelligent to work in complete isolation, but the technical and practical requirements of modern science make collaboration a practical necessity. On a more basic level, unless each scientist is to start like Descartes from a position of absolute ignorance, then the scientific community is required to provide the contemporary encyclopedia of background knowledge that is assumed in the practice of contemporary science, even given a certain degree of disagreement: 'we have to find some loyal compromise for actual collaboration, without suppressing personal convictions' (Neurath 1946a/1983, 236).²⁷

The implied epistemic agent in most traditional philosophy is inherently passive and receptive. The epistemic agent receives information, be it empirically or a priori, and applies certain criteria (although these criteria haven't yet been conclusively established). This is perhaps especially true of the empiricist tradition, and the caricature with which logical empiricism is commonly equated. This received view, Dewey called the 'spectator theory of knowledge' (Dewey 1930, 26). Dewey argues that traditional philosophy models the process of knowledge acquisition on vision. An external object stimulates knowledge in a person just as it stimulates visual images in the eye. In this interaction only the person is effected; 'the operation of inquiry excludes any element of practical activity that enters into the construction of the object known' (Dewey 1930, 25). Knowledge and action are sharply separated, the former being necessarily passive and the latter active. Knower and known

²⁷ Neurath broadly shares with Kuhn the view that normal science is not a process of perpetual radical revision, but of supplying tools for solving problems. See Kuhn (1962/2012, 35–42).

are sharply distinguished, the known preceding the process of inquiry, and the relationship between them is understood as strictly one-way. The epistemic agent is therefore seen as a passive participant in the process, absorbing knowledge from external sources whether they try to or not. But Neurath, like Dewey, rejects such a conception of knowledge, and the sharp separation of knowledge from action; 'differences between thinking and action are only of degree' (Neurath 1913/1983, 2). Neurath's agents are necessarily active. They are participants in the social production of knowledge by the scientific community, be it through the practice of day-to-day scientific experimentation, or the metatheoretical provision of better tools and methods for scientific practice. Knowledge is not merely received, but pursued, targeted and constructed. Again we see a strong vein of pragmatism in Neurath's philosophy.

But Neurath's agents are active in another, more unique way. In the traditional picture, the epistemic agent does little more than sort received information into pre-determined categories according to given criteria for justification or evidence.²⁸ For Neurath, this picture is false in two ways. First, the reality of underdetermination prevents all information being conveniently, conclusively sorted by criteria. It is pseudo-rationalism to believe that the data will always decisively favour one theory over another, or decisively disprove one hypothesis; that the data will always show one theory to be uniquely justified amongst the competing alternatives. Pragmatic decisions between competing but incompatible conclusions are an unavoidable part of science. And this role as decider is the crucial aspect of the epistemic agent's role in scientific practice. Additionally, there are no pre-determined criteria and methods for the practice of science. We must decide on these too. We are creators of the standards of scientific methodology, themselves being constantly revised on the basis of our decisions. The epistemic agent is not only active in constituting the norms, values and criteria of scientific practice, but takes an active role in making decisions that cannot be definitively solved by these criteria. It is this unique twofold activity that constitutes Neurath's epistemic agent. A perfect example of this active epistemic agent, and the role of the behaviouristics of scholars, is provided by Neurath's theory of protocol statements, the topic of the next chapter.

²⁸ While most philosophers would not argue that we know what these criteria are, most would argue that we are working towards finding them. On this assumption, once we find the right criteria, all we need to do from then on is apply them.

Chapter 3

Neurath's Conception of Protocol Statements



This chapter will provide an account of Neurath's protocol statements as fully as possible in their own terms. This means taking them in their historical context within the protocol sentence debates, and in relation to Neurath's broader philosophical project and motivations as explored in the previous chapter. The role of protocol statements in Neurath's conception of science is central, and yet for the first-time reader, they can appear completely opaque. It is clear Neurath considers them important, but less clear exactly what the import is. Not only does Neurath never provide a comprehensive account of protocol sentences as he understands them, he also never gives a detailed account of their role or position within science generally. His works therefore require more exegesis than other philosophers'. A little historical context is important for interpreting Neurath's works on protocol statements. The majority of Neurath's work on protocol statements occurs either during or in the aftermath of the Vienna Circle's protocol sentence debates. As a result, many of Neurath's works on protocol statements are polemical, beginning *in media res* and addressed to an audience assumed to be at least somewhat familiar with the argumentative context. The most blatant such examples are 'Radical Physicalism and the "Real World"' in 1934, and 'Pseudorationalism of Falsification' in 1935, in polemic with Schlick and Popper respectively. But even where less explicitly polemical, none of his articles are framed as introductory. Here, I attempt to provide what Neurath did not; a detailed account of and argument for Neurath's conception of protocol statements from the ground up.

3.1 The Protocol Sentence Debates

Protocol statements are the evidence statements of science which act as the basis of scientific testing. At the most basic level, a protocol statement is a scientific datum. They are the empirical data against which theories are tested for consistency. It was

the form this data took, precisely what its content was and their epistemological status that was contested within the Vienna Circle and became the locus of what became known as the protocol sentence debates.

What is now called the protocol sentence debates covered the period from 1929 to 1936, and involved three key participants; Schlick, Carnap and Neurath. Whilst the other members of the Vienna Circle and the adjacent philosophical community unquestionably exerted an influence, it is these three central thinkers who must take central stage in any account.¹ The debates grew initially out of discussions of Carnap's *Aufbau* (Carnap 1928a/2005). But from these limited beginnings, the debates developed into a discussion that encompassed what constitutes empiricism, and ultimately what epistemology ought to be. Uebel argues for the separation of the protocol sentence debates into four main stages, distinguishable both historically and thematically: roughly, 1929–1930, general discussion of the *Aufbau*; 1930–1932, Carnap and Neurath's discussion of the form, content and status of protocol sentences; 1934–1935, Schlick's arguments against physicalism; 1935–1936, Carnap's attempt to reconcile Schlick and Neurath (Uebel 2007d, 27). 1936 is the obvious endpoint of the debates, as it not only saw the death of Schlick, but the reorientation of Neurath and Carnap's correspondence; Carnap's adoption of Tarski's semantic theory of truth in 1935 precipitated a dispute over the permissibility of semantics and truth-talk.² From 1936 both Neurath and Carnap's conceptions of protocol statements are entrenched and no longer the direct topic of discussion, and consequently remain largely unchanged throughout their ongoing debate until Neurath's death in 1945.³

A simple way of framing Neurath's role in the protocol sentence debates is as the advocate of anti-foundationalism, especially in the clash between Schlick and Neurath in 1934. This framing, whilst not entirely incorrect, fails to tell the full story and orients the debate too narrowly around the issues of doxastic justification and epistemic foundations. The disagreement over certainty and fallibility is simply one facet of the broader opposition of two conceptions of protocol statements. A second obvious issue at play in the Neurath-Schlick phase of the debates is that of truth, with Schlick defending a correspondence theory and Neurath arguing vehemently against it.⁴

But lying behind these obvious disagreements is a deeper dispute. More than simply an epistemological debate between foundationalism and anti-foundationalism or between competing conceptions of truth-talk, the Schlick-Neurath clash must be recognised as simultaneously a meta-epistemological dispute, a clash between two conceptions of what epistemology ought to be. Where Schlick intended to provide an account of protocol statements that satisfied the demands of traditional

¹ Popper and Hempel were important peripheral contributors.

² This dispute is discussed in detail in Chap. 6.

³ For a more detailed discussion, see Uebel (2007d, 24–30).

⁴ Exactly what position Neurath was arguing for has been variously interpreted, as discussed further in 6.5.

epistemology, Neurath's account was intended to satisfy the naturalistic criteria explored in the previous chapter; to be sensitive to and useful for the sciences.

Recognising this meta-epistemological component of the debate is essential to properly grasping the conception of protocol statements that Neurath proposed. His were proposals for the required structure and content of scientific evidence statements to allow for entry into the body of scientific knowledge. What is masked by considering Neurath's position simply in connection to foundationalism is the other ways in which Neurath's naturalistic conception of protocol statements break with typical accounts of observation statements from the empiricist tradition. A good way to bring these differences out is via comparison to Neurath's contemporaries, and more traditional empiricists, Ayer and Russell.

3.2 What Is a Protocol Statement?

As sympathetic but disconnected observers, Russell and Ayer provide a useful basis for comparison when interpreting the Vienna Circle. They equated the Vienna Circle's talk of protocol sentences to what they called 'basic sentences', those 'which need not wait upon other propositions for the determination of their truth or falsehood, but are such that they can be directly confronted with the given facts' (Ayer 1936, 138). 'For every individual, there are, at every moment, "basic" propositions, i.e., propositions which he believes in virtue of some particular experience, and which, for the time being, he cannot be made to doubt.' (Russell 1937, 6). The sort of statements they have in mind are those like "I see red" or "I have a headache". For both Russell and Ayer basic sentences are first-personal reports of the given, directly verifiable and incorrigible. 'All basic propositions in the above sense are personal'; they are inaccessible to and therefore unverifiable for everyone but the observer themselves (Russell 1940/1995, 139). For the observer, these basic sentences can be known with absolute certainty, and are consequently incorrigible; 'they alone are directly and conclusively verifiable' (Ayer 1956, 54).⁵ For Russell and Ayer, the first-personal and private nature of such statements gives them an epistemic privilege, an immunity from doubt that other factual statements lack. These basic sentences fulfil a familiar role for the foundationalist empiricist; they serve as a secure and epistemically privileged basis of certain knowledge upon which science can be constructed. Ayer specifically presents such sentences as a reply to the sceptic (Ayer 1956, 51–68).

For many subsequent readers, this account has seemed to plausibly capture the position of the Vienna Circle too. That the Vienna Circle's protocol statements were the basic statements of traditional empiricist epistemology is consistent with the widespread misunderstanding that foundationalist epistemology was the Vienna Circle's *modus operandi*. But to what extent did the Vienna Circle actually maintain such a position?

⁵ They recognize that someone can lie about or misremember their experiences. The crucial point is that one cannot be mistaken about one's own basic statements in the moment of the experience.

Of our three protagonists, Schlick's position is perhaps closest to that described by Russell and Ayer. This is unsurprising since Schlick's understanding of the Vienna Circle's project has most in common with traditional epistemology. However even for Schlick, protocol statements are not epistemologically basic. The protocol statements are not the 'fundamental propositions' (Schlick 1934/1979, 377). Schlick makes a distinction between protocol statements and the Konstatierungen or 'affirmations' (Schlick 1934/1979, 381).⁶ Affirmations are present-tense reports of private, first-personal experiences, of the form 'Here now so-and-so' (Schlick 1934/1979, 385). By contrast, protocol statements, the evidence statements utilised by science, make mention of the specific time and place of an observation, who the observer was, and the act of perception (as we will see below was emphasised by Neurath). But such information is absent from the affirmations:

When I make the affirmation 'Here now blue', that is *not* the same as the protocol proposition "On such-and-such a date in April 1934, at such-and-such a time and place, Schlick perceived blue"... In protocol propositions the reference is *always* to perceptions (or they are to be supplied mentally; the personal identity of the perceiving observer is important for a scientific protocol), while in affirmations they are *never* mentioned (Schlick 1934/1979, 386)

These uniquely first-personal statements are immediate and self-contained, which grants them a privilege and privacy that all other non-analytic propositions lack.

These experiences, Schlick argues, are 'immune from all doubt' (Schlick 1934/1979, 379). They are inherently confirmed, in a way analogous to the self-confirmation of an analytic statement, but only in the moment of the experience. The affirmations and their certainty are fleeting, and memories of them are as open to doubt as any other. The ephemerality of affirmations means that 'no logically tenable structure can be erected, for they are already gone at the moment building begins' (Schlick 1934/1979, 382). Instead, the affirmations function as a point of finality, a terminus in certainty, rather than the more familiar desire for certainty as a starting point. Schlick claims that the affirmations act as the 'unshakeable points of contact between knowledge and reality' (Schlick 1934/1979, 387). In a unique way then, Schlick's affirmations fulfil the traditional epistemologist's desire for a point of certainty. They are 'the absolutely fixed points; we are glad to reach them, even if we cannot rest there' (Schlick 1934/1979, 383). Schlick cannot then be simply classified as a run-of-the-mill empiricist foundationalist, but can be considered a foundationalist of some kind.⁷ Yet whatever the idiosyncrasies of his position, Schlick's is in keeping with the empiricist tradition in that his affirmation's privileged status is a consequence of being first-personal and private: 'whatever the world-picture I construct... my own observation propositions would always be the final criterion' (Schlick 1934/1979, 380).

⁶ Here we again see a separation of science and philosophy, as discussed in Sect. 1.2.2. Those statements that serve as the basis of scientific knowledge are not sufficient to ground the possibility of scientific knowledge.

⁷ The uniqueness of Schlick's foundationalism cannot be fully explored here. See Uebel (2020a).

Carnap's views undergo the most significant changes, moving from the right to the left of the circle.⁸ Early in the debate, Carnap adopts a position close to Schlick, Ayer and Russell, taking protocols to be 'sentences about the immediately given' (Carnap 1932b/1959, 166). They, 'being an epistemological point of departure, cannot be rejected' (Carnap 1932b/1959, 191). Only once they have been translated into the physicalist language are they revisable. As phenomenalist protocols, they are incorrigible. But over time, Neurath's warnings about the impossibility of a private, phenomenalist language were heeded and, after a brief and somewhat puzzling period during which he argued that any statement can be used as a protocol (Carnap 1932d/1987), Carnap arrives at his mature position that protocol statements are physicalist statements in a 'thing language' about medium sized objects (Carnap 1936–37, 466). Here, Carnap has clearly moved away from Schlick, and (as we will see next) much closer to Neurath.

Neurath's views bear minimal resemblance to Ayer and Russell's basic sentences:

The protocol statements are statements of medium complexity and uncertainty like those familiar to us in current language. In no way do they correspond to the ideal that so many people desire, of possessing 'atomic statements' with which one could compose 'molecular statements', the statements that in fact one uses in ordinary life and in science (Neurath 1936c/1983, 152–153)

Neurath rejected any notion of protocol statements referring to the immediately given. Instead he insisted that protocol statements are third-personal and physicalist, referring to publicly accessible spatio-temporal events in the physical world. Neurath's rejection of private protocols is a consequence of his broader conception of language. As we saw previously, Neurath rejected the possibility of private language, and it is in the midst of the protocol sentence debates that Neurath provides his most clearly elaborated, but still highly terse, private language argument (Neurath 1932a/1983, 96–7). This requirement of physicalism ensures the empirical controllability of all statements, and therefore prevents the inclusion of isolated metaphysical statements (Neurath 1931b/1983, 54).

As Neurath is at pains to emphasise, this insistence on the physical and public nature of protocol statements need not limit our means of expression or the things about which we can be expressive.

I have no objection to use all shades of a painter's or a connoisseur's stories, when we transform them into a proper "physicalistic" shape. Statements of the type: "this entrance hall of a building thrills me" can be regarded as physicalist ones because they are observation statements (Neurath 1941/1983, 221)

Despite what physicalism may suggest, Neurath does not abandon the elements of experience that we typically consider mental. Protocol statements 'are connected with seeing, hearing, touching and other "sensations" (as physical occurrences)' (Neurath 1931c/1983, 65). What Neurath does reject is that such statements must be cordoned off from the rest, in the realm of the "mental" (Neurath 1941/1983, 221).

⁸ Uebel traces five stages in the development of Carnap's conception, and two in Schlick's (Uebel 2007d, 442–444).

He was desperate to prevent the false impression of 'a special "world of things" and a special "world of thinking" juxtaposed in opposition to one another' (Neurath 1937c/1983, 180). Similarly, he worries that the delineation of "inner experience" from "outer experience" functions to erect boundary lines between disciplines, and between mental and physical (Neurath 1939/1983, 209). He consequently avoids the terms "mental" and "inner experience", for fear of invoking this spectre and allowing the reader to slip into metaphysical thinking.⁹ Neurath insists on referring to experiences as physical processes because they take place in the same spatio-temporal world as everything else, not an additional and uniquely mental world.¹⁰ But this in no way impoverishes the potential subject matter of unified science. 'The description of spatio-temporal processes covers the whole of psychology' (Neurath 1931d/1973, 360). For our purposes, we can continue to refer to them as mental as is colloquially done, but with the understanding that the mental in this sense is physical in Neurath's sense.

Ayer argues that when 'protocol statements which served to describe our experiences are transformed into statements about the condition of our bodies' the connection 'between scientific knowledge and experience' is destroyed (Ayer 1963, 280). Russell similarly worries that if we reject comparison with reality 'the whole basis of empiricism, namely the appeal to experience, is gone' (Russell 1937, 10). This is an intuition many empiricists have historically adhered to. But Neurath breaks with the tradition, maintaining that public physicalist statements can still play the definitively empiricist role of connecting experience to knowledge.

Thus for us striving after knowledge of reality is reduced to striving to establish agreement between the statements of science and as many protocol statements as possible. But this is very much; in this rests empiricism (Neurath 1934/1983, 109)

For Neurath, the agreement between observation reports and the other statements of science takes the place of the agreement between word and world. Neurath adopts this position because he rejects the notion of comparison that underpins the positions of Ayer, Russell and Schlick; 'that statements can be compared with Facts' (Schlick 1935a/1979, 400).

For Neurath, comparison is only possible between like things:

we may make a comparison-statement dealing with two or more items which are characterized by qualificatory terms of the same kind. We may compare the length of a rod with the length of a tree and we may compare the character of one person with the character of another, but I should not speak in the same way of "comparing a statement with a fact" (Neurath 1944/1971, 5)

Neurath recommends that when making comparisons, we 'always indicat[e] with reference to what the comparison is made' (Neurath 1936a/1983, 135). Things can be compared quantitatively or qualitatively, or in terms of personal preference, or ordered. But none of these senses of comparing captures the supposed corre-

⁹ For Neurath's attitude to avoiding 'dangerous terms', see (Neurath 1941/1983, 217–218).

¹⁰ There is the additional factor that Neurath was fighting against the categorization of the social sciences as "mental sciences", distinct from natural sciences (Neurath 1931c/1983, 68–71).

spondence relation between world and word. Schlick, Ayer and Russell's use of "compare" describes an incoherent process. But this rejection does not mean severing the connection between our experiences and language. To connect experience to language we must formulate statements, which are then capable of comparison with the rest of our language, about our experiences. This is the role fulfilled by protocol statements. They are records of observations. As Neurath describes them, they are 'more precisely formulated observation statements' (Neurath 1934/1983, 109). The formulation of an observation statement is what allows us to talk about experience; far from severing the link, it establishes one.

As Neurath's contemporary L.J. Russell correctly notes, the conception of public physicalistic protocol statements abandons the typical ownership relation between observer and observation report, the assumption that a protocol is inherently, uniquely connected to the person who reports it. As he notes, 'to speak of a report as "so and so's report" is merely a convenient abbreviation. "So and so" is merely a part of the conditions under which the report occurred' (Russell 1934, 183). That a protocol was reported by a specific observer is an important piece of information, but also a contingent piece of information. A similar observer under the same conditions could produce the same protocol. Two protocols from different observers are no more inherent to those observers than two readings from different thermometers (Russell 1934, 183). Neurath cites L.J. Russell approvingly, framing this rejection of self-privilege as a democratisation; 'My proposal is to treat all observation statements democratically, irrespective of whether they are made by the same person at different times or by different persons' (Neurath 1941/1983, 228). The problem Neurath sets for an adequate account of observation statements is 'to keep, on the one hand, the "ipseity" of empiricist statements and on the other hand to remove the "EGO" from the traditional pedestal' (Neurath 1941/1983, 228). In other words, to retain the important fact that a specific person was the observer, an important spatio-temporal fact about the circumstances of observation, whilst rejecting the solipsistic emphasis on the self. Neurath's solution is to 'replace the term "I"—to avoid traditional pseudo-problems—by mentioning the observers name' as many times as required (Neurath 1934/1983, 102). This will become clear below, when we consider Neurath's proposed structure for protocol statements.

In abandoning this exclusive ownership relation between observer and observation statement, Neurath also sets the stage for his rejection of the methodological solipsist assumption that there is a special authority to one's own private experiences. This is crucial break with the empiricist tradition, which went against the popular consensus amongst many of his contemporaries in the Vienna Circle, as well as within contemporary epistemology. For Neurath, protocol statements have no unique security or additional warrant, but are always open to doubt and revision just as any other statement is. As proof of this revisability, Neurath imagines a scholar called Kalon who is able to write with both hands at once, and records two incompatible protocol statements simultaneously (Neurath 1932a/1983, 95). If protocol statements cannot be rejected, then we are forced to accept two mutually contradictory statements. The obvious solution then is that we must reject (at least) one of these protocols. And with that, the incorrigibility of protocol statements is shattered, the possibility of revision has to be conceded. Additionally, the adoption

of protocol statements is never imposed upon us, but those protocol statements that we adopt 'are selected on the basis of decisions' (Neurath 1934/1983, 102). Here we see another example of Neurath's decisionism opposing pseudo-rationalism; the 'metaphysical endeavour to put the unambiguity of "atomic statements" or affirmations as the eternal unambiguous reality' (Neurath 1934/1983, 107).

Neurath's physicalistic protocol statements are then almost the polar opposite of the basic sentences advocated by Ayer and Russell. Against the certain, incorrigible, first-personal, private, phenomenalist basic sentences Neurath argues for fallible, revisable, third-personal, public, spatio-temporal statements. And against many of his contemporaries, including not just Ayer and Schlick, but also Russell's Logical Atomism and *Tractatus*-era Wittgenstein's picture theory, Neurath rejects the notion of language as corresponding to or mirroring reality. And yet, whilst rejecting methodological solipsism, epistemological foundationalism, reduction to the given, and the notion of correspondence between word and world, Neurath aims to preserve the essential link between our experiences and knowledge which is central not only to empiricism but to the scientific method.

For some readers of Neurath, by rejecting the assimilation of protocol sentences to basic sentences, he is also 'led to refuse any privilege to protocol sentences' (Ruytinx 1983, 41). Consequently, 'any empirical sentence in a theory can be chosen and formulated as a protocol sentence' (Ruytinx 1983, 40). The latter remark is simply wrong since there are clear criteria for legitimacy as a protocol as we will see below, but the former captures a common epistemic intuition. The problem here is that there are two possible usages of "privileged". The one typical to most philosophers, including Ayer, Russell and Schlick, is epistemological privilege; the sort of additional warrant provided to a unique subclass of statements. Typically, this means incorrigibility, certainty or being subject to a different means of justification. This is clearly not what Neurath attributes to protocols. The second sense is a methodological privilege; that a certain subset of statement plays a special role in the practice of science and is therefore of more practical importance. In this sense protocols definitely are privileged by Neurath:

Statements of this type are neither "simple" nor "primitive" but they are of the type that is used by scientists and by the man in the street when they are discussing "factual statements," "hallucinatory statements" and other items of this type (Neurath 1941/1983, 220)

Protocol sentences therefore have a significance above a simple statement, but this privilege is not epistemological or inherent. It is a consequence of the methodological role of protocol statements in scientific testing which broadly reflects, as Neurath indicates, ordinary practice of report acceptance. Specifically, protocol statements are the terminus of scientific testing. 'Within a system of sentences, protocol sentences are the ultimate beyond which we cannot go' (Neurath 1932b/1987, 2). This methodological role is a consequence of Neurath's empiricism and replaces perpetual reference back to private experience. But whilst Neurath's protocols are not certain, they do have a stability that separates them from other statements.

Whereas in traditional accounts basic sentences are separated by their certainty, a consequence of their subjectivity, Neurath completely inverts this paradigm. Rather

than certainty or incorrigibility, ‘protocol statements have the merit of greater *stability*’ (Neurath 1935b/1983, 129). Stability here does not mean certainty but staying power of a very specific type. Neurath means by stability that protocol statements retain utility over time; that they ‘can be used today in about the same way as some centuries ago’ (Neurath 1936c/1983, 149). Neurath gives the following example:

The statement: ‘In the sixteenth century people saw fiery swords in the sky’ can be retained whereas the statement ‘There were fiery swords in the sky’ would have to be deleted. (Neurath 1935b/1983, 129)

The latter is simply rejected as a false sentence, where the former is retained (provided, as intended, that ‘saw’ is read non-factively) because it takes the form of scientific testimony. By “retained”, Neurath does not mean that the putative observation sentence has not been falsified. What Neurath means is that his protocols can be retained where a pure observation statement cannot, and thereby have a methodological utility that others do not. This stability means preserving and growing the body of evidence available to the sciences. This ‘stability explains how it is possible that people of different epochs and different ethnic groups can understand each other so well on a great variety of things concerning ordinary life’ (Neurath 1936c/1983, 150). Stability facilitates understanding, the spread and retention of information. It helps to ground the social process of science and the communal project of furthering our knowledge. Stability is therefore essential for science to make cumulative advances, and for science to function as a progressive force in the world. It is this stability, the openness to future use, that is definitive of the methodological role of a protocol statement.

A historical example is illustrative of Neurath’s point. George Gamow describes the historical data recorded by an anonymous eleventh century Chinese astronomer of the appearance of a ‘guest star’ in the sky, which after a year ceased to be visible (Quoted in Gamow 1949, 19). When modern astronomers pointed telescopes at the coordinates given by the Chinese astronomer, they saw the Crab Nebula. Measuring the rate of expansion of the nebula gave a timescale for its expansion that lined up with the Chinese astronomers report. And from his report that the guest star was ‘as visible by day as Venus’, modern scientists were able to calculate the brightness of the star at the time of its explosion (Quoted in Gamow 1949, 19). Here is a concrete example of the capacity of properly recorded historical observational data to be used by contemporary scientists. But perhaps most interesting is the lament by Gamow, himself a theoretical physicist, about historical observations:

One clear December night about 2000 years ago a bright star appeared on the eastern horizon. . . We all know and enjoy this beautiful Biblical legend, but unfortunately it lacks an important piece of information: the three wise men did not bother to measure the right ascension and declination of the new star, or to record at least its position with respect to the known constellations. . . The annals of human history record several other tantalizing accounts of this kind, connecting the appearance of a new star with the coronation of some famous king, or an attack by the enemy, or a pestilence, but failing to give us the details by which we might judge whether the phenomenon was really a “new star” or, for example, a comet. (Gamow 1949, 19)

If only the three wise men had known the correct form for protocol statements! The annals of human history are littered with observation reports that are insufficient for the purposes of contemporary science. The important question then is how we can ensure that observation reports retain their significance for posterity. What is it about a protocol statement that facilitates this stability? To find the answer, we must look in detail at Neurath's proposal for the structure and usage of protocol statements.

3.3 The Structure of Neurath's Protocol Statements

The foregoing has explained the point of Neurath's conception of protocol statements: his physicalist anti-foundationalist empiricism. But the form and structure of Neurath's protocol statements has proved equally if not more controversial. The archetypical example of a Neurathian protocol statement is:

Otto's protocol at 3.17 [Otto was word-thinking at 3.16 (in the room at 3.15 was a table perceived by Otto)] (Neurath 1941/1983, 220)¹¹

Prima facie, this seems an absurd analysis of the form that scientific evidence takes. When asked to give evidence for our beliefs, these are not the sort of replies we give. Nor is this how we expect scientists to communicate with one another. Critics of Neurath have accused him of failing to make sufficiently clear his reasons for providing such a complex conception of a protocol sentence. Even those somewhat sympathetic to Neurath have found his proposal 'cumbersome' (Reisch 2005, 7), 'awkward indeed' (Creath 1987b, 474), or 'baroque' (Richardson 1997b, 212). Amongst Neurath's contemporaries, even nominal allies associated with logical empiricism criticised his 'peculiar' proposal (Ayer 1936, 139–40) as 'insufficiently realistic' (Ayer 1982, 124). Schlick argued that Neurath renders the distinction between protocols and non-protocols 'meaningless' (Schlick 1934/1979, 374). Russell concluded that Neurath rendered protocol statements 'superfluous' (Russell 1940/1995, 143). Popper struggled to 'see what part the protocol sentences are supposed to play' (Popper 1935/2002, 79).¹² However, Neurath's failure is not one of conception, but of presentation. Although Neurath never presents an in-depth argument for his proposal, he does have very strong reasons. I do however think it is legitimate to criticise Neurath for making these reasons so opaque, thereby requiring more interpretative work of the reader than would otherwise have been necessary. So what are these reasons?

Let's return to that typical Neurathian protocol statement:

Otto's protocol at 3.17 [Otto was word-thinking at 3.16 (in the room at 3.15 was a table perceived by Otto)]

¹¹ This wording is a minor modification of his earlier formulation (Neurath 1932a/1983, 93).

¹² For more on the Popper-Neurath dispute see (Cat 1995).

The most obvious and striking feature here is the multiplicity of embedded clauses. This has caused confusion. Bertrand Russell took it at face value, interpreting Neurath's proposal as a conventional sentence reporting an experience, albeit one containing a lot of brackets. So he expands the brackets in the typical grammatical way. From this, Russell concludes, we get the true form of scientific data.

Thus according to Neurath the data of empirical science are all of the following form: "A certain person (who happens to be myself, but this, we are told, is irrelevant) is aware at a certain time that a little while ago he believed a phrase which asserted that a little while before that he had seen a table." That is to say, all empirical knowledge is based upon recollections of words used on former occasions (Russell 1940/1995, 146)

Russell is right that such an account is unacceptable. That the data of science are all of this bizarre form, recollections of beliefs about assertions of observations, is implausible. But Russell is wrong to attribute it to Neurath. An important lesson on how to read Neurath can be gained from Russell's understandable but mistaken interpretation here.

Crucially, what Russell misses, is that Neurath's protocol statements are not to be read as typical grammatical sentences. Consequently, the use of brackets by Neurath is not for typical grammatical purposes, but rather schematic, structural-analytical purposes. The brackets are being used as a means for visual separation, their typical linguistic usage is irrelevant. The protocol structure Neurath proposes is not intended as an analysis of the true form of scientific data, or an empirical account of real-world language use. Rather, what Neurath provides is a schema which explicitly visualises what is implicitly contained in useful observation reports, and our criteria for receiving them. The brackets are not being used to separate linguistic clauses, but to visually separate distinguishable factors in the appraisal of observation reports.

Recognising this can also help to assuage a second understandable worry: that Neurath requires all observers to speak like automatons. The purpose of the schema is not prescriptive or descriptive, but reconstructive. Neurath is under no illusions that this is how scientists typically talk, and nor is he suggesting that it should be. What Neurath *is* prescribing is that observations are recorded in such a way that they contain sufficient information to make reconstruction in his schematic form possible. As he puts it, his schema 'enabled us to ask the question, "When, where, and how?"' (Neurath 1946a/1983, 233). Neurath requires that observation reports be held to the same standards that any other scientific report would be: knowing who made the observation, where and when. It must therefore be possible to reconstruct a report as a Neurathean protocol statement. But we don't have to speak Neurathese. Take the following hypothetical. An ornithologist's book of sightings does not need to read like this:

Peter's protocol at 3.17 [Peter was word-thinking at 3.16 (at Alderley Edge at 3.15 was a woodpecker perceived by Peter)]

As Neurath says, such a 'complicated expression can admit of abbreviations depending on the case' (Neurath 1936c/1983, 151). The contextual information required to formulate a protocol in Neurathese needs to be recorded for the report to

have the necessary stability. But such information can be recorded more efficiently. If each entry in our ornithologist's bird diary begins with a clearly specified date and location then, along with the list of what they observed, a Neurathian protocol could be easily constructed. To function as public stable evidence, if shorthand methods are used they need to be sufficiently clear or well established to allow transparency for subsequent readers. But Neurath's requirements in no way prevent the adoption of norms, abbreviations and shorthand for the more efficient recording of observational data. This would completely contradict Neurath's fidelity to the actualities of science.

With this schematic purpose in mind, what can we learn from Neurath's proposed protocol structure? Well, Neurath is clear that his protocol statements are capable of structural decomposition:

This [protocol] sentence is so constructed that, after 'deletion of brackets', further factual sentences appear, which, however, are not protocol sentences: 'Otto's speech-thinking was at 3:16 o'clock: (at 3:15 o'clock there was a table in the room perceived by Otto)', and further: 'At 3:15 'clock there was a table in the room perceived by Otto'. (Neurath 1932a/1983, 93)

Here we must be careful in interpreting Neurath. If we take his talk of deleting brackets as referring to the expansion of brackets, then we are led to the same confusion Russell exhibits above. Rather, having recognised Neurath's proposal as a schema and not a typical sentence, we should understand "deletion of brackets" as meaning the deletion of the entire clause contained within the brackets.¹³ What Neurath tells us here is that within a protocol statement there is contained a series of sub-statements that are revealed by successive deletions of the outer-most clause.

The details of this process of deconstruction were first explicated by Uebel (2007d, 383–84).¹⁴ The initial protocol statement can be depicted as containing four clauses:

- (i) protocol (thought [stimulation state {observable fact}])

The successive deletions, and the sub-statements they reveal, are then broken down as follows:

¹³ This confusion may be born of translation. In German, "Klammer" (bracket) can be used not simply to refer to the symbols themselves, but the symbols and the clause contained within, as with "parenthesis" in English.

¹⁴ One may question Uebel's providing a four-part analysis when the above quote by Neurath would appear to suggest a three-part proposal. In terms of a textual basis, Neurath explicitly adopts a four-part analysis in 1941, as is clear from the categorisation tables he provides, used below in §3.7 (Neurath 1941/1983, 220). As Uebel shows, Neurath had also provided a slightly different four-part analysis in 1936 in correspondence with Kaufmann (Uebel 2007d, 379). There is then a clear textual basis for Uebel's analysis. More fundamentally, Neurath experimented with alternate formulations, and never decisively settled on a final or definitive formulation, which warrants the interpreter of Neurath to choose the formulation that best suits the spirit and purposes of Neurath's work. What follows will demonstrate the successes of Uebel's analysis for Neurath's purposes. For more, see Uebel (2007d, 378–381).

- (ii) thought [stimulation state {observable fact}]
- (iii) stimulation state {observable fact}
- (iv) observable fact

Importantly, the sub-sentences (ii)–(iv) are factual sentences, but not protocol sentences like (i).

Each layer of embedding corresponds to a condition for the acceptance of protocol sentences. Uebel gives a preliminary designation of these four corresponding conditions as follows (Uebel 2007d, 383–84):

- *(i) the institutional condition: somebody made the explicit claim that somebody thought that somebody was stimulated as if she perceived an object in the room;
- *(ii) the intentional condition: somebody thought that somebody was stimulated as if she perceived an object in the room;
- *(iii) the sensory condition: somebody was stimulated as if she perceived an object in the room;
- *(iv) the negative coherence condition: there is no evidence available that would contradict the object's being in the room.

The one condition that may be expected for a scientific evidence statement but is missing from Neurath's is a truth condition. For many people, it seems intuitive that only a true statement can be considered evidence, or that only truths would be used in the sciences. Crucially though, the conditions captured in the structure of a protocol are acceptance conditions. They are conditions for the acceptance of another person's reports, not for the truth of the statement, and the two are not necessarily coextensive. 'Neurath gave conditions not for when I know, but for when we may take it that some third person knows' (Uebel 2009a, 6). What this means is that for Neurath, protocol statements are instances of scientific testimony and their explicated structure exposes the conditions of intersubjective acceptance. As Neurath points out, we cannot expect to personally check every statement that is added to the encyclopedia. Science is a communal process, and we have to take some knowledge on trust. In fact, we have to take huge swathes of knowledge on trust. On board Neurath's boat, there is nothing we can do other than work within the encyclopedia as it currently exists; there is no way to start from scratch. But clearly science equally cannot simply take all and any reports as accepted. Neurath's conditions are intended to allow science to remain sufficiently discerning regarding the acceptance of scientific testimony without requiring personal verification of all observation reports (although as we will see below, the possibility of verification should be maintained if possible). This conception of protocols as scientific testimony sharply separates Neurath's theory of protocols from the foundationalist epistemological projects for which Ayer and Schlick's protocols were intended.

Neurath's protocol sentences are not simply first-personal reports of private experiences, like Schlick's *affirmations*. They are not even simple observation sentences. They are context-explicating observation statements. Neurath argues that protocol statements must provide the information "When, where, and how?" and it is for this reason that his protocol statements include the repetition of the observers name, the time and the context. To accurately categorize different types of observation report, and make decisions on their acceptability, we require this

contextualising information to be supplied along with the observation report itself. The multiply-embedded structure is designed to exhibit exactly this information to facilitate synoptic understanding of all the factors relevant to the acceptance of observation reports and, their inter-relations.¹⁵ This understanding is not achieved through lists of conditions treated in isolation from one another. Rather, synoptic understanding requires the presentation of protocol statements as constituted by concatenations, simultaneously comprehensible wholes. It is this simultaneous presentation of the parts and the whole that allows synoptic understanding. Protocol structure is intended to render these factors quite literally open to view.

Here, a possible (albeit not direct) comparison can be made to Neurath's picture language, ISOTYPE. Whilst ISOTYPE was primarily concerned with the communication of quantitative data, some of the principles are clearly applicable in this context to his protocol statements. Both are forms of visual presentation that allow for the derivation of complex information from a relatively simple image. Consider Nemeth's description of how to engage with ISOTYPE:¹⁶

Neurath's pictures should prompt those looking at them to go back and forth between at least two constellations (normally more than two) of elements, figuring out for themselves what the comparison is all about (Nemeth 2019, 130)

In the case of Neurath's protocol structure, the different constellations to go between are the multiple embedded clauses, and the sub-statements that they make up. What the comparison with ISOTYPE shows is that instantaneous, *prima facie* understanding was not the expectation. The information to be conveyed is complex, and as such requires a degree of complexity when visually presented. There are multiple inter-connected elements that are relevant to the appraisal of observation reports, and this is reflected in the complexity of Neurath's protocol structure. It is only through active engagement with and consideration of the protocol structure by the epistemic agent that the information conveyed there-in is effectively communicated.

What is essential to understand is that Neurath provides a schema, and the structure is designed with a very particular schematic purpose in mind. His proposal prioritises the presentation of the factors relevant to the acceptance procedure in a way that maximises both ease of understanding and practical utility. Maximum clarity about when these conditions are met and the ease with which these statements can be made use of in practice are of much greater significance for his purposes than providing his proposals in proper logical form. Consequently, rather than a protocol sentence as a statement of standard predicate calculus with typical logical form, Neurath gives a synoptic schema that visualises the factors

¹⁵ Neurath describes the purpose of his Encyclopedia as allowing science to be 'presented to us synoptically in its totality' (Neurath 1937d/1987, 141). I think the same motivation underlies the proposed structure for protocol statements.

¹⁶ See also Angélique Groß's discussion of the '*presentation concept*' of ISOTYPE where she highlights the 'methodical structurization... provid[ing] concentration on the essential and... the human capacity for creating association' (Groß 2019, 191).

and process of protocol acceptance with maximum perspicuity. The cost of this maximisation of epistemological perspicuity is a loss of logical simplicity. But Neurath's emphasis on the pragmatics of science over the logic of science, make this trade-off understandable.

To see why this trade-off is necessary, we can provide a rudimentary attempt at a proper logical form for Neurath's protocols. A protocol statement is acceptable if all four of conditions $*(i)–*(iv)$ are met, that being the case in which all four sub-statements are themselves accepted. The correct logical form for a valid protocol statement is then formalizable as: $(i) \wedge (ii) \wedge (iii) \wedge (iv)$. Initially, this logical form does not seem too complicated. But, when a concrete example is used, the unsuitability of this formulation for Neurath's purposes becomes become apparent. Take the example of a protocol given by Neurath:

Karl's protocol: Karl formulates: Karl is seeing: In the room there is a round table (Neurath in a letter to Felix Kaufmann, quoted in and Translated by Uebel 2007c, 386)

For this example, $(i) \wedge (ii) \wedge (iii) \wedge (iv)$ becomes:

"Karl protocolises that Karl formulates that Karl sees that in the room is a round table" $\wedge$ "Karl formulates that Karl sees that in the room is a round table" $\wedge$ "Karl sees that in the room is a round table" $\wedge$ "In the room is a round table"

This is a congested and unwieldy sentence. Even this rudimentary approximation of a proper logical form is so complex that our ability to grasp it as a whole composed of inter-connected parts is lost. The desired synoptic understanding is not facilitated by this logical form. By contrast, Neurath's structure exhibits the whole and the parts simultaneously, thereby highlighting the inter-connections, and allowing for engagement by the observer.

The deletion of brackets also plays a practical role, to facilitate quick and easy checking of conditions $*(i)–*(iv)$. To see how, we can use an example. Again, take the protocol sentence:

Karl's protocol: (Karl formulates: [Karl sees: {In the room is a round table.}])

With it, we can decide whether condition $*(i)$ is met. Having done so, we can then "delete the bracket", i.e. delete the first clause. We then have the sub-statement (ii):

Karl formulates: [Karl sees: {In the room is a round table.}]

This allows us to check condition $*(ii)$. Deleting brackets again gives sub-statement (iii):

Karl sees: {In the room is a round table.}

We can now check $*(iii)$. And finally, deletion gives sub-statement (iv), the object sentence:¹⁷

In the room is a round table.

¹⁷ I am using "object sentence" in the sense introduced by Uebel: the sub-statement contained in the inner-most bracket expressing an observable fact (Uebel 2009a, 6). In an earlier work, he uses the term 'content statement' (Uebel 2007d, 388).

This allows us to check *(iv). These three successive deletions allow all four conditions to be quickly and easily checked, allowing for simple categorisation of protocol statements (via charts like those Neurath supplies, which we will see in Sect. 3.7).

In real-world situations, deletion could be achieved by covering the “deleted” clauses with your hand, striking them through with a pen or any other physical means for successively removing each clause. During the process of successive deletions, one could simply leave a tick or cross for each successive check. The result could then be immediately compared with the categorisation charts (discussed below). This gives a sense of how Neurath intended these protocol statements to be made use of practically. Most importantly for our purposes though, far from being baroque or cumbersome as his critics have argued, Neurath's protocol statements, when understood in their proper context, are eminently practical by design. Neurath's structure, far from being a mistake, is specifically designed for and successful in terms of his purposes.¹⁸

3.4 The Four Conditions for Testimony Acceptance

With this context in place, we can now address the specifics of Neurath's proposal. Taking protocols as instances of scientific testimony, what criteria does Neurath impose on them? Well, we know that he specified four conditions that must be met for a protocol to be considered valid. The first is the institutional condition: was the report publicly and explicitly avowed? If so, then we have a genuine protocol, which is to say, an instance of scientific testimony. This first condition makes performativity a necessary component of any instance of scientific testimony. Contrary to Ayer, who considered Neurath's usage of the ‘title of *Protokollsätze*... arbitrary and misleading’, Neurath's choice of terminology is actually both appropriate and illuminating (Ayer 1936, 140). The German term “Protokoll” is not a specifically philosophical one but rather refers to a report or record. A “Versuchsprotokoll” is a record of the set-up, execution and result of an experiment. For Neurath, a protocol is something being deliberately put into the permanent public record, ‘like the minutes of a faculty meeting or a statement made to the police’ (Cartwright and Cat 1996, 81). The choice of term indicates that for Neurath, these were statements that were being specifically entered into the public arena with the purposes of being preserved for posterity and remaining accessible for future use. The form and content of protocol statements must facilitate this.

¹⁸ An interesting suggestion on this issue was made to me by Elisabeth Nemeth. She wondered whether it is only philosophers who find the structural and visual component of Neurath's schema so unintuitive and confusing. Would a graphic designer or perhaps an architect recognize Neurath's intentions as more transparent and understandable? Is it simply that philosophers do not typically dabble in these kinds of visual experiments outside the narrow confines of agreed upon formalizations for logic?

This requirement is consistent with Neurath's broader views on science and language. Both are necessarily social and public, so a private protocol would have no utility. But the requirement is also rooted in Neurath's sensitivity to the actual nitty-gritty of scientific practice, especially in the social sciences with which he was more personally familiar. Whilst this requirement of performativity may seem superfluous or pedantic, it is actually grounded in considerations of what preserves the utility of historical observation reports. The explicit act of protocolisation is confirmation to future scientists of definite intent to report (although not of course sincerity). This seems banal, but cases without such explicitness make clear the advantages. Imagine stumbling upon a notebook, that contains the following: "On Thursday 3rd March 2020, I saw an enormous green bird in Whitworth Park". Whilst the credibility is not established, that this is intended as a report of an observation is clear. Compare that to finding the following note: "Big green bird!" Is this an observation report? A dream diary? An idea for a band name or short story? It is unclear whether the initial writer even intended to make a report, so deciding how or whether to utilise it becomes a complicated and drawn-out task. The former, explicit report requires no such considerations, we can simply move on to the next condition. Neurath's requirement for explicit protocolisation is an attempt to ensure that protocol statements remain as useful as possible for as long as possible in as many contexts as possible, a desiderata few people (let alone scientists) could oppose.

Considerations of the above sort however raise a slight issue with the institutional condition: does it require that a claim is made publicly? Is the process of protocolisation a necessarily social one, or can a protocol be recorded in a private diary (albeit of course, not in a private language)? Neurath refers to travel journals, specifically giving the example of Darwin's journals from his voyage on the Beagle, as 'collections of data' (Neurath 1935b/1983, 126). But is this because Darwin's journals were intended for publication, because they were actually published, or would they still provide valuable data if they had been discovered in his private collection posthumously? This is a nuance that Neurath is never sufficiently clear on. However, some brief thoughts on his behalf can make a first step toward an answer. Imagine for example, that Darwin had died unexpectedly, and that consequently some part of Darwin's journal had been lost amongst his papers, and was thus only published posthumously. Or that some of his journals were lost in the post, and only found much later. Or imagine a personal diary explicitly intended for publication, but only after certain members of the author's immediate family had died so as to avoid scandal or embarrassment. Such a diary would surely remain a valuable source of data, and the intent to preserve for posterity is clearly present. What such thought experiments suggest is that the timescale for when the reports are made public seems irrelevant. Perhaps then, the institutional condition ought to be phrased: *somebody intended to make the explicit claim that somebody thought that somebody was stimulated as if she perceived an object in the room*. Ultimately, Neurath didn't supply sufficient detail. But the above thoughts suggest (at a minimum) that the intent to, not simply the act of, making public is sufficient.

Details aside, passing the institutional condition is all that is required to qualify as a legitimate act of testimony. For Neurath, the nature of scientific testimony is relatively simple: scientific testimony is the delivery of an observation report, structured to provide maximum stability and utility, and entered into the public record for future use. But what makes successful testimony? After all, not every act of testimony is successful. How are those cases of genuine transmission distinguished? For Neurath, these are the cases in which all three of conditions *(ii)–*(iv) are met. Only then is the fullness of the original report absorbed so as to fulfil its role in scientific practice.

Going forward, I will adopt Uebel's terminology for differentiating between types of protocols (Uebel 2009a, 7). An "accredited" protocol has met the institutional condition and consequently counts as a proper instance of scientific testimony. A "valid" protocol meets all four conditions, and consequently counts as accepted scientific testimony. It is only valid protocols that provide transmission of content (which will be explored below). All valid protocols are accredited, but not all accredited protocols are valid. As is clear from Neurath's criteria, the process of recognising a protocol as either valid or non-valid is not immediate. Deciding whether the conditions have been met requires extensive clarification (and decision-making).

The process of receiving testimony for Neurath therefore has two stages. The first stage (checking the institutional condition) resolves the issue of whether we accredit the speaker's act: did they or did they not record their observation report for posterity? The second stage resolves the question of a protocol's validity. It is conditions *(ii)–*(iv), the intentional, sensory, and negative-coherence conditions, that resolve this and consequently determine how a protocol is to be processed and used as evidence. The first stage of testimony reception is a far less complex and less extensive process than the second. Questions of competency and sincerity are irrelevant to the process of accreditation. Whether or not sufficient information was provided to allow formulation in the form of Neurath's schema is a simple matter of fact, as is whether it was at least made (potentially) publicly available. If so, then it is accredited. Whether or not the speaker was a liar, a fool or hallucinating is irrelevant to the question of accreditation. A report of a lie, of an hallucination, or simply a mistake are still for Neurath examples of accredited scientific testimony, just not valid testimony. The second stage, establishing the validity (or lack thereof) of a protocol is a far more extensive and more complicated procedure, with each condition playing a different role in establishing the reliability of a report. It is in this second stage of protocol reception that the issues occupying much contemporary epistemology of testimony become relevant. Assessment of both aspects of reliability, speaker competence and sincerity, are covered by Neurath's conditions *(ii)–*(iv).

Each of conditions *(ii)–*(iv) has a separate and specific role to play in establishing the validity of a protocol. Unsurprisingly, the intentional condition *(ii) covers the issue of speaker sincerity. The purpose of the intentional condition is to guarantee the authenticity of the report captured in a protocol. Lying involves a failure to fulfil it. Lies therefore, although accredited protocols, are non-valid. But

authenticity for Neurath involves more than simply sincerity. As Uebel has shown, '[the intentional condition] asserts that the observer conceptualized the event or state of affairs at issue in the terms cognate with those employed in the protocol' (Uebel 2007d, 387). The notion of authenticity in this condition is therefore twofold. On the one hand, it refers to the intention of the speaker, that they are not lying or aiming to deceive. But it also means that the protocol is formulated in the voice of the original speaker. This requires that the protocol is not rephrased or reinterpreted, that it retains the marks of the world-conception and encyclopedia in which it was initially formulated.¹⁹ To use Neurath's example, were we to rephrase the reports of medieval people and express what the report told us as "medieval peasants reported seeing meteor showers", we would violate the authenticity of the original protocol. It is crucial that hearers do not impose their own conceptual frameworks onto past reports, lest those reports be altered and corrupted by it.

We can easily imagine such a case of corruption. Imagine that instead of receiving the reports of the eleventh century Chinese astronomers, we had only a second-hand report via European renaissance astronomers. These renaissance astronomers dismissed the idea of a guest star, but concluded the report was evidence of some rare phenomena in the celestial spheres. Now, imagine that only their reinterpretation survived, which says "In the eleventh century, Chinese astronomers saw an anomaly in the spheres". For contemporary astronomers, what use is that report? We can tell that a reinterpretation has happened: eleventh century Chinese astronomers wouldn't cognise in terms of celestial spheres. But then what use can be made of the report when it has been corrupted, when the concepts expressing it have been altered? At the very least, the report is significantly less reliable. The process of reinterpretation undermines the value of the original observation report. In the absence of the original report, in authentic form, we have only a second-hand reinterpretation, which because it lacks authenticity, lacks (or at least has diminished) utility. Here it becomes clear how important this requirement on authenticity is, beyond the requirements of simple intellectual integrity, for providing protocol statements with their requisite stability.

This does not mean however that we cannot make use of such reports by attempting to separate them from their original conceptual framework, but in doing so we must also be careful to recognise and preserve their original formulation. That they can be separated in this way is essential to them having any utility at all. Neurath gives an example about a Roman boat as a case where a report still functions despite conceptual change:

"In a certain year B.C. a ship moved up the waters of the Tiber in the direction of Rome." The terms of this statement can be used today in about the same way as some centuries ago, although what corresponds in science to the common term "water" has today a definition that is different from that of centuries ago (Neurath 1936c/1983, 149)

¹⁹ A Kuhnian could say that a protocol must retain the marks of the paradigm in which it was initially recorded.

A more detailed historical example is found in Hacking's account of the experimental work of the astronomer William Herschel, who painstakingly recorded the different amounts of heat transmitted through his telescopes when different coloured filters were used, and further records of observations of refraction and reflections at different angles, intensities, and so on. Hacking says we 'shall not find a better sense-datum report than this, in the whole of physical science' (Hacking 1983, 176). And part of what makes them so significant is that, despite Herschel himself having a Newtonian theory of light as rays of particles, his body of observation reports retain their significance to us now, with a post-Newtonian understanding of wave-particle duality, even though Herschel himself ultimately gave up on his hypothesis for explaining what he had observed. We therefore see a concrete example of how a scientist's protocols can transcend both the theory into which they were born, and the work to which their original observer put them.

It is then left to *(iii) and *(iv), the sensory and negative-coherence conditions, to establish the competency of the speaker. Despite playing a similar role however, it is essential that the two conditions are separate. The sensory condition requires that the speaker was stimulated 'in certain areas of perception in the brain' as described in the report (Neurath 1931b/1983, 55). Condition *(iii) refers to the physical process of sensory stimulation, where *(ii) refers to the cognitive awareness and conceptualisation of that stimulation by the observer. Put more simply, *(iii) requires that the speaker had the sensory sensation they reported. This is a self-evident requirement of an accurate observation report; the speaker needs to stand in an appropriate causal nexus. Crucially, *(iii) does not ensure that the experience was veridical. By isolating *(iii) as a separate condition in this way, relating only to the speaker's states, 'changes... only within the human body', and not 'spatio-temporal changes that have taken place outside the man', Neurath's theory is able to distinguish between cases where experiences correlate with the external world and cases where they do not (Neurath 1931b/1983, 55). If factual accuracy and sensory stimulation were captured by the same condition, then the criteria would be unable to differentiate between certain kinds of lying, mistaken claims, hallucinations and dreams. By maintaining the separation of the sensory condition, Neurath ensures his criteria provide a more fine-grained basis for categorising different types of report, and the reasons for their acceptance or rejection as valid. This is perhaps the most important reason for the adoption of the four-part over the three-part analysis, and therefore one of the most important justifications of Uebel's approach.

The negative-coherence condition is the most wide-ranging of Neurath's conditions. In the terminology of contemporary epistemology, the negative-coherence condition requires that a piece of scientific testimony have no defeaters. Whilst we do not require decisive reasons in favour of accepting a piece of testimony, Neurath does require that we do not have good reasons for rejecting it. These can be simple. 'When a man wrote, "I have seen a zebra in the zoo," then somebody may say: "There is a zebra in the zoo, but you are a liar as you did not visit the zoo."' (Neurath 1941/1983, 220). In practice, there will very frequently (if not always) be some reasons for rejecting a piece of testimony. No encyclopedia is ever perfectly consistent, there will always be certain incompatibilities and

contradictions. As Neurath notes, ‘it frequently happens that a certain theory, useful in a determined field, contradicts another theory useful in a different field’ (Neurath 1941/1983, 215).²⁰ If we reject any piece of testimony for which there is even the slightest reason to doubt, we would never accept testimony at all. The question is what level of counterevidence is required for rejection, and this will typically be a matter of decision. For instance, a protocol that disagrees with just one or two statements of our encyclopedia would typically pass the negative-coherence condition. After all, no encyclopedia is ever entirely self-consistent. But a protocol that contradicts vast swathes of our data or well-established theories is likely to be rejected. In some cases, we may find it necessary to conduct further investigations to judge the reliability of the protocol in question if our pre-existing body of protocols is insufficient to determine reliability. Protocols are not therefore restricted to a dichotomy of rejected and validated; ‘At each moment there are statements about which a decision has not yet been made’ (Neurath 1936a/1983, 135). Neurath recognises that sometimes we are not yet in a position to make a qualified judgement, and that in that position the rational thing to do is withhold judgement.

Since our ability to judge the validity of a protocol statement is dependent on our current body of knowledge, the status of a protocol is perpetually open to revision:

Sentences recorded earlier as “reality statements” can be turned, e.g., into “hallucination statements”. We say Otto perceived the man, but nobody else could perceive him. We may well hypothesize that we can only use a sentence about changes in Otto’s nervous system, not one about a perceived man external to him. (Neurath 1932b/1987, 5–6)

The appraisal and categorisation of a protocol statement relative to the four conditions is never final or decisive, but always contingent on our current encyclopedia and stock of protocol statements. Neurathian protocols allow for both reuse and revision, despite their content remaining the same, because of the use we choose to make of them. The possibility of revising protocols, of reconsidering whether they meet the four conditions, also reaffirms the importance of the intentional condition, and the authenticity it guarantees.

Importantly, conditions *(ii), *(iii) and *(iv) are all checked by comparison with the rest of our encyclopedia, and most importantly, other protocol statements. This is obvious for condition *(iv). But it applies to *(ii) and *(iii) too. After all, how are we to demonstrate that these conditions are met, that a speaker was speaking sincerely, or that the speaker did have the experience they reported, than by resorting to further evidence? This may be evidence about the speaker generally, about whether they are typically trustworthy or not. It may be evidence from other protocols about the same time and place, from different speakers. Neurath requires that the reliability of a speaker be empirically determined.

²⁰ The obvious example is contemporary physics, where quantum mechanics and Einsteinian relativity are incompatible and yet equally indispensable and successful in their respective areas.

3.5 The Pragmatic Condition

Beyond the four formal conditions for the acceptance of protocols spelled out above, there is one final pragmatic condition. This is simply the decision whether or not to accept a protocol as “binding”. Consider that there will always be some internal contradictions within an encyclopedia. Neurath recognised the existence of Kuhnian anomalies which despite being inconsistencies, were tolerated as non-fatal:

We can very well imagine that a falsifying hypothesis that Popper would call “confirmed” is pushed aside by a successful scientist because, on the basis of very serious general considerations, he deems it an impediment to the development of science that itself would show how this objection is to be refuted (Neurath 1935b/1983, 124)

Clearly some inconsistencies are considered fatal. But even those cases are not as neat as Popper portrays them. A famous example highlighted by Lakatos is the case of Mercury’s path around the sun, which contradicts the predictions of Newtonian mechanics. Yet, contrary to naïve Popperian falsification, ‘eighty-five years elapsed between the acceptance of the perihelion of Mercury as an anomaly and its acceptance as a falsification of Newton’s theory’ (Lakatos 1970, 30). It was only the arrival of the superior explanation provided by Einstein’s theory that ‘transformed a dull anomaly into a brilliant ‘refutation’ of Newton’s research programme’ (Lakatos 1970, 72). Such examples demonstrate that which anomalies are treated as refutations and which are treated as benign is a matter of decision, we choose which protocols we attribute significance to.

To capture this difference Uebel introduces the notion of binding, which he defines as follows:

A valid protocol is considered binding iff (a) a test statement (derived from a theory) is implied by the content sentence of the protocol and the protocol is thus conceived as a confirming instance of that theory, or (b) the test statement is incompatible with the content sentence of the protocol and the protocol is conceived as a disconfirming instance of that theory... A protocol is non-binding, if it is not conceived as a confirming or disconfirming instance of that theory, irrespective of whether the test statement (derived from that theory) is implied by or incompatible with the content sentence of the protocol (Uebel 2007d, 388–89)

As Uebel points out, this distinction between binding and non-binding is implicit in the Neurath principle that when faced with recalcitrant protocol we can alter our encyclopedia or the protocol. Without licencing academic malpractice and permitting scientists to re-write the content of previous observation reports, revision of a protocol can only amount to a rejection of its significance (Uebel 2007d, 389).²¹ For those who want science to provide decisive answers, for the scientific method to unequivocally prescribe what scientists should do in all circumstances, the idea that we decide which protocols are binding can seem deeply troubling. But this

²¹ It is also possible that a recalcitrant protocol is rejected as non-valid. What is important about Uebel’s addition here is that, even given valid protocols, it is still only on the basis of decision that they play a role in confirmation.

distinction between binding and non-binding protocols allows Neurath to once again capture the actual practice of the sciences. An interesting case of this sort is noted by Neurath himself in his correspondence with Carnap, where he makes reference to ‘the Ehrenhaft discussion’ (Neurath, 1/04/1944, in Cat and Tuboly 2019, 618).

The Ehrenhaft-Millikan dispute was a long-running (1910–1925) debate over the size of the elementary electrical charge.²² Millikan argued that the electron was the smallest unit, whilst Ehrenhaft argued for the existence of subelectrons of smaller charge. Their experiments involved spraying droplets of oil or water between two oppositely charged metal plates. Occasionally, a droplet would collect an extra electron (or several) and, now negatively charged, would be attracted to the positively charged plate. By measuring the droplet’s speed as it moved toward the positive plate, the charge could be calculated. By repetition, a common factor could be found, to isolate the charge of a single electron.²³ Both scientists repeatedly conducted similar experiments and produced similar data. Despite this, they still disagreed. The dispute was not over the data itself, but how to interpret it. The ‘Millikan–Ehrenhaft controversy... demonstrate[es] how two well-trained scientists can interpret the same set of data in two different ways’ (Niaz 2000, 480). More specifically, how they can disagree over which data ought to be held binding. Ehrenhaft maintained that all the data must be accounted for, including cases where the charge indicated was less than that of an electron. He seems to have been committed to the ideal of ‘the traditional scientific method’ (Niaz 2005, 699). Millikan on the other hand openly admitted to discarding certain readings that did not agree with his proposal (Holton 1978, 193–94). ‘Millikan’s data selection procedure depended primarily on his commitment to his presuppositions’, specifically, his belief in the electron as the smallest unit of charge (Niaz 2005, 681).

Ehrenhaft lost the dispute. Millikan received the Nobel Prize for Physics in 1923 and contemporary physics recognises electrons as the smallest discrete units of charge. We would expect Neurath to celebrate Millikan, as an example of non-falsificationist methods being vindicated in practice. But Neurath also defends Ehrenhaft:

I convinced some people that Ehrenhafts remarks are more sound than they thought and nevertheless said that their opposition remains sound from point of view [sic] of the positive instances (Neurath, 1/04/1944, in Cat and Tuboly 2019, 618)

Neurath’s point is not that Ehrenhaft was right and Millikan wrong. His point is that Ehrenhaft was not irrational or foolish, and that being on the losing side of the debate

²² That this is the “Ehrenhaft discussion” Neurath refers to is the most plausible interpretation, but not absolutely certain. Neurath’s discussion is unfortunately oblique. Cat and Tuboly previously identified Felix Ehrenhaft as the most likely referent (Cat and Tuboly 2019, 618 fn. 269). And the central role of Felix Ehrenhaft personally in this debate makes it the obvious candidate for the name “the Ehrenhaft discussion”. Perhaps most telling however is Neurath’s reference to Ehrenhaft’s arguments ‘against the electricity elementary unit [sic]’ (Neurath, 1/04/1944, in Cat and Tuboly 2019, 618). The debate also illustrates Neurath’s points exceptionally well, as we will see.

²³ For more details, see Holton (1978).

did not make him a bad scientist. But crucially, the same is true of Millikan. It is possible, even common, that both sides of a scientific dispute have good reasons and evidence in their favour. In making these decisions, scientists exercise judgement. But it is possible for such judgements to disagree, and this does not render one party irrational or unscientific. We must therefore move beyond the easy caricature of scientific debate: the bold visionary struggling against the recalcitrant traditionalist who refuses to change despite the data.²⁴ Such narratives are common in the popular imagination: Galileo and Darwin against the religious establishments of their day are the most obvious such examples. But these black and white stories, whilst making for more compelling and digestible narratives, are oversimplifications. Scientific rationality is rarely if ever isolated on one side of any such dispute. It is only by recognising the role of pragmatic decision-making, foregoing the desire for a scientific decision-calculus that unilaterally prescribes correct and incorrect (or rational and irrational) choices in all circumstances, and recognising the importance of binding that we can actually understand such historical cases.

Neurath's ideas have since been vindicated. Contemporary literature on the debate emphasises the extent to which both scientists were justified. And the since opened archives of the Nobel committee show that as late as 1920, Millikan was refused the prize because the dispute was still considered unresolved amongst contemporary physicists (Niaz 2005, 698). Neurath's conclusions also pre-empt the conclusions of more recent scholarship amongst historians and philosophers of science. Holton for instance explains Millikan's behaviour via 'the scientist's ability during the early period of theory construction and theory confirmation to hold in abeyance final judgments concerning the validity of apparent falsifications of a promising hypothesis' (Holton 1978, 212). In other words, the importance of the scientist's ability to treat dis-confirmatory data as non-binding. Similarly, Niaz notes how 'Millikan's methodology consisted precisely in not abandoning his guiding assumptions despite anomalous data' (Niaz 2000, 492). Holton even specifically notes how cases like this prove problematic for Popperian falsificationism (Holton 1978, 213).

There is then a further step, beyond meeting the four formal conditions for validity, that must be taken for a protocol to be used in testing. We have to make a pragmatic decision: 'should we regard something as a serious shaking or should we simply disregard it for the time being' (Neurath 1935b/1983, 127)? This decision is made by deciding which of the existing protocols within the encyclopedia are taken as binding. As with everything else, which protocols are considered binding is revisable. And the basis on which we decide which protocols to accredit are exactly those discussed in the previous chapter: the norms and values we are involved in creating in our roles as active epistemic agents.

²⁴ This opposition to presenting the history of science via dichotomies emerges in Neurath's early work on the history of optics (Neurath 1915/1973, 1916/1983).

3.6 Utilising Testimony: Extraction

We have so far explained how Neurath's conditions specify the criteria for the accreditation and validation of protocols, but we have not yet explained how these protocols are processed to become useful data for the sciences. How does a piece of testimony become usable as evidence? If we establish that a protocol is valid, processing into evidence is relatively simple: extract the object sentence embedded in the protocol, and add it to our encyclopedia as a piece of evidence. By establishing that a protocol is a reality statement (Neurath's term for a valid protocol), we determine it to be an instance of reliable testimony. This licences the extraction of the object sentence at the protocol's core as a piece of testimonial knowledge, a piece of scientific data. But there is a subtlety to the process that can be overlooked. For Neurath, so long as a protocol is accredited it can be added into our encyclopedia. It is not incorporation into the encyclopedia that distinguishes valid and invalid protocols, but whether the object sentence is also added to the encyclopedia:

If we incorporate part of the above mentioned protocol, the statement "in the room was a table perceived by Charles" along with the whole protocol into the body of science, then we can speak of a 'reality formulation', whereas we would speak of a 'dream or hallucination formulation', if we accept the whole protocol but not the part "in the room was a table perceived by Charles". (Neurath 1934/1983, 107)

To make sense of this, we need then to differentiate two different areas (metaphorically speaking) of an encyclopedia. To capture this, I propose a distinction between two different but related bodies of information: the protocol bank, and the data bank.²⁵ The protocol-bank is essentially a comprehensive list of all available accredited protocol statements. It acts as an enormous archive, a compendium of scientific testimony, from which scientists can draw as they require. Inclusion of a protocol in the protocol-bank requires only accreditation, not validity. The data bank is the body of statements that makes up the data of our best current scientific theories. Unlike the protocol-bank however, the contents of the data-bank are not protocol statements. The data added to the data-bank are the object sentences of all the valid protocol statements.

Lets once again return to Neurath's example from above:

Karl's protocol: Karl formulates: Karl is seeing: In the room there is a round table

As an accredited protocol, the whole statement can be added to the protocol bank. For the sake of argument, let us also accept that this protocol is valid. In this case, as well as adding the whole statement to the protocol-bank, we are licenced to extract the object sentence "In the room there is a round table" and add it to the data-bank. This physicalistic statement about the presence and shape of a table is the data that we receive through the act of testimony. This makes clear why Neurath denies that our own protocols have any privileged epistemic status. We

²⁵ The proposed terminology is my own.

can derive information from another person's experience in exactly the same way as we would derive information from our own protocol. The process of protocolisation necessarily puts the protocol statements of one's own on the same epistemic footing as those of everyone else. The datum itself is not a complexly-structured report, but a simple physicalist statement about medium-sized objects, expressing a spatio-temporal state of affairs. Contrary to Russell's misinterpretation then, Neurath's account does not render the data of empirical science implausibly complex. It is the protocol schema, and the process of acceptance it embodies, that exhibits such complexity. The complexity of the delivery mechanism does not undermine the simplicity of the datum it supplies. The complexly structured protocol is a means for delivering this simple physicalist datum.

The data bank is more exclusive and smaller than the protocol bank. Any scientific datum in the data-bank must have at least one corresponding protocol in the protocol-bank. However, not all protocols in the protocol-bank will have corresponding data in the data-bank: only the valid ones, the reality statements. Whilst any protocol can be added to our encyclopedia as a report to be used as a potential piece of evidence, only valid protocols are absorbed into our encyclopedia alongside the object sentence embedded at its core. A valid protocol gives you both the report *qua* report, and the data delivered by the report. It is this additional incorporation of the communicated content that distinguishes the role of a valid protocol statement. All reports provide potential evidence, even lies and mistakes. But valid testimony brings with it additional evidence of a factual sort, scientific data, by virtue of its validity providing warrant to add it to our encyclopedia.

This interpretation of Neurath's protocols may initially seem unintuitive but, it is demonstrated explicitly in the private correspondence between Neurath and Felix Kauffman from which the above quote is taken:

Karl's protocol: Karl formulates: Karl sees: In the room is a round table.
and

Karl's protocol: Karl formulates: Karl touches: In the room is a round table.
are both statements featuring the part:

In the room [is] a round table.

(Quoted in and Translated by Uebel 2007c, 386)

In other words, both supply the same piece of scientific data (about the table) but the reports supplying the data are different, one being prompted by a visual stimulus and the other by a tactile stimulus.²⁶ Clearly the object sentence itself can be meaningful separated from the protocol that initially delivers it, and be understood and used in isolation from it.

What we must also remember is that there is a dual-sense in which a valid protocol provides evidence, both as an observation statement and as conduit for the

²⁶ As Uebel notes, this 'allows for the convergence of reports in different sense modalities to be clearly displayed' (Uebel 2007c, 387). When I claim that "There is bread on the table in front of me", that I smelled, saw, touched, and tasted it all supply distinct pieces of evidence for the same object sentence. This also has the advantage of making multiple reports delivering the same data not redundant, but rather mutually reinforcing.

object sentence. The use of valid protocols as reports is exactly the use they would be put to if they were non-valid. It is to the use of non-valid protocols that we now turn.

3.7 Utilising Testimony: Abduction

So far, we have seen the way that valid protocol statements are utilised. But what about the non-valid protocols? They cannot be utilised in the same way, since the extraction of their data sentence is not warranted. But it would be nice if they have some use. After all, the protocol-bank is full of non-valid protocols being preserved for posterity. In fact, Neurath's own example of a stable protocol statement, the swords in the sky, is a non-valid one. Clearly then, Neurath believed that non-valid protocols still have a utility for the sciences. The task now is to explain this utility.

It seems intuitive that when processing testimony, we sort it into the true and the false. We can then preserve the true and discard the false. But what use is a false report? Well in science, Neurath tells us, potentially a great deal! Even a protocol determined to be a lie remains a potentially useful piece of evidence, relative to specific purposes. For instance, such protocols may be used as evidence that the speaker is not trustworthy in future appraisals of their testimony. These lie-protocols could function as evidence that the speaker's future protocols fail condition *(ii), the intentional condition. Or such protocols may be useful for future biographers or historians as evidence of the speaker's deceitful, conniving character. Such reports are also useable in multiple non-exclusive ways. Take for example Neurath's example of the swords in the sky. For the astronomer, the statement can be used as evidence of a cosmological event in history. But the same protocol could be used by an historian or anthropologist studying the attitudes of medieval peoples to astronomical and meteorological phenomena for entirely different purposes. They are still usable as evidence, although not via the extraction of the object sentence. Non-valid protocols therefore supply a different kind of evidence, but valuable evidence all the same.

Depending on which of conditions *(ii)–*(iv) are and are not met, and on the purposes for which the evidence is being used, we make use of information from non-valid protocols in different ways. Neurath provided tables with such categorisations in mind (Neurath 1941/1983; 220):²⁷

²⁷ The schema is taken from Neurath, but changes have been made for consistency with Uebel's terminology as adopted here. Neurath's proof reading appears to have been inattentive: his 'C (zebra)' should have been (iv), not (iii), and his 'D (person perceiving)' should have been (iii), not (iv). Accordingly, in the case of hallucinatory statements, the verdict of line 3 should have been 'accepted', not 'rejected', and that of line 4 'rejected', not 'accepted'. (Also note that in the case of lying, lines 3 and 4 could also receive verdicts opposite to the ones indicated here and in the original, as Neurath recognised with the inclusion of 'Type of').

Factual:	Hallucinatory:	(Type of) Lying:
* <i>(i)</i> *, <i>(ii)</i> , * <i>(iii)</i> *, * <i>(iv)</i> : accepted	* <i>(i)</i> , * <i>(ii)</i> , * <i>(iii)</i> , * <i>(iv)</i> : accepted	* <i>(i)</i> , * <i>(ii)</i> , * <i>(iii)</i> , * <i>(iv)</i> : accepted
* <i>(ii)</i> , <i>(iii)</i> *, * <i>(iv)</i> : accepted	* <i>(ii)</i> , * <i>(iii)</i> , * <i>(iv)</i> : accepted	* <i>(ii)</i> , * <i>(iii)</i> , * <i>(iv)</i> : rejected
* <i>(iii)</i> , * <i>(iv)</i> : accepted	* <i>(iii)</i> , * <i>(iv)</i> : accepted	* <i>(iii)</i> , * <i>(iv)</i> : rejected
* <i>(iv)</i> : accepted	* <i>(iv)</i> : rejected	* <i>(iv)</i> : accepted

These tables would be entirely superfluous if Neurath's conditions are intended solely for determining the validity of protocol statements. We know that all four conditions must be met, so why bother distinguishing different types of non-valid protocols? But when we recognise that non-valid protocols are themselves potential evidence, there is obvious utility to being able to categorise different types of non-valid protocols, and the uses scientists make of them. With this understanding of Neurath's protocol-tables, we begin to see how conditions **(i)*–**(iv)* function not merely as criteria for determining validity, but as the basis for an empirical methodology for schematising the usage of different types of scientific reports. After differentiating different varieties of non-valid protocol, the scientific meta-theorist can then undertake the task of an empirical and historical study of their uses, such that regularities can be recorded. With such a body of data, recommendations for their more effective utilisation through the adoption of certain practices or methodological norms could be made.

Once again, we see the active role of the Neurathian epistemic agent. We are not receivers of discrete bundles of information through a process of passive transmission. We are active in the process of deriving, interpreting and integrating the information provided by the protocol statement. And again, how this information is incorporated is a matter of decision. Crucially, both reports in the protocol-bank and data from the data-bank can be utilised by scientists, albeit in different ways. The protocol-bank acts as a permanent source of potential evidence, for use in inferential or inductive reasoning. The data in the data bank function as scientific facts, accepted statements of a scientific theory.²⁸

Let's return once again to Neurath's examples of the swords in the sky. This example is another instance in which a non-valid protocol is a useful piece of scientific evidence. Its acceptance conditions stand as follows:

- **(i)*: accepted
- **(ii)*: accepted
- **(iii)*: accepted
- **(iv)*: rejected

This results, according to Neurath's tables, in categorisation as an hallucination statement. One initial lesson to be drawn here is that Neurath's category of 'hallucinatory statements' ought to be understood as widely as possible, to include hallucinations but also more common misidentifications of the apparent object of

²⁸ To call them facts here is not to imply that they are true, or unchanging. Rather, I mean "scientific fact" in the sense of a statement that within a specific paradigm/encyclopedia are simply accepted.

perception. Perhaps a more appropriate name for the category would be “misinterpreted sensory stimulations”.

Given our current body of evidence we conclude that the report was publicly protocolised, is sincere, was authentically formulated, and reports a genuine sensory occurrence but is rejected on the basis of the negative-coherence condition. On what grounds? The exact thing for which the report itself is being used as evidence: that what caused the report was actually a meteor shower. The use we make of the sword report is as a basis for abductive reasoning, or inference to the best explanation (IBE). The significance of IBE in the reception of testimony was stressed by Peter Lipton. Testimonial inference to the best explanation (TIBE) is differentiated from typical IBE, ‘because it takes the central datum to be explained not to be some natural phenomenon... but rather the fact that the speaker said what she did’ (Lipton 2007, 243). As Nick Jardine emphasises, this inferential process is typically contrastive. Jardine frames the process of abductive reasoning as genealogical: it attempts to establish the origins of the instance of testimony under discussion. We accept testimony when it has a ‘vindictory genealogy’ that better explains the occurrence of testimony than any available ‘critical genealogy’ (Jardine 2008, 167).

In Neurath’s example, our posit of meteor-showers explains why medieval people reported seeing fiery swords in the sky. And the occurrence of a meteor shower is concluded to be a better explanation, in light of our encyclopedia, than the report itself being valid. The occurrence of a meteor shower is consistent with our understanding of gravity and the properties of swords in ways the medieval observation report is not. In this case then, the critical genealogy is preferred to the vindictory. And yet despite this critical conclusion, the report nevertheless retains utility, just as Neurath argued it did. The same is true of Gamow’s guest-star case (Gamow 1949, 19). Taken at face-value, the protocol must be rejected. But as the basis for abductive reasoning, the report functioned as an excellent piece of evidence, and was crucial in leading to the vindictory experiments of contemporary astronomers.

3.8 Conclusion

What should be clear by now is that, whilst initially unintuitive, Neurath’s conception of and schema for protocol statements provides a rich account of the nature of scientific testimony, and the process of receiving and appraising scientific observation reports. The conception of protocol statements Neurath arrives at is a testament to the value of his pragmatics of science as a methodology: the insights gleamed from attention to the actual and historical practices of science are the basis for the unique insights that his account of scientific evidence supplies. Most strikingly, Neurath’s conception of protocol statements anticipates an ongoing area of research in the philosophy of science which emphasises the significance of the context of scientific practice for understanding the relevance of scientific evidence. Nora Mills Boyd has recently argued that for empiricism to keep pace with the

development of modern science, the conception of scientific evidence must go beyond the notion of personal experience, and be enriched to include metadata, data recording the circumstances under which evidence was recorded and how it has been processed (Mills Boyd 2018, 407). The spatio-temporal indexing of the circumstances in which observations are made, encoded in Neurath's protocol statements and preserved in the data-bank are metadata of exactly this sort. Mills Boyd even highlights how such metadata allows for the use of evidence beyond their original epistemic context, just as Neurath argues that his protocol's have the advantage of stability. Neurath's work on protocols prefigures this ongoing work in both spirit and content.²⁹

²⁹ For more on Neurath's conception of evidence, and its similarities to contemporary work in the philosophy of science, see Bentley (2022).

Chapter 4

Carnapian Explication



In the previous two chapters we saw what Neurath's naturalistic pragmatics of science can do: provide a theory of empirical evidence that ultimately combines the relevant insights of the psychology of perception with the sociology and history of institutionalised collective endeavours with normative directives arrived at in the light of agreed objectives. Let us now turn to Carnap. What contributions might a logician of science be able to make beyond simply analysing the logic of pre-existing scientific theories? Important as that task may be, one must also ask whether Carnap's logic of science can, in a comparable and complementary way to Neurath's pragmatics, make a positive contribution towards shaping the future development of scientific practice. The answer, as will be demonstrated, is yes. Carnap's method of explication allows the logician of science to contribute to the scientific endeavor both and increased awareness of, and consequent deepening, of scientific practice but also the provision of linguistic and logical innovations.

4.1 Explication and the Bipartite Metatheory

The later Carnap and his proposal for the method of explication was for a long time overlooked, lazily dismissed, or subjected to a straw-man caricature and consequent simplistic rejection. In recent years, attention has returned to Carnap's explicationism. For some, this has been simply to defend the plausibility of the project in principle (Carus 2007a; Brun 2016, 2017; Dutilh Novaes and Reck 2017). Others have gone a step further and have begun applying the explicationist methodology to specific topics in philosophy and science, including probability (Maher 2010), ecology (Justus 2012), biology (Kitcher 2008), experimental philosophy (Shepherd and Justus 2015) and the Gettier problem (Olsson 2015). This recognition is certainly overdue. But whilst these recent studies understandably focus on explication as a method, they don't typically emphasise the role of explication in Carnap's

broader philosophy. The exception here is Carus who situates the explicationist methodology as the centre of Carnap's entire philosophical project as a renewal of the Enlightenment project. While I am very sympathetic to this perspective, my aim in this chapter is to discuss the method of explication specifically relative to the bipartite metatheory project, and Carnap's logic of science specifically. This means demonstrating that naturalistic impulses that animate Carnap's explicationism.

The purpose of this chapter is twofold. First, drawing on the recent work in favour of explication, I aim to show that the none of the critiques of Carnapian explication hold up to scrutiny. Secondly, I aim to show that they fail primarily because they misinterpret Carnap's proposals. Broadly, the criticisms of Carnapian explication can be separated into two types. The first type are specific technical criticisms, concerned with the actual applicability of Carnap's proposals. These, I argue, are relatively easily combatted, and can frequently be shown to be based on a misunderstanding or oversimplification. The second sort, the "deeper" sort, amount to the demand for traditional philosophical answers from a naturalist methodology. There is very little that can be said to such criticisms, beyond reverting back to a defence of naturalism itself. To show this, I have addressed two main critics. One, Strawson, is a contemporary of Carnap. The other, Reck, is a current critic of Carnap. By focusing on these two, I hope to show how little the criticisms of Carnapian explication have changed, and how they fundamentally rely on the assumption of anti-naturalism. In so doing, I aim to show the viability of Carnapian explicationism as a method for epistemological naturalism.

4.2 Carnapian Explication

4.2.1 *Explication: Process and Goals*

'The task of explication consists in transforming a given more or less inexact concept into an exact one' (Carnap 1962, 3). This requires the introduction of a new concept, the meaning of which is exactly specified, which is intended to capture as closely as possible the spirit, usage and scope of the pre-existing concept, in such a way that the explicatum (the new concept) supplants the explicandum (the pre-existing concept) for specific purposes. Brun calls this a process of 're-engineering' (Brun 2016, 1211). To make the creation of an exact concept possible, the explicatum will often be within a formal language-system. The fundamental difference between analysis and explication is that where an analysis seeks to find truth, the correct understanding of a pre-existing concept, an explication is a proposal for a new, albeit functionally similar, concept.¹ Where the analyst is a detective, the explicationist is an engineer.

¹ Brun raises the seemingly banal but easy to overlook point that not only is explication only one form of conceptual engineering, but that sometimes explication simply may not be possible (Brun 2016, 1234).

The provision of an explication is both descriptive and stipulative, with neither taking unilateral pre-eminence.² So ‘the question whether the solution is right or wrong makes no good sense because there is no clear-cut answer. The question should rather be whether the proposed solution is satisfactory’ (Carnap 1962, 4). When deciding on whether to accept or reject a proposed explication, it is ultimately a matter of pragmatic judgement when comparing the utility of different explications of the same concept. The most important element in making these decisions is the purpose for which the explication is intended, because it is only with this purpose in mind that the instrumental value of the explicatum can be judged on practical grounds. It is for this reason that making clear one’s purposes is such a crucial part of the informal first phase of explication.

Inexact terms often have a multiplicity of uses, and each may well require different explications. Wittgenstein’s famous example of family resemblance concepts, held together by overlapping relations of similarity and shared features rather than any singular unifying criterion would perhaps welcome a multiplicity of explications. Perhaps the various facets of “game” that Wittgenstein rightly locates within our intuitive concept could be fruitfully disentangled and sharpened via Carnapian explication. It is therefore possible that some terms of natural language will, through the method of explication, give way to multiple similar new terms, each introduced to capture one specific purpose of the original term. An obvious candidate for such disambiguation would be “man”. In this term, the sociological gendered component, biological sexual component, and anthropological-historical usage as equivalent to “mankind” are all separable (and there are probably others).³

Explication as a process can be separated into two main phases. The first, the ‘informal clarification of the explicandum’, involves specifying the term to be explicated (Carnap 1963b, 933). Since the explicandum is in need of clarification, and therefore inexact, a perfectly precise definition will not be found. Rather, the concept is ‘made as clear as possible by informal explanations and examples’ (Carnap 1962, 3). Exemplary cases and borderline cases can be highlighted. These, along with the degree of vagueness in a pre-existing concept, can be roughly established through some empirical linguistics. We can survey people and ask them, for a specific term, to choose three classes: those to which the term applies, those to which it does not, and those for which they are unsure. The larger this third class, the vaguer the predicate (Carnap 1955a, 35–36). But the process is not simply extensional. The uses and conceptual core of a concept are also highlighted. For those vague concepts with multiple uses, these will need to be disambiguated, distinguished and the relevant one specified. Such concepts will allow of multiple non-exclusive explications. In tandem with this, the purposes of the explication itself

² This is a key difference between explication and definition. According to Brun, definitions are either reportive (capture current usage) or stipulative (introduce a new usage). The process of explication may involve stipulative definition, but is not equivalent to it Brun (2016, 1229–32).

³ As both Haslanger and Dutilh-Novaes have argued, this kind of conceptual engineering and ameliorative definition can be put to progressive social and political purposes too (Haslanger 2007; Dutilh Novaes 2020).

needs to be specified. We need to know what use the concept being engineered will be put to. Shepherd and Justus have plausibly suggested the possible utility of X-phi in fulfilling these functions, which they call ‘explication preparation’ (Shepherd and Justus 2015, 391). There is however a balancing act here. Too little specificity makes explication impossible, but some inexactness is inherent. Just as we don’t need to fix something that works, we don’t need to explicate clear concepts. The method of explication exists in a Goldilocks zone of concepts that are not-too-vague but not-too-clear.

The second phase in the process involves the introduction of a new clearer concept, the explicatum, through the specification of the rules for use of this new concept and how it fits into our system of concepts. The success of this explication can then be judged according to four criteria (Carnap 1962, 5–8):

- i. Similarity: how closely it captures the usage and spirit of the explicandum
- ii. Exactness: how well elaborated the rules for use are
- iii. Fruitfulness: how useful it is for the practice of science, in particular for forming universal statements
- iv. Simplicity: how simple the definition is, and the laws that connect it to other concepts are

Of these criteria, simplicity is the least important, most frequently being relevant as a means for deciding between competing alternative explicatum (Carnap 1962, 7). Similarity may *prima facie* seem the most important. Obviously, complete synonymy of explicandum and explicatum would render the explication a pointless enterprise, so some divergence must be allowed. As Dutilh Novaes and Reck emphasise, fruitful explication requires a sacrifice of similarity (Dutilh Novaes and Reck 2017, 206). But the degree of divergence Carnap allows can be significant, more than intuition may initially suggest. Not only does the principle of tolerance licence experimentation, but study of the practice of scientists shows that sometimes these serious deviations are the most productive for scientific practice (Carnap 1962, 6). What Carnap could have stated more explicitly is that the requirement of similarity is always relative, ‘different requirements for different situations’ (Carnap 1963f, 945). In each case, the requirement is “similar in the relevant sense”. The similarity requirement is therefore determined by the use to which the explicatum is to be put, because it is the use we wish to make of the explicatum that determines which elements of the explicandum it must be similar to.

As Justus highlights, precision is crucial for Carnap because it is only non-vague statements that can generate confirmable or disconfirmable hypotheses, they are the basis of testing (Justus 2012, 169). This again is not an abstract demand, but a demand based in the actual practices of science. Precision therefore provides fruitfulness. So precision is not so much valued in itself, but as a means for achieving fruitfulness (Shepherd and Justus 2015, 388). It becomes clear then that the priority of explication is fruitfulness, that the other criteria are somewhat subservient to it, and desirable potentially as a means to it.⁴

⁴ The primacy of fruitfulness is also recognised by Dutilh Novaes and Reck (2017, 202).

What the importance of fruitfulness shows is that for Carnap, explication is a means not an end. Explication is a process to be undertaken with a specific purpose in mind, against which this fruitfulness can be judged. But ultimately, in weighing these criteria against one another, there are no hard and fast rules, nor means for quantifying relative significance in a concrete situation. There is an inescapable ‘conventional component’ in explication (Carnap 1962, 6). Every explication involves instrumental decision-making: ‘In my view, a language, whether natural or artificial, is an instrument that may be replaced or modified according to our needs’ (Carnap 1963b, 938). How we decide to change our language depends on the goals to which the change is oriented.

Despite it being of paramount importance to his explicationist methodology, Carnap gives very little detail about his requirement of fruitfulness. He does say that fruitfulness is comparative, a ‘matter of degree’, as the explicatum is judged relative to the explicandum (Carnap 1956a, 62). The desiderata may therefore more accurately be captured as “more fruitful than the explicandum”. Carnap also says that ‘A scientific concept is the more fruitful the more it can be brought into connection with other concepts on the basis of observed facts’ and, a concept is fruitful if ‘useful for the formulation of many universal statements’ (Carnap 1962, 6–7).⁵ Clearly then, fruitfulness is connected to a concept’s capacity to generate testable statements. Fruitfulness is directed at the conduct of empirical science. But, as Dutilh Novaes and Reck have highlighted, Carnap’s own examples imply a more developed notion of fruitfulness. As Dutilh Novaes and Reck understand him, Carnap here suggests that a more fruitful concept ‘delivers “results” that could not be delivered otherwise’ (Dutilh Novaes and Reck 2017, 206). A more fruitful concept doesn’t simply generate testable statements, but statements that would not have been generated (or at least not as easily) with the original concept. For Carnap then, explication is ‘a cognitive tool leading to discoveries and new insights’ (Dutilh Novaes and Reck 2017, 206).

Reck and Dutilh-Novaes argue for a ‘paradox of adequate formalization’, that there is a tension between the criteria of similarity and fruitfulness (Dutilh Novaes and Reck 2017, 196). Specifically, ‘an adequate formalisation is one that is faithful to the target phenomenon *and* reveals something new about it; there is an obvious tension between these desiderata’ (Dutilh Novaes and Reck 2017, 211). But the problem here is overstated. For one thing, we know where priority is typically weighted: fruitfulness. For another, tension between desiderata and paradoxicality are totally different things. There are many contexts in which desiderata are in opposition and cannot be maximised simultaneously, necessitating compromise. A famous metaphor from football management describes poor teams as like a short blanket: they can’t cover all areas effectively. As a result of a team’s limitations, the two desiderata of attacking threat and defensive solidity are in tension. But to claim that bad football teams are therefore paradoxical is ludicrous.

⁵ There are obvious similarities with Carnap’s prioritization of inter-connection to Neurath’s rejection of isolated statements as meaningless.

Dutilh Novaes and Reck adopt the language of “paradox” because they draw a comparison to the paradox of analysis. But there is a crucial disanalogy which they do not attend to. The paradox of analysis says that a successful analysis requires both correctness (that the analysans and analysandum must be identical) and informativeness (which requires a difference between analysandum and analysans). It therefore seems that no analysis seems capable of being both correct and informative. They rightly note that Carnap would reject the notion that a correct analysis requires identity (Dutilh Novaes and Reck 2017, 211). But they don’t explore why, and this is a crucial difference.⁶ Analysis starts from the assumption that there is a “real” meaning of a concept. But Carnap simply doesn’t make this assumption. This divergence, as we will see below, is fundamental.

4.2.2 *Criteria for Explication*

The criteria Carnap provides for judging the success of explication have been challenged. Hanna argues that Carnap’s notion of explication is too broad, and thus too vague (Hanna 1968, 34). Similarly, Bonioli argues that Carnap’s criteria for explication are too ‘imprecise’, with three of Carnap’s four criteria for explication (similarity, fruitfulness and simplicity) being ‘concepts that no philosopher of science... has ever been capable of formulating with precision’ (Bonioli 2003, 291–292). So, can we clear up what Carnap means? To a degree, yes. Simplicity is the least important, and I think the least controversial. Firstly, the criteria of simplicity is used primarily as a deciding factor. So we only have to be able to compare relative simplicity of competing explications.

More difficult is similarity. In the context of Carnapian explication, the demand for strict criteria for similarity is misguided. The criteria of similarity is not rigidly enforced by Carnap. And we have to remember that explication is always pragmatically guided, determined by the specific purposes for which the explicatum is intended. I think a definition of similarity would be something like:

An explicatum is sufficiently similar to the explicandum, for purposes X, if the explicatum is usable in most cases, for purposes X, in which the explicandum was used

Now immediately, the use of “most” will ring alarm bells as too vague. But it is inherent to the pragmatic nature of explication. ‘Conceptual flexibility is at the heart of explication: the license to deviate from the intuitive meaning of concepts to reap epistemic rewards’ (Justus 2012, 162). It is this flexibility that makes explication so rich in possibilities. This is why Stein refers to certain pre-systematic notions as ‘usefully vague’ (Stein 1992, 281). Rigid standards for similarity would remove this.

⁶ This paradox is missing in Dutilh-Novaes’s more recent paper, she may have reconsidered (Dutilh Novaes 2020).

What about exactness? Well, critics treat Carnap as if he only acknowledges the ideal case, in which an explication is provided within an exactly specified formal language. In such a scenario, complete precision could be provided in the explicatum. This is the ideal scenario, but Carnap was realistic enough to recognise that most science is not conducted in such languages.

The use of symbolic logic and of a constructed language system with explicit syntactical and semantical rules is the most elaborate and most efficient method. For philosophical explications the use of this method is advisable only in special cases, but not generally (Carnap 1963b, 936)

He acknowledges that explications can be provided in natural language too, although such explications would not attain the complete precision of those provided in a formal language. In those cases, the criteria is not complete clarity, but greater clarity. ‘The only essential requirement is that the explicatum be more precise than the explicandum’ (Carnap 1963b, 936). There is no specific degree of clarity required universally. Explication succeeds in so far as the explicatum is clearer than the explicandum. But perfect clarity within an exactly specified formal language still remains as the ideal.

It is regarding fruitfulness that the criticism seems most accurate, and most damaging. Carnap does not give much detail on how we are to interpret the demand of fruitfulness, and what he does give us is narrowly focused on the formal cases. Specifically, he gives the definition that ‘the more it can be used for the formulation of laws’, the more fruitful an explication is (Carnap 1963b, 6). However, this clearly doesn’t exhaust the notion of fruitfulness. Carnap does provide some further suggestions though: ‘A scientific concept is the more fruitful the more it can be brought into connection with other concepts on the basis of observed facts’ (Carnap 1963b, 6). In other words, fruitfulness is measured or determined by the provision of confirmable (or perhaps well-confirmed) empirical hypotheses. As Shepherd and Justus point out, this is the same way ‘in which science measures epistemic success’ (Shepherd and Justus 2015, 382).

I think however, that this on its own is too narrow and too formalistic a conception of fruitfulness to be universally applicable. The notion of fruitfulness as understood in the sciences typically includes a value for novelty. Similarly, a notion of fruitfulness should discriminate between valuable and trivial generalisations. Fruitfulness is not only quantitative, in terms of how many predictions are provided, but also values new predictions specifically because they are new. This is not mentioned by Carnap, or Justus and Shepherd, but I do think it needs to play a role in evaluations of fruitfulness. And yet, there seems to be no good way to give priority to one over the other in general. In different contexts and for different purposes, it is possible that we may well value the quantitative sense of fruitfulness over novelty, and vice versa. There just is no obvious basis for making hard and fast criteria like this. To demand such criteria amounts to pseudorationalism. I agree with Shepherd and Justus that it is with this in mind that Carnap avoids a detailed elaboration of the requirement of fruitfulness, despite its obvious significance (Shepherd and Justus 2015, 397).

4.2.3 *Explication as Naturalism*

As established in Chap. 1, one aspect of the naturalized epistemology adopted by the left wing of the Vienna Circle is the acceptance of the methodological continuity of unified science. All the special sciences have the same array of methods available to them. This is not to say that all sciences must be conducted identically. Some techniques will be more relevant and more fruitful for certain special sciences than others (physics and anthropology for example require very different methods for collecting and analysing data). But importantly, there is no difference in kind. There is no cleavage between empirical science and a priori meta-reflective philosophy. All and only the methods of the sciences are available to Carnap, but this is perfectly consistent with his later work. The formal a priori methods available to scientists are also available to the metatheorist.

But more concretely, as Lutz has also recently argued, the specific methodology of conceptual engineering itself has precedent within science. Carnap himself never claimed to have originated the methodology, he acknowledged that ‘scientists, and mathematicians make explications very frequently’ (Carnap 1962, 8). Carnap gives the examples of the explication of “Pisces” from “Fish”, according to the desire for greater precision and specificity required by zoologists, and the explication of the quantitative concept “temperature” from the pre-scientific qualitative concept of “warm” (Carnap 1962, 8–10). Lutz rightly argues that whilst epistemological naturalism typically assumes scientific claims are synthetic, and that philosophical claims ought to follow suit, this is an unnecessarily limited approach (Lutz 2020, 2). Specifically, he argues that conceptual engineering, the provision of ‘suggestions for (typically new) concepts rather than analyses of concepts we currently have’, should be understood as an instance of naturalised epistemology that deals with conventional analytic statements (Lutz 2020, 2). He is right, and the sort of conventional analyticity he describes is the sort I will argue that Carnap and Neurath allow. Lutz also notes some interesting cases from recent studies of the history and methodology of science which arrive at conclusions like Carnap’s, that conceptual engineering plays a key role in science (Lutz 2020, 15–16).

But Carnap’s explicationism is not simply compatible with epistemological naturalism in general. Carnap’s naturalism has clear intersections with Neurath’s specific naturalistic epistemology of science, as outlined in Chap. 2. This is supported particularly by Carus’s account of Carnap’s explicationism. Possibly their most obvious point of agreement is their understanding of language as fundamentally instrumental, and frequently described as a tool. Not quite as obvious however is Carnap’s agreement with Neurath that gradual improvement must not be misunderstood as approximation of an ideal system, but as a process of indefinite and incremental adjustment:

my work and that of a few friends in the step for step solution of problems should not be regarded as leading to “the ideal system”, but rather as a step for step improvement of an instrument (Carnap 28/04/1962, quoted in Reisch 1991, 267)

Explicationism also exhibits a constructive decisionism akin to that we saw from Neurath. The role (and necessity) of non-rule-guided decisions is as central to Carnap's explicationism as Neurath's picture of science would lead us to believe. And Carus also recognises in Carnap the role of the active, constructive epistemic agent that I attributed to Neurath. For Carnap, we are 'not just representors, and consumers of representations, but also *agents* in the world so represented' (Carus 2017, 179). But perhaps most fundamentally, Neurath's definitive image of the scientific process is also applicable to Carnap's process of explication. A particular phrasing of Carus's makes this clear: 'The ideal of explication is one of piecemeal upgrading; it starts *in medias res*, not from first principles' (Carus 2007a, 276). Or, to put it another way, explication takes place on board Neurath's ship. Just as Neurath argues it must, Carnap's explicationism takes the natural language as it exists in the now—socially-constructed, historically contingent, and riddled with unclarity—and recognises that this language can be our only starting point. And, as we will see below, the adoption of explicationism as a method is a tacit endorsement of Neurath's rejection of pseudo-rationalism. There is fundamental agreement here.

Recognising this fundamentally naturalist outlook is, I argue, essential to understanding and moving beyond many of the criticisms that have continued to plague Carnap's method of explication. It is to these we now turn.

4.3 Carnap's Critics

Perhaps the most famous and influential critique of Carnapian explication was Strawson's 1963 response to Carnap. Strawson's arguments are frequently repeated and endorsed by more recent critics of Carnap. However, these arguments lack penetration on account of a fundamental misunderstanding of Carnap's purposes. In presenting these criticisms, not only will it be shown how not to interpret Carnap, but in Carnap's replies we can see fully the naturalistic reasoning at play in his explicationist methodology.

The core of Strawson's criticism is that to explicate 'is not to solve the typical philosophical problem, but to change the subject' (Strawson 1963, 506). For Strawson:

however much or little the constructionist technique is the right means of getting an idea into shape for use in the formal or empirical sciences, it seems *prima facie* evident that to offer formal explanations of key terms of scientific theories to one who seeks philosophical illumination of essential concepts of non-scientific discourse, is to do something utterly irrelevant – is a sheer misunderstanding, like offering a text-book on physiology to someone who says (with a sigh) that he wished he understood the workings of the human heart. (Strawson 1963, 505)

Reck similarly claims that for all explication's successes, there is a 'philosophical residue' left over which explication cannot resolve (Reck 2012, 114). For Strawson, solving philosophical problems is only possible by paying close attention to the concepts that generate them. By looking for new concepts we ignore both the real problem and its source.

There is obvious disagreement between Carnap and Strawson over the appropriate degree of similarity that must be retained in conceptual analysis. For Strawson, the degree is high, and trumps all other factors. Carnap's other criteria (exactness, fruitfulness, simplicity) are irrelevant if extreme similarity of meaning is not preserved.⁷ Where for Carnap the loss of synonymy via a change in meaning but a consequent increase in the fruitfulness and exactness of a concept is not only permissible but advisable, for Strawson this loses the philosophical point of such analysis in the first place. So for Strawson, Carnap's weak similarity requirement is simply not fit for purpose. There is a uniquely philosophical issue, above and beyond the task of finding appropriate scientific terminology, rooted in the existing concepts of natural language.

What is Carnap's response to the accusation that explication amounts to changing the subject? He argues that explication can provide the solution to philosophical problems. He also attempts a reconciliation of his own method with Strawson's, granting the relevance of Strawson's methods in the initial clarificatory stages of explication, making the two approaches 'mutually complementary' (Carnap 1963b, 940).

First, Carnap rebukes Strawson's characterisation of when philosophical problems arise. Strawson describes the need for philosophical elucidation as arising especially in cases where a desire for further elucidation accompanies mastery of the usage of a concept (Strawson 1963, 509). Carnap argues that this self-confidence is frequently deceptive, and that we are not certain in the usage of these contested terms. He argues that while these terms are used for the most part without difficulty, in certain circumstances this ease of usage breaks down (Carnap 1963b, 935). For example, most people are comfortable using "True" and "False" until faced with the Liar Sentence. What these cases illuminate is that the usage was not unambiguous to begin with. These are not uniquely important cases of a special philosophical confusion, but indications that our original confidence in our mastery of the concept was unwarranted.

Carnap also rejects the accusation that his method amounts to a change of subject, pithily utilising the image of languages as tools:

[Strawson's] thesis is like saying that by using a special tool we evade the problem of the correct use of the cruder tool. But would anyone criticize the bacteriologist for using a microtome, and assert that he is evading the problem of correctly using the pocketknife? (Carnap 1963b, 939)

There is nothing sacred about natural language that demands preservation, it is simply a tool for us to use. This tool analogy has been controversial:

Carnap's analogy allows for the obvious rejoinders that (1) pocketknives are not replaceable by microtomes for most ordinary uses and (2) someone who was having trouble using a pocketknife in an ordinary circumstance would not be helped in the least by being shown the workings of a microtome. So it is not obvious that Carnap's analogy adequately answers

⁷ Strawson stops short of expecting complete synonymy from conceptual analysis, but any deviations are clearly expected to be minimal.

Strawson's charge of the irrelevance of explication for unraveling perplexity involving ordinary notions. (Loomis and Juhl 2006, 291)

These obvious rejoinders themselves have obvious rejoinders.

Whilst (1) is true, it is irrelevant. A microtome is a highly specialised scientific instrument, used for specific experimental purposes. It was not designed to slice bread. An explicatum is also introduced with a specific purpose in mind. But an explicatum analogous to the microtome would be a highly specialised term of scientific jargon, not intended for use in domestic life. Carnap doesn't expect this term to become a part of everyday language. As we already saw, explication involves refinement, not necessarily replacement. This criticism relies on a caricature of Carnap's position. He never intended the vocabulary of specialised science to replace everyday language in everyday contexts. And this is clear in Carnap's own use of the metaphor: 'we shall try to discover the cause of the failure, and then either use the knife more skilfully, or replace it *for this special purpose* by a more suitable tool' (Carnap 1963b, 939, emphasis added).

Loomis and Juhl's argument mirrors one made by Strawson:

it seems in general evident that the concepts used in non-scientific kinds of discourse could not literally be *replaced* by scientific concepts serving just the same purposes; that the language of science could not in this way *supplant* the language of the drawing-room (Strawson 1963, 505)

This claim implies two significant mischaracterisations. First, Carnap never argues that explication is a process of supplanting and eliminating. For Carnap, providing an explication does not necessitate abandoning the explicandum altogether. 'That the explicandum Fish has been replaced by the explicatum Pisces does not mean that the former term can always be replaced by the latter' (Carnap 1962, 6).⁸ Rather, the explicatum replaces the explicandum *for certain purposes*. But this one explicatum may not capture all the myriad uses of a vague pre-systematic term. Just as we can retain older drafts of an essay, we can retain older versions of a concept, as they may be a useful basis for further explications, or simply still be useful for the practicalities of everyday life. Carnap recognises the extent to which vagueness pervades our natural language, and whilst he thinks the provision of more precise terms will be useful, he is never so unrealistic as to hope that all human interaction will ultimately be replaced by perfectly clear discourse.

Carnap also never pretends that all uses of a vague concept can always be captured in a single explication. As mentioned above, he recognises that one concept may require multiple alternative explications, each directed at different ends, with each explicatum capturing a different usage of the explicandum. He gives the example of the prescientific classificatory concept "Warm", which can be explicated as both the comparative concept "Warmer" and the quantitative concept

⁸ This quote comes from the work Strawson is specifically responding to. Strawson presents as criticisms what Carnap has readily accepted already.

“Temperature” (Carnap 1962, 13).⁹ Which explicatum we opt to use is guided by our purposes: are our goals best suited to the use of a comparative or a quantitative concept? This leads us to the second of Strawson’s misrepresentations, that the explicatum must replace the explicandum in all cases. That this is false is directly contradicted by Carnap in the above quote. And yet, these accusations continue to reappear.¹⁰

As for (2), there is an ambiguity here. It could be understood in one of two ways. One way is that someone has a knife and a purpose in mind, but lacks the skill-knowledge of how to use it. This would then be an irrelevant case of the need for know-how, which cannot be supplied by any definition or analysis. The second way involves a person with a knife who wants to know the use of a knife. Here we simply have a case of ignorance. This isn’t even Strawson’s exemplar of philosophical questioning, in which familiarity with use accompanies the desire for further questioning. It’s simply a case of ignorance. What (2) really argues is that explication will not help if we lack a sufficient understanding of the explicandum to begin with. This is, again, true but irrelevant. Carnap is aware of the need for sufficient familiarity with the explicandum prior to explication. And he never suggests that the purpose of an explicatum is to clarify the explicandum for someone entirely ignorant of it.

However, as Maher points out, Carnap’s analogy here is not entirely without the need for clarification. The bacteriologists problem was not the pocketknife’s true purpose, but a problem of medical science which required a more specialised tool. But for the critic, it is exactly the proper use of the pocketknife that is at stake (Maher 2007, 333). Maher provides an alternative response on Carnap’s behalf. He argues that whilst explication by itself may not be sufficient to answer the original philosophical problem, explication can help us on the way to finding such a solution by ‘provid[ing] insights and lines of argument that we may not discover if we reason only in terms of the vague explicanda’ (Maher 2007, 334). And if it is helpful, then it can’t be irrelevant. It therefore seems that there is the potential for complementarity rather than opposition between Carnap and Strawson. This is further reinforced by Carnap, whose response to Strawson stresses the possibility of reconciliation. Carnap admits the need for an understanding of the explicandum prior to explication, especially the uses to which it is put, and that Strawson’s analyses of natural language are well-suited to this purpose. Insofar as they are empirical, analyses of actual usage like Strawson’s, which Carnap calls ‘linguistic naturalism’, have a key role in the ‘informal *clarification of the explicandum*’ (Carnap 1963b, 933).

However, this is as far as the reconciliation of the two approaches can go. As Justus highlights, ‘Strawson utilizes a more substantive conception of conceptual

⁹ Another example comes from Carnap’s collaboration with Bar-Hillel. They conclude that ‘there is not *one* explicandum [of] ‘amount of information’ but at least two’ (Bar-Hillel and Carnap 1953, 150).

¹⁰ See for example Eagle (2004, 372).

analysis than the empirical version Carnap could accept' (Justus 2012, 173). For Carnap, the study of natural language is limited in two ways: it is only a preparatory step that lays a basis for explication, and it is purely descriptive. The study of natural language can inform us only how concepts are used. This directly contradicts Strawson, who thinks that the distinctly philosophical task of analysis 'attempts to show the natural foundations of our logical, conceptual apparatus', the discovery of which brings with it normative force (Strawson 1963, 516). For Strawson, the study of natural language provides not only descriptive knowledge, but prescriptive knowledge of how concepts ought to be used. The force supplied is sufficient to rule against any deviation from it, which explains Strawson's strictness over the similarity requirement (Justus 2012, 173). Here there is a fundamental divergence between Strawson and Carnap. As we saw, for Carnap explication is never right or wrong, but more or less useful. For Strawson, there is a correct use of a concept, derived from its natural foundations. There is an objectively right and wrong way for concepts to be used. By contrast, Carnap's proposals are only ever instrumentally normative; an explication should be adopted if one has specific purposes in mind.

With these differences between the projects of Carnapian explication and Strawsonian analysis in mind, we will now return to Maher's defence of Carnap. Maher attempted to show that explication is not irrelevant, because it can be helpful for the analysis of natural language concepts. Maher describes how an explication of a term *X* can help determine whether or not a sentence of natural language containing *X* is in fact true or not (Maher 2007, 334). It should now be clear that Maher's defence of explication is not a defence of *Carnap's* explication, but rather an argument for the potential utility of something like explication for Strawsonian projects. Whether or not Maher is attempting to provide an answer Carnap would agree with is unclear. Regardless, he does not provide one. More importantly though, Carnap does not need the defence Maher provides.

Whilst Maher does not dwell on it, there is an important illumination provided by his expansion of Carnap's pocketknife metaphor. For Strawson, and seemingly Maher, Carnap's metaphor leaves out the essentially philosophical attitude. But for Carnap, the microbiologist's attitude is the only coherent attitude. He asks to what use a tool can be put, or what tool best suits his purposes. For the critic, this ignores the more basic philosophical question of what the tool's *real* purpose is. This absence from Carnap's arguments is not an oversight. The reply implicit in everything Carnap says, is that this question has no answer. There is no essential purpose to a tool, just as there is no essential meaning to a concept. The question Carnap's critics demand an answer to is the sort of specifically philosophical one that has no place in Carnap's project.

Here we arrive at the core of the critic's argument. Strawson identifies the failings of Carnap's method as a consequence of his failure to take the demands of philosophy seriously (Strawson 1963, 507). Carnap retorts that he wouldn't have devoted his life's work to solving philosophical problems were this the case (Carnap 1963b, 935). What is at issue here is how we understand this demand to take philosophy seriously.

At this point, it is necessary to separate two potential demands that a critic of Carnap could be making. One demand is that Carnap must supply an account of how concepts are used in the real world, rather than only supplying potential alternative definitions. A more fundamental demand is that conceptual analysis must attend to the real meaning of concepts in a way that goes beyond simply charting real world usage, but that attends to the inherent or essential meaning of a concept. Here a distinction introduced by Kitcher is useful. Kitcher describes the conflicting positions of conceptual monism and pluralism. Monists think there is a correct concept, they think our ‘classificatory task is like that of carving a beast at its joints’ (Kitcher 2008, 118). By contrast, pluralists reject the idea that there is one correct account of a concept. As we have seen, Carnap is explicitly pluralist, viewing concepts as instruments intended for specific purposes. Carnap, one suspects, would have agreed with Kitcher’s conclusion: ‘To think that there’s a structure set in nature to which concepts must conform is, I believe, an article of metaphysics... we are better off without’ (Kitcher 2008, 119).

One can plausibly interpret Strawson as a conceptual monist: ‘the philosopher labours to produce a systematic account of the general conceptual structure of which our daily practice shows us to have a tacit and unconscious mastery’ (Strawson 1992, 7).¹¹ This is also suggested by Strawson’s repeated use of “philosophical” as a qualifier. Concepts are ‘philosophically puzzling’ (Strawson 1963, 505). Scientific theories cannot provide ‘philosophical illumination’ (Strawson 1963, 505). There is an extra category of ‘philosophical problem’ (Strawson 1963, 506). Similarly, Reck argues for the necessity of ‘philosophical arguments’ (Reck 2012, 114). On such an account, vague concepts provide philosophical puzzles to be solved by conceptual analysis. The need for philosophical elucidation arises especially in cases where a desire for further elucidation accompanies mastery of the practical usage of a concept (Strawson 1963, 509). This embodies the quintessential philosophical nigggle, the feeling that one knows how a word is used, but doesn’t know what it *really means*. Providing new concepts free of these puzzles does not amount to a solution. Avoiding a problem is not the same as solving a problem. A real solution requires us to ‘understand the foundation of our concepts in natural facts’ (Strawson 1963, 517). Implicit in this, clearly, is that these are legitimate problems capable of being solved. On this account the method of analysis is more than merely descriptive, it is normative. Analysis provides both a correct description of the concept, but also rules for how the concept should be used.

The repeated use of the qualifier “philosophical” suggests a separation of the domain of “philosophy” as an autonomous field, with unique methods and answers. And this is the exact conception of philosophy that naturalism serves as a critical response to. There are no uniquely philosophical questions with uniquely philosophical answers. Concepts are simply instruments, and there is no deeper truth or reality to access. What Carnap is attempting in his response to Strawson is a form

¹¹ As Quine notes, the assumption of ordinary language philosophers like Strawson seems to be that philosophical analysis involves ‘the uncovering of hidden meanings’ (Quine 1960/2013, 238).

of philosophical deflationism. He tries to show that what appear to be uniquely philosophical confusions are linguistic. Carnap even uses a very Wittgensteinian turn of phrase, referring to the process of explication as 'therapy' (Carnap 1963b, 937). Where Strawson thinks philosophical problems can be solved through a greater understanding of our concepts, Carnap refers to philosophical difficulties as 'problems and perplexities' to 'eliminate' or 'overcome' (Carnap 1963b, 936–937).

The difference between these two approaches is that for Carnap, showing a philosophical problem to be the result of conceptual and linguistic confusion *is* the solution. In the case of most philosophical problems, 'what seems to be a clash between two opposing positions, is actually a difference in the interpretation of a concept' (Carnap 1963b, 934). In these cases, dissolution of the original confusion by clarifying our use of language is as much of a solution as can be requested.¹² There is no special domain of uniquely philosophical problems. Those questions that cannot be answered scientifically are confusions to be resolved by close attention to the use of language. In dispelling such illusions Carnap is taking philosophy seriously, albeit not in the traditional philosopher's way. Though Carnap, being forever diplomatic, did not say so, in Neurath's spirit we can recognise that Strawson's demand (at least if understood as above) is an essentially pseudo-rationalist one. He demands a final answer where there is none. He demands singularity where there can only ever be plurality.

It is possible however that I have mischaracterised Strawson here, and that his demands amount only to the less strenuous requirement of an adequate account of the everyday use of natural language concepts. If this is the case however, Strawson's claim that Carnap fails to take philosophy seriously has even less force. As we have seen, Carnap takes pains to accurately chart the use of natural language concepts prior to attempting an explication. This is the process he calls the informal clarification of the explicandum. In this sense, Carnap takes the issue of conceptual analysis exactly as seriously as Strawson does. Carnap does not ignore the requirement of close attention to actual usage, but he also sees a purpose in going beyond simply attending to current usage and proposing potentially more useful alternatives.

For those critics who are unhappy with this deflationist approach, who maintain that Carnap's method ignores the fundamentally "philosophical" question, Carnap simply rejects their outlook. Their disagreement with Carnap is a meta-

¹² A perfect example of this kind of disambiguating explication as a solution to a philosophical dispute is Carnap's own explication of "probability". Surveying existing accounts of probability, Carnap concludes that the 'assumption of a unique explicandum common to all authors is untenable' (Carnap 1945, 517). There are rather, broadly speaking, two concepts being referred to by "probability". Carnap therefore provides two explicanda: probability₁, the degree of a hypothesis's confirmation, and probability₂, the relative frequency of a type of event (Carnap 1945, §3–§4). Whilst both are numerical, in their metric versions the former is semantic (attributing a value determined by logical relations between hypotheses and the evidence for them) and the latter is empirical (the value being determined by relevant observations). In so doing, Carnap has shown how to move beyond the 'futile polemics' of those advocates of either conception who believe they have arrived at the singular, correct analysis of "probability" (Carnap 1945, 532).

philosophical one, a disagreement over not only what a philosophical problem is, but how philosophical problems are solved. Such criticisms assume the perspective of a non-naturalist. So for non-naturalists, including many of Carnap's critics, these arguments are not just compelling, but obvious. But this is preaching to the choir. Similarly, defenders of Carnapian explication include Justus, an advocate of experimental philosophy, and Kitcher, an avowed naturalist, thinkers who consequently see little value in the demand for "real" philosophical answers. At this point, the disagreement over the value of explication as a naturalistic methodology reverts to the argument between naturalism and traditional philosophy.

Chapter 5

Quine, Carnap, and Analyticity



I have argued that Carnap should be considered an epistemological naturalist. For many readers, this will provoke an immediate response: “But Carnap maintains the analytic/synthetic distinction!” Amongst contemporary naturalists, under Quine’s influence, it is assumed that the rejection of the analytic/synthetic distinction is necessary for or equivalent to naturalism. Quine brought naturalised epistemology into the philosophical mainstream, and many accounts treat his route to naturalism as not merely a contingent historical one, but a conceptual one. And Quine’s rejection of the analytic/synthetic distinction is famously developed in opposition to Carnap’s defence of analyticity. The engagement between Quine and Carnap, and their numerous and infamous disagreements, have frequently been interpreted as a process of Quine’s naturalism growing out of a rejection of Carnap’s (supposed) reductionism. This framing however obscures that Carnap ought to be understood as a methodological naturalist. Quine’s intellectual pathway to naturalism is not the only one. Moving beyond the common misunderstanding means recognising that the issues of naturalism and the analytic/synthetic distinction are not coextensive, and that the one does not necessarily determine the other. This allows us to see that a specific conception of analyticity is perfectly consistent with epistemological naturalism, and that just such a conception is maintained by Carnap. Really then, this is a debate between two competing versions of naturalism; one allowing for a conventionalised, semantic notion of analyticity, and another that rejects any attempt to divide analytic from synthetic.

5.1 The Development of Quine’s Thinking About Analyticity

One can describe the years between 1934 and 1951 in Quine’s intellectual development as a process of finding his own voice through criticism of a hero. Whilst in 1934 he could consider himself a disciple of Carnap, by 1951 he had not only

rejected Carnap's understanding of analyticity, but any such notion and Carnap's project (as Quine understood it) more broadly. This disillusionment seems to have occurred gradually; there is no single point at which Quine can be seen to switch from Carnapian to anti-Carnapian. Rather, as time went on the list of disagreements became longer and longer, and the underlying reasons for these disagreements must have started to become apparent to Quine, and perhaps Carnap too.

5.1.1 *Initial Enthusiasm*

Quine's early engagement with Carnap is enthusiastic, even fawning.¹ Quine would later call his 1934 lectures on Carnap 'abjectly sequacious' (Quine 1991, 266). This is despite Quine having voiced some concerns to Carnap over the latter's conception of analyticity in personal conversation in 1933.² In the lectures, Quine defines analytic judgements as 'consequences of definitions, conventions as to the uses of words' (Quine 1934/1990, 47–48). He adopts (what he believes is) Carnap's conventional notion of analyticity. The conventionality is twofold; analytic propositions are true by convention, and which sentences become analytic is a matter of convention too (Quine 1934/1990, 64). At this stage, Quine equates the analytic with logical truths established by definition. Analytic truths gained from these via substitution of synonyms as yet play no role. Quine already notes the arbitrary nature of these definitions. Having established the analyticity of maths and logic, he asks 'why stop here?' (Quine 1934/1990, 61). Why not extend this process to incorporate all terms? He concludes that 'we do best to render only such sentences analytic as we shall be most reluctant to revise' (Quine 1934/1990, 63). But this is pragmatic, not principled. What will later become a deep worry is merely hinted at here.³

Quine's first explicit doubts, raised in "Truth by Convention", are friendly.⁴ Again, Quine notes that there is no principled basis for not expanding analyticity indefinitely: the 'method can even be carried beyond mathematics, into the so-called empirical sciences' (Quine 1936/1966, 93). But now this is more concerning, and becomes the thrust of his argument. As Greg Frost-Arnold says, 'we can read Quine as making a slippery slope argument: once we permit one sentence to be true by linguistic fiat, there is no principled ground for stopping the unlimited inflation of such truths' (Frost-Arnold 2013, 85). Publicly, Quine presents himself

¹ Quine said he was '[Carnap's] disciple for six years' (Quine 1970, xxiii). Seemingly, the period between 1932 and 1938.

² Quine quotes Carnap's unpublished notes from 31st March 1933, discovered by Neil Tennant, but concedes that '[e]vidently Carnap allayed my misgivings for awhile' (Quine 1991, 266).

³ Hylton has similarly argued for tensions in these lectures which pre-empt later disagreements (Hylton 2001), as has Gary Kemp (Kemp 2014).

⁴ Janssen-Lauret calls Quine's reservations 'fairly mild' (Janssen-Lauret 2018, xix). At the very least, Quine's endorsement 'now seems less wholehearted' (Ben-Menahem 2005, 253).

as essentially in agreement with Carnap, but providing some necessary amendments and clarifications. However, there are reasons to suspect that privately, Quine's doubts ran deeper. Specifically, the last line of the penultimate paragraph, where Quine asks 'what one adds to the bare statement that...they are firmly accepted, when he characterizes [statements] as true by convention', provides a first glimpse of the position he will come to in 1951, that it is only the extent of our willingness to revise statements that distinguishes them (Quine 1936/1966, 99).

5.1.2 *Growing Concerns*

This basic attitude of polite (but growing) doubt seems to have persisted throughout the conversations at Harvard between Quine, Carnap and Tarski over the academic year 1940–41. These discussions show Tarski furthering and cementing Quine's worries about the lack of a non-arbitrary dividing line and the indefinite expansion of the analytic. Tarski showed that indefinite expansion in the other direction is also possible; there is just as little stopping us from expanding the domain of the synthetic to incorporate everything, including maths and logic (Frost-Arnold 2013, 86). Quine still understands analyticity as logical truth, although he is beginning to think about synonymy (Carnap, 20/01/1941, in Frost-Arnold 2013, 155). But more clearly, the discussions show the development of Quine's thinking (as hinted at in 1936) that unrevisability is what distinguishes the analytic. Crucially, Frost-Arnold shows, in these discussions Quine makes clear that he understands the question of whether a statement is analytic to be one for empirical, behaviouristic study (Frost-Arnold 2013, 103). The pursuit of a behaviouristic account of the analytic, will occupy him in the coming years.

Following these discussions, attempts to reconcile Quine's increasing concerns are played out in correspondence with Carnap, most importantly over the course of 1943. It is in these letters that Quine first explicitly voices concerns and arguments that will recur prominently in their debates for the next twenty years. Quine emphasises that we ought to be concerned with natural language, rather than formal ones: 'What I have in mind is our actual scientific language, or something approaching it' (Quine, 5/1/1943, in Creath 1990b, 296). Quine also expresses his uneasiness with the artificiality of Carnap's procedure:

Certainly little progress is made toward clarifying the term 'analytic' in any of its pre-existing usage, if in the face of every statement which is not explicitly a logical truth (like 'No woman not married is married') we have to conclude, 'Whether this is analytic or not depends on what constitution system we adopt, and we aren't going in fact to adopt any. (Quine, 5/1/1943, in Creath 1990b, 297)

Quine makes clear his dissatisfaction with a language-relative notion of analyticity, arguing that it cannot capture our pre-existing conception. Evidently, these concerns were long gestating in private before he published them.

Perhaps most importantly, in this period we see Quine's move beyond the narrower notion of logical truth, and towards a definition of analyticity in terms of synonymy. This is expressed most explicitly in "Notes on Existence and Necessity", published in 1943, where Quine says 'we can define an analytic statement as any statement which, by putting synonyms for synonyms, is convertible into an instance of a logical form all of whose instances are true' (Quine 1943, 120). This is a natural development from his thinking in 1941.

It is clearer, I think, to shortcut the question of definitions in connection with the relation between the analytic and the logically true, and to speak directly, rather, of the relation of *synonymy* or sameness of meaning. Given this notion, along with that of logical truth, we can explain analyticity as follows: a statement is analytic if it can be turned into a logical truth by putting synonyms for synonyms (Quine, 5/1/1943, in Creath 1990b, 297)

This necessitates a slight reorientation of the task at hand:

The problem remains, of course, to explain this basic synonymy relation... The definition of this relation of synonymy, within pragmatics, would make reference to criteria of behavioristic psychology and empirical linguistics (Quine, 5/1/1943, in Creath 1990b, 297–298)

Quine remains somewhat hopeful about the prospects for a behaviouristic account of analyticity in terms of synonymy. What is here presented with cautious optimism re-emerges in "Two Dogmas" with the opposite message, that this was the best possible attempt and it failed.

5.1.3 *Quine's Break with Carnap*

Quine's first explicit rejection of the notion of analyticity (in the broad sense, via substitution of synonyms) appeared in 1951 in "Two Dogmas of Empiricism". The first part of the paper argues that the notion of analyticity is unclear, as are the concepts of meaning and synonymy, and that attempts to define analyticity in terms of definitions, semantical rules, or interchangability ultimately rely on prior notions of synonymy, and are therefore equally unclear. Such definitions become circular or pointless. That there is a distinction between analytic and synthetic at all is the first eponymous dogma of empiricism. The second part of the paper argues against what Quine calls reductionism, the thesis that each statement has uniquely confirmable empirical consequences. The thesis of reductionism is the second dogma. This position Quine interprets as a (potential) definition of synonymy. Two statements are synonymous if they have the same method of confirmation. If this is true then we have a notion of synonymy, in terms of which we can define analyticity. It is on this basis that Quine claims the 'two dogmas are, indeed, at root identical' (Quine 1951a/1980, 41). But Quine argues against reductionism on the basis of confirmation holism, and therefore undermines reductionism. With no clear concept of synonymy, the best basis for an account of analyticity fails. From this failure Quine argues, that 'there is a distinction to be drawn at all is an unempirical dogma,

a metaphysical article of faith' (Quine 1951a/1980, 37). The distinction between analytic and synthetic ought to be abandoned. Quine concludes:

it is nonsense, and the source of much nonsense, to speak of a linguistic component and a factual component in the truth of any individual statement. Taken collectively, science has its double dependence upon language and experience; but this duality is not significantly traceable into the statements of science taken one by one (Quine 1951a/1980, 42)

Quine instead gives a model of knowledge as a field of interconnected beliefs. In his new system, those beliefs we are least disposed to revise are those furthest from the periphery.

"Two Dogmas" sees the fruition of three themes that recur throughout Quine's preceding works. The first is revisability. This idea arises as far back as the 1934 lectures, but takes on a new form here. Fixity is now determined to be insufficient as an account of analyticity.⁵ Mathematics and logic are still the most fixed beliefs and lie in the centre of our field of beliefs. But this does not make them immune from revision. There is no domain of uniquely unrevisable statements to be identified as analytic. The second theme is Quine's focus on synonymy. Since at least 1943, Quine had suggested synonymy as the most promising path to a viable account of analyticity. The crucial step in "Two Dogmas" is his recognition that even this best route fails. In this sense, Quine's argument in the paper is as much a reckoning with his own prior proposals as it is with Carnap's. The third key theme is Quine's demand for behaviouristic criteria. Although this demand is not explicitly made in the paper, I agree with Creath that it is this demand for behaviouristic criteria that underlies Quine's claim that the notion of analyticity is unclear (Creath 2004, 49).⁶ Although not explicitly mentioned in the paper, it is in "Two Dogmas" that these behaviouristic concerns finally become a part of Quine's more concerted argumentative strategy. The argument of "Two Dogmas" then, whilst the first explicit rejection of the analytic/synthetic distinction, can (at least retrospectively) be recognised as a natural development of Quine's thinking over the previous fifteen years or so.

In 1954, Quine goes a step further than he did in "Two Dogmas". In "Carnap and Logical Truth", Quine rejects even the possibility of logical truth, conceived of as truths of the meaning of logical constants alone. Quine develops his worries about a separation between the logical and factual components of truth:

consider the logical truth "Everything is self-identical"... We *can* say that it depends for its truth on traits of the language... but we can also say, alternatively, that it depends on an obvious trait, viz., self-identity, of its subject matter (Quine 1954/1966, 107)

⁵ In fact it goes back further than 1934. In Carnap's notes of a meeting with Quine in March 1933, when Quine first visited Carnap in Prague, Carnap describes Quine's view of analytic sentences: 'they are the sentences we want to hold fast' (Quoted in Quine 1991, 266).

⁶ Creath provides an extensive list of cases in which Quine makes these behaviouristic demands both before and after Two Dogmas (Creath 2004, 63). Verhaegh also emphasises the importance of these behavioural criteria for Quine, although he notes that these criteria are only implicit in Quine's 1930s work (Verhaegh 2018, 116–122).

Quine rejects the notion of logical truth by undermining the notion of a sentence whose truth depends solely on the language, on meaning alone, and not at all on the facts. Quine thus rejects what he had lauded Carnap for in 1934. Quine began by endorsing analyticity as logical truth in 1934, then moved to analyticity in term of synonymy in 1943. In “Two Dogmas”, he rejected analyticity in terms of synonymy, and finally here he abandons logical truth too. He has come full circle.

The outright rejection of analyticity, Creath dates to the summer of 1947 (Creath 1990b, 35).⁷ He argues that in correspondence with White and Goodman, Quine’s scepticism hardened into outright denial. These discussions led to the publication of Goodman’s “On Likeness of Meaning” in 1949, White’s “The Analytic and Synthetic” in 1950, and Quine’s “Two Dogmas” in 1951.⁸ Mancosu, however, suggests a much earlier break. Leaning on Quine’s correspondence from 1942, Mancosu suggests Quine had rejected analyticity by 1942 at the latest, and perhaps as early as 1936 (Mancosu 2005, 331). Frost-Arnold provides a timeline that attempts to partially accommodate both Creath’s and Mancosu’s interpretations. On this account, Quine gave up on Carnap’s contemporaneous language-relative explication of analyticity by 1942, but didn’t give up analyticity entirely until 1947 (Frost-Arnold 2013, 115).⁹ Accordingly, from 1943 onwards, Quine abandons Carnap’s semantic, a priori characterization of analyticity, instead attempting a characterization along empirical, behaviouristic lines. I think Frost-Arnold’s combined timeline seems accurate. Quine’s 1943 change of emphasis to an account of analyticity in terms of synonymy (to be understood behaviouristically) not only lines up with this timeline, but can be explained by it; dissatisfaction with Carnap’s contemporaneous account motivates the pursuit of an alternative. I therefore agree that the decisive period, during which Quine’s scepticism of analyticity transforms into hostility and rejection, occurs around 1947. After 1947, we can consider Quine as having both rejected Carnap’s definition of analyticity, and having concluded that no coherent definition is attainable.¹⁰

What the above should highlight is that there is an underlying dissatisfaction throughout Quine’s engagement with Carnap. It initially manifests as a worry over indefinite expansion, and later transforms into a worry that Carnap’s analytic/synthetic distinction lacks epistemological significance. In the rest of this chapter, by considering and evaluating Quine’s arguments against Carnap and defending Carnap from them, I will tease out the conviction underlying this dissatisfaction, as I think it is the basis for the various disputes between Quine and Carnap. Specifically I hope to show how Quine’s misunderstanding of the

⁷ Verhaegh suggests the break comes slightly later, in 1949, caused by a broadening of holism precipitated by Quine’s thinking about nominalism over 1947 and 1948 (Verhaegh 2018, Chapter 6).

⁸ An emphasis on the importance of these discussions is also found in Isaac (2011), Murphey (2012) and Verhaegh (2018).

⁹ Carnap’s conception is outlined in detail below.

¹⁰ That is, until Quine accepts a version of analyticity in 1974, which we will return to below.

purposes of Carnap's logic of science project and the reorientation of philosophy it entails, is itself a result of Quine's misunderstanding of Carnap's empiricism. This misunderstanding of Carnap's goals and motivations facilitates an expectation in Quine from Carnap's work that Carnap's actual project prohibits him from ever providing: the drawing of a principled, non-arbitrary epistemological dividing line between analytic and synthetic statements.

5.2 What Is Carnap's Notion of Analyticity?

To understand Carnap's conception of analyticity, the first key step is to understand his reconceptualization of logic.¹¹ Traditionally, it was taken for granted that there is one true logic. Carnap argued that not only are there multiple logics, as many as we want to make, but there is no sense in which a system of logic is right or wrong. Logic is not a doctrine, but a set of rules, rules we can choose to adopt or reject and which give structure to languages. In choosing a language, we accept a certain set of propositions. These are the assumptions, conventionally accepted, that are definitive of the language system. And 'given a specific conventional selection of basic assumptions and principles of inference, the truth of a great many other claims is thereby decided' (Creath 1991, 350). These claims, those whose truth within the language is antecedently determined by the choice of axioms that determine the language, are the analytic statements of that language. Such truths are true by definition in as much as their truths are a simple consequence of the language system as originally established. For such sentences, 'the truth is independent of the contingency of facts' (Creath 1991, 308). The difference between analytic and synthetic statements in Carnap's epistemology is perfectly captured in Creath's characterisation of them as constitutive and substantive respectively. Synthetic statements 'genuinely have content and describe the world' while analytic statements 'serve instead to structure or constitute the language in which we describe the world' (Creath 2007a, 334).

This reorientation, from seeing logic as a doctrine or set of truths to seeing it as a constitutive rule-system for language, was a decisive intellectual turning point. It leads Carnap to his famous '*Principle of Tolerance: It is not our business to set up prohibitions, but to arrive at conventions*' (Carnap 1934d/2001, 51–52).¹² Carnap spells out the consequences of this principle further:

In logic, there are no morals. Everyone is at liberty to build up his own logic, i.e. his own form of language as he wishes. All that is required of him is that, if he wishes to discuss

¹¹ The interpretation of Carnap's conception of analyticity defended in this chapter was developed in collaboration with Thomas Uebel. It has been defended in (Bentley and Uebel 2023).

¹² Exactly what precipitated this adoption is still under investigation by contemporary scholars. See Carus (2007a, 250–254), Awodey and Carus (2007), Creath (2012), and Leitgeb and Carus (2020).

it, he must state his methods clearly, and give syntactical rules instead of philosophical arguments (Carnap 1934d/2001, 52)

We can no longer talk about the correct logic, or the correct language. ‘There is no question of right or wrong, but only a practical question of convenience or inconvenience... of its suitability for certain purposes’ (Carnap 1936–37, 4). The choice between languages, or logics, is always made on practical grounds. No logic is more accurate, no language truer than any other. The choice is always made according to utility.

Carnap’s conception of logic is therefore fundamentally conventionalist. The logical postulates on which a language is established, what Carnap calls a language’s L-rules, are adopted on the basis of pragmatic decision, by convention. Analytic statements, those true by the rules of the language, are analytic by convention.¹³ But, this is not the same as simply declaring analytic truths to be true by stipulation. In choosing a set of axioms, we are essentially defining the terms involved. The process of selecting axioms is in effect a process of implicit definition. ‘These conventions, namely, the rules of formation... are, however, not arbitrary. The choice of them is influenced, in the first place, by certain practical methodological considerations’ (Carnap 1934d/2001, 320).¹⁴ These sentences, whilst true by convention in one sense, also serve to establish the meanings of the terms involved: ‘we have so chosen our language as to make these particular claims true’ (Creath 1991, 6). Analytic truths beyond the axioms, are those that in virtue of their meaning, are made true by the adoption of these axioms. So, analytic statements are both true by convention, and in virtue of the meanings of the terms alone. (As we will see, both “made true” and “in virtue of” are in need of further analysis.)

This is not to say that there cannot be statements of the framework adopted simply by convention. There are certain non-analytic statements that we may simply choose to include within the framework principles, but this does not make them true by convention. Carnap introduces so-called P-rules as ‘*extra-logical rules of transformation*’ (Carnap 1934d/2001, 180). He allows that we may sometimes wish to adopt a language system with such non-logical transformation rules. The choice whether or not to include P-rules in a language ‘is not logico-philosophical problem, but a matter of convention and hence, at most, expedience’ (Carnap 1934d/2001, 180). The obvious candidates are so-called laws of nature, universal empirical statements. Carnap considers that it may prove useful for the purposes of analysing particular scientific theories to elevate certain universal statements to framework principles. The inclusion of such P-rules within a language’s framework may prove useful for understanding how these theories function.¹⁵ But the P-rules

¹³ Carnap sometimes called the explicated notion of analyticity ‘L-truth’, referring to analytic statements as being ‘L-true’ (Carnap, 1947/1988, 10). But he also continued to talk in terms of the pre-systematic notion “analytic”, as I will in this chapter.

¹⁴ To avoid the connotation of arbitrariness, Carnap later preferred the characterisation ‘*truth based on meanings*’ (Carnap 1963a, 916).

¹⁵ An example of this kind of P-rule choice is explored below, in Sect. 5.3.7.

are not true in virtue of meaning, and they do not play a constitutive role in structuring a language-system. Analytic truths therefore cannot simply be equated with conventionally adopted statements of the framework.

Two features of Carnap's conception of analyticity are particularly noteworthy. Firstly, because of the relationship between language-formation and analyticity, each language has its own corresponding notion of analyticity:

That a certain sentence S is analytic in L_n means only something about the status of S within the language L_n ; as has often been said, it means that the truth of S in L_n is based on the meanings in L_n of the terms occurring in S . (Carnap 1963a, 921).

There is, therefore, no general language-independent definition of analyticity. A truth can only ever be analytic-in- X , never analytic simpliciter. Analyticity is a consequence of the meaning of words, but since meanings can vary between languages, so can what is true in virtue of them. One further upshot of this is that 'the concept of analyticity has an exact definition only in the case of a language system, namely a system of semantical rules, not in the case of an ordinary language' (Carnap 1952b/1990, 427). Because the analytic truths of a language are given as the consequences of a directly specified set of language-defining axioms, in the case of natural languages which lack these exactly specified axioms, the concept is not obviously applicable. As Leitgeb and Carus highlight, the exact details of how analyticity is defined vary between different linguistic frameworks, being dependent upon the purposes for which the framework is intended (Leitgeb and Carus 2020, Supplement B). But crucially, despite these variations in detail, the functional role of analyticity remains essentially the same:

to keep track of those statements in a linguistic framework that "come with" the framework, that express the framework's own resources, that are presupposed in the framework, and which do not have any empirical content on their own (Leitgeb and Carus 2020, Supplement B)¹⁶

The second important feature of Carnap's notion of analyticity is that it must not be understood as a definition or conceptual analysis of the pre-existing notion of analyticity. Rather, it exemplifies Carnap's method of explication. As discussed in the previous chapter, the important difference between analysis and explication is that where an analysis seeks to find truth, the correct understanding of a term, an explication is a proposal for a new, albeit very similar, term. Relative to the purpose for which the explication is initially intended, success is to be judged according to how closely it captures the usage and spirit of the explicandum. In this case, Carnap informally characterises the explicandum "analyticity" as '*truth based on meanings*' or, those statements for which 'to understand them is a sufficient basis for the determination of their truth' (Carnap 1963a, 916). These key features of the

¹⁶ They compare this to Friedman's (2001) interpretation of Carnapian analyticity as an explication of '*relative or constitutive apriority: the relational property of sentences being presupposed by, and enabling, the formulation, truth-conditionality and empirical justification of scientific hypotheses belonging to a joint scientific paradigm*' (Leitgeb and Carus 2020, Supplement B).

explicandum are used as the basis for a criterion of adequacy for explications of analyticity:

A sentence S_i is **L-true** [analytic] in a semantical system S if and only if S_i is true in S in such a way that its truth can be established on the basis of the semantical rules of the system S alone, without any reference to (extra-linguistic) facts.

This is not yet a definition of L-truth. It is an informal formulation of a condition which any proposed definition of L-truth must fulfil in order to be adequate as an explication for our explicandum. Thus this convention has merely an explanatory and heuristic function. (Carnap 1956b/1988, 10)

The epistemological features associated with the traditional notion of analyticity are a key component of its traditional usage, and therefore important for judging the success of an explication. But the way in which the analytic/synthetic distinction itself is drawn is semantic, not epistemological.

5.3 Which of Quine's Arguments Pose a Problem For Carnap?

Whilst it is clear that Carnap and Quine have significant disagreements about analyticity, it is possible for one person's worries to be unproblematic from within another person's project. For current purposes, the question is whether Quine's worries should (not whether they actually did) worry Carnap. If we adopt Carnap's position, can we answer Quine's worries? I think the answer is yes, and that in doing so, a more fundamental, foundational disagreement between Quine and Carnap is illuminated. The obvious caveat here is the "if". These defences of Carnap from Quine are largely premised on the acceptance of Carnap's goals, and more broadly, the project of the logic of science. In particular, as Carnap repeatedly emphasises in his responses to Quine, the methodology of explication and the principle of tolerance play central roles in the answer to many of Quine's worries. These are defences that are internal to the Carnapian project. All of the arguments below falter if we reject Carnap's project from a starting point entirely different to his. But that is a different topic altogether.

5.3.1 "Analyticity" Is Unclear

Quine's most rudimentary objection to the notion of analyticity is that it 'is not even moderately clear in anybody's mind' (Quine 1946a/2008, 24). The first half of "Two Dogmas" is devoted to showing that this unclarity cannot be overcome. As we know, Carnap's solution to this lack of clarity is explication. But Quine is insistent that this

approach fails in the case of analyticity.¹⁷ As we saw, Carnap's explications gave us language-relative notions of analyticity. But what does a definition of analytic-in-L actually tell us? Quine says there are two options. One interpretation is that the new definition being given still includes the unclear concept of analyticity. In this case, 'before we can understand a rule which begins "A statement *S* is analytic for language *L*₀ if and only if ...", we must understand the general relative term "analytic for"' (Quine 1951a/1980, 33). Alternatively, we should think of it as a definition for a new symbol, in which case it would be better to give the notion a new name, 'so as not to seem to throw light on the interesting word "analytic"' (Quine 1951a/1980, 33). In the first case, this new definition leaves the original notion exactly as unclear as before, and consequently so is the notion of analytic-in-L. In the second case, the new notion is clear, but irrelevant to understanding the earlier notion of analyticity. The only link between the two is the shared name-symbol "analytic".

The most obvious response to Quine is that he is exaggerating. The explicandum "analytic" is unclear, but nowhere near as unclear as Quine pretends. In fact, Quine's argument in the first half of "Two Dogmas" suggests quite the opposite. If he has no idea what analytic means, then it is a remarkable coincidence that his attempts to provide a non-circular definition come so tantalisingly close to the shared intuition he is apparently unaware of. As was famously argued by Grice and Strawson, there is a longstanding, established use of "analytic" in philosophy. There is significant agreement within the philosophical community over which cases are analytic, which are not, and which cases are the most difficult or ambiguous, a consistency which extends not only to familiar cases but to new ones as well (Grice and Strawson 1956, 142–3). This is insufficient to prove that the concept in our everyday language is perfect, but it does demonstrate a shared intuition in need of explication, and that is all Carnap needs.

There is a minimal sense in which Quine must accept the word "analytic" is meaningful for him to even talk about it. Borrowing a term from Gillian Russell, call this meaning-lite (Russell 2014, 186–7). Quine's counter argument would be that, whatever the arguments supplied by the likes of Grice and Strawson, all this demonstrates is meaning-lite. But meaning-lite in this absolutely minimal sense is by no means a guarantee of conceptual legitimacy. Harman makes this Quinean point: 'philosophical talk about analyticity no more shows there is such a thing as analyticity than the former prevalence of talk about witches shows there were once witches' (Harman 1967, 125). This is a legitimate argument. The fact we have a word for something is no guarantee of its existence. The history of science is full of substances like aether and phlogiston that we now consider non-existent. And I

¹⁷ Quine does not outright reject explication as a methodology. In fact, he embraces a version of explication. But his explication is not Carnap's. For one thing, Quine's explication is eliminative (Quine 1960/2013, 239). For more on the difference between Quine and Carnap's conceptions of explication, see Gustafsson (2014).

think Carnap would have agreed. As we will see below, he does eventually make certain concessions to Quine's demands for clarification of the explicandum.

Quine's repeated insistence that the notion of analyticity is unclear can sometimes appear as stubbornness, but there is a principled reason behind it. The motivation for this demand is an essentially empiricist one. Quine, is a strict empiricist.¹⁸ As far as he is concerned, any notion that lacks empirical content is meaningless.¹⁹ Quine's worry or accusation is that the distinction between analytic and synthetic is empirically empty (Creath 2004, 49). Hence why he refers to it as 'an unempirical dogma, a metaphysical article of faith' (Quine 1951a/1980, 37). And although this claim is made most obviously in 1951, the demand for empirical criteria is in evidence in Quine's desire to define analyticity in terms of synonymy, and his conviction that only a behaviouristic account can supply a sufficient definition of synonymy. It is also a key topic in the correspondence:

It is only by having some general, pragmatically grounded, essentially behavioristic explanation of what it means in general to say that a given sound- or script-pattern is analytic for a given individual, that we can understand what is intended when you tell us (via semantical rules, say) "the following are to be analytic in my new language" (Quine, 10/5/1943, in Creath 1990b, 338)

We can now see exactly why Quine was dissatisfied by Carnap's explicatory proposals. Quine thinks that whether or not a statement is analytic is determined behaviouristically, by empirical study of if and when people would be willing to give it up. At the very least, analyticity as a concept must allow of behaviouristic characterisation. And without clear specification of the behaviouristic conditions under which we would judge something analytic, then the concept remains unempirical and metaphysical.

Given that Carnap's notion of analyticity is a proposed explication, Quine's demands can potentially be understood as applying to either (or both) the explicandum and the explicatum. As applied to the explicandum, Quine's argument amounts to is this: explication of a concept can only be successful if we have sufficient behaviouristic criteria for the initial concept, otherwise it is unclear what the explication is of and what purpose the new explicated concept is supposed to provide. This seems a legitimate demand. Now obviously we cannot be expected to provide criteria to show the notion to be perfectly clear. Perfectly clear concepts don't require explication. But some criteria would do the work of demonstrating to Quine that the initial notion is not unintelligible. As Creath highlights, Quine never attempts to argue that such criteria cannot be given (Creath 2004, 52).²⁰ Rather, it is incumbent upon those who do take the notion of analyticity to be intelligible to demonstrate this through the provision of appropriate criteria.

¹⁸ Creath suggests he is employing an empiricist criterion of meaning here (Creath 2007a, 328).

¹⁹ Whether such a demand is legitimate is dubious, given that analyticity is a meta-linguistic notion, and therefore should not be held to the same criteria as the object language.

²⁰ Although a more cynical reader may argue that he does seem to heavily imply it.

Carnap interprets Quine along these lines:²¹

As I now understand Quine, I would agree with his basic idea, namely, that a pragmatistical concept, based upon an empirical criterion, might serve as an explicandum for a purely semantical reconstruction, and that this procedure may sometimes, and perhaps also in the present case, be a useful way of specifying the explicandum. (Carnap 1963a, 919)

So, does or can Carnap provide successful behaviouristic criteria for the pre-existing notion of analyticity? Well, as Creath notes, Quine himself does the work for Carnap (Creath 2004, 59). In his later work, having embraced a version of analyticity, Quine gives a behaviouristic characterisation of analytic statements as those we learn the truth of in coming to understand the terms of the statement (Quine 1974, 78–80). This is a fair characterisation of the explicandum. And from this basis, Carnap is free to make whatever explicatory proposals he wants. Carnap does also provide his own account of how to provide behavioural criteria for meaning specification of natural language terms Carnap (1955a). If Quine's demand is merely that the explicandum should allow of behaviouristic characterisation, then this demand can be met.

Carnap makes clear however that, although such behaviouristic criteria can be useful for clarificatory purposes, they cannot be required in principle of every explicatum, in particular not of the explicatum analyticity:

A philosophical thesis on logic or language, in contrast to a psychological or linguistic thesis, is not intended to assert anything about the speaking or thinking habits of the majority of people, but rather something about possible kinds of meanings and the relations between these meanings. In other words, a philosophical thesis does not talk about the haphazard features of natural languages, but about meaning relations (Carnap 1963h, 1003)

But Quine appears to have demanded that Carnap's explicated notion of analyticity should also display such criteria. What Quine demands of Carnap is that he provide behavioural criteria not only for the explicandum, but also the explicatum. But for Carnap, this demand is simply illegitimate. The role of the analytic statements of a language is structural, meaning-constitutive. They are not descriptive, they do not have empirical content, and it is empirical content that Quine's demands for behavioural criteria require. Richardson for this reason calls Quine's demand a 'philosophical sleight of hand' (Richardson 2003, 7).

For our purposes, the important sense of this question is whether, having accepted the distinction between substantive and constitutive, we should grant that some behavioural criteria must be applicable to the constitutive components of a language system. Carnap would argue that it is not. Unsurprisingly, Quine does not recognise a distinction between the substantive and constitutive components of a language. As Creath notes, his demand is indicative of a fundamental difference between Carnap and Quine as to whether or not the statements of logic have content (Creath 2007a, 330). The possibility of Quine's demands here being met rest upon starting assumptions that are incompatible with Carnap's. The question therefore returns us to those starting points. But, as we will see, the defence of Carnap is not restricted simply to stating such incompatibilities.

²¹ See also Carnap (1955a, 35).

5.3.2 *Natural vs Formal Languages*

Another point of constant disagreement between Quine and Carnap was which languages are relevant to discussions of analyticity. For Carnap:

I believe that we cannot construct an exact and workable theory of concepts like “true”, “analytic”, “meaning”, “synonymous”, “compatible” etc. if we refer merely to the actually used language of science. It seems to me that we can use those concepts only if we replace the given language by a system of rules (Carnap, 21/1/1943, in Creath 1990b, 309)

For Carnap, these concepts as used in natural languages lack the clarity required, because natural language itself lacks sufficient clarity. Consequently, Carnap deals almost exclusively with formal, exactly specified language-systems, in which terms can be given precise definitions for purposes of explication. The construction and study of such logically precise language-systems is achieved in the domain of semantics. Carnap distinguished very sharply between the semiotic domains of semantics and pragmatics (more on which in Sect. 5.3.8). Unlike the former, only the latter is concerned with the study of the meaning of natural language terms in the wild and their given logical imperfections. Carnap’s notion of analyticity is a semantic notion.

As we saw above, Quine did not agree to this strategy, insisting on the relevance of natural rather than artificial languages. This is one case in which Quine and Carnap seem to have recognised the nature and scale of their disagreement. They both advocate different projects for the philosopher, and to try and demonstrate the value of one project to an advocate of another is an argument that cannot be won. And both Quine and Carnap agreed. According to Stein’s testimony, the two agreed that the only real means of vindicating one project over the other was to pursue them, and see which proved more successful. He recalls Carnap saying:

In my view, both programs - mine of formalized languages, Quine’s of a more free flowing and casual use of language - ought to be pursued; and I think that if Quine and I could live, say, for two hundred years, it would be possible at the end of that time for us to agree on which of the two programmes had proved more successful (Stein 1992, 279)

I think this agreement is the best solution to the debate. As we will see later, the difference of emphasis on the part of either Quine or Carnap is a result of a fundamental difference in the project being pursued. When philosophical disputes are this foundational, there seems to be no argumentative possibility for one side prevailing over the other.

5.3.3 *Quine’s Observation*

As mentioned above, in “Truth By Convention” Quine takes logical constants to have a specific meaning in virtue of our stipulating the truth of certain sentences containing them. But there are an infinite number of such sentences for each constant, so some general convention must be adopted according to which all

sentences of a certain type are valid. However, to establish the truth of any specific statement on the basis of the general convention requires a logical inference from the general to the specific: 'if logic is to proceed *mediately* from conventions, logic is needed for inferring logic from the conventions' (Quine 1936/1966, 97). And this will have to make use of constants like "all" and "every". This leads to an infinite regress. Quine compares this paradoxical conclusion to Lewis Carroll's in the paper "What the tortoise said to Achilles" (Quine 1936/1966, fn. 21). Following Ebbs, I will call this regress puzzle 'Quine's observation' (Ebbs 2011, 195). This has often been interpreted as posing a problem (frequently an insurmountable one) for Carnap's conception of logical truth:

Carnap, like other logical empiricists, held that the adoption of a system of logic is fundamentally a matter of linguistic convention, so that logical truth is "truth by convention"... Quine observed, however that the logical implications of a logical truth cannot themselves be a matter linguistic convention, on pain of requiring an infinite number of such conventions, and thus that the role of convention in logic can amount at most to the adoption of certain fundamental principles (Baldwin, Quoted in Ebbs 2011, 193–194)²²

One type of response to Quine's observation is provided by Paul Boghossian.²³ Boghossian intends to defend a Carnapian conventionalism about logic. He argues from a distinction between explicit, deliberate adherence to rules, and behaviour in accordance with rules identified in terms of behavioural dispositions (Boghossian 1996, 381). In order to follow a rule *R*, one need not explicitly state that rule. The regress can therefore be circumvented. This solution is in the tradition of modern responses to Carroll's original paradox.²⁴ Pavese for instance has recently argued that Carroll's paradox shows that 'giving an argument by a rule is possible by following that rule', and that following a rule is a matter of pre-supposing certain entailment relations (Pavese 2020, 34). A solution along these lines is actually suggested by Quine himself in the original paper:

It may still be held that the conventions (I)–(VIII), etc., are observed from the start, and that logic and mathematics thereby become conventional. It may be held that we can adopt conventions through behaviour, without first announcing them in words... So conceived, the conventions no longer involve us in a vicious regress (Quine 1936/1966, 98)

But Creath notes that this argument of Quine's parallels one of Quine's from 1934 in defence of Carnap, and argues that it cannot therefore be construed as levelled against Carnap (Creath 1987a, 493). He thus offers a more fundamental rejection of the supposed challenge posed by Quine's observation: Quine's observation is not intended as a problem for Carnap's conception of analyticity. Where Boghossian attempts a response to Quine, Creath's interpretation suggests Carnap doesn't need

²² For further examples of this interpretation of "Truth by Convention" and the challenge it poses Carnap, see Ebbs (2011, fn. 1).

²³ This has been met by an explicitly Quinean response from Gilbert Harman. However, their debate is too technical and granular to detail here. See Harman (1996); Boghossian (2003).

²⁴ Carroll's original puzzle concerns how it is possible to make use of a logical rule, specifically *modus ponens*, in an argument.

to respond.²⁵ This same basic position has been defended more recently, albeit from a different angle (and in scrupulous detail), by Gary Ebbs. Ebbs also argues that Quine never intended his regress observation as a problem for Carnap.²⁶ What Ebbs also demonstrates is that Quine's observation only proves problematic for a metaphysical conception of analyticity, one according to which conventional stipulation *makes* things true. But as we will see below, Carnap maintains no such conception. Questions of truth-making are simply not within the Carnapian remit. Ebbs rightly recognises that Carnap's notion of analyticity is provided with scientific and not traditionally philosophical purposes in mind:

The point of [Carnap's] account is not to explain what "makes" logical truths true, or to show that our knowledge of logical truths can be based on linguistic conventions alone, without relying on logical truths and rules of inference in a metalanguage in which we specify those rules, but to provide inquirers with metalinguistic resources they can use to articulate language systems that facilitate their inquiries (Ebbs 2011, 214–215)

Quine's observation therefore simply is not, and was not intended to be, problematic for analyticity as Carnap understood it.

5.3.4 *Factual vs Non-factual Truth*

Another important criticism of analyticity, one that occurs in the work of both Quine and more recent Quineans, concerns a supposed difference in the kinds of truth captured by analytic and synthetic statements. According to this argument, the analytic/synthetic distinction relies on or amounts to an unsustainable split between fact-independent and fact-dependent truths respectively. This division, Quine argues, is false because 'in general the truth of statements does obviously depend both upon language and upon extralinguistic fact' (Quine 1951a/1980, 41). The problem posed for the defender of the analytic/synthetic distinction is how to distinguish between truths that are factual and truths that are not. For example, statements of the form " $a=a$ " are often considered as obviously true in virtue of their meaning. But Quine asks why they are not considered true in virtue of the fact of self-identity (Quine 1954/1966, 106). Essentially the same argument is repeated by Harman (Harman 1967, 129).

In replying to this Quinean argument, it is useful to utilise a distinction drawn by Paul Boghossian, who distinguishes two different interpretations of analyticity. The first he calls the epistemological notion; 'a statement is "true by virtue of its meaning" provided that grasp of its meaning alone suffices for justified belief in

²⁵ This does not however invalidate Boghossian's analysis of Quine's reading of Carnap. That Quine attributes a metaphysical notion of analyticity to Carnap remains a plausible interpretation, as we will see presently.

²⁶ His argument is convincing, but too long and involved to detail here. See also Ebbs (2014, 2017) and Leitgeb and Carus (2020).

its truth' (Boghossian 1996, 363). This understanding is overtly epistemological because it concerns how we come to know or be justified in certain beliefs. The other notion of analyticity Boghossian calls the metaphysical, according to which "true in virtue of meaning" means 'it *owes its truth value completely to its meaning*, and not at all to "the facts"' (Boghossian 1996, 363). Here, the conception is metaphysical because it concerns that in virtue of which a statement is made true, rather than how we come to know it as such. The distinction is subtle but important, and historically the two positions are often conflated. For example, Harman gives the two conceptions consecutively, as if equivalent or only trivially different: 'The meaning of such a sentence guarantees its truth; knowledge of meaning is enough for knowledge of truth' (Harman 1967, 131). According to Boghossian, these two claims are drastically different. The former suggests the controversial notion of truth-makers (which suggests a correspondence theory of truth). The latter does not, and as such is preferable. It merely makes the claim that some sentences are such that the rules of language let us know they are true (or justify our belief in them); this is a much more modest claim. And yet the two are often treated as equivalent. They are not, and the epistemological notion does not entail or require the metaphysical notion.

What is particularly valuable in Boghossian's analysis is that he highlights how Quine criticises and attributes to Carnap the metaphysical notion. Take Quine's wording on the first page of "Two Dogmas":

Leibniz spoke of the truths of reason as true in all possible worlds. Picturesqueness aside, this is to say that the truths of reason are those which could not possibly be false. In the same vein we hear analytic statements defined as statements whose denials are self-contradictory (Quine 1951a/1980, 20)

The "in the same vein" is an equivocation here. The latter statement is compatible with a much weaker, conception of analyticity such as Boghossian's epistemological notion or Carnap's semantic notion. The Leibnizean formulation is considerably stronger, and necessarily metaphysical. Here, Quine blurs the two together.

Elsewhere, Quine associates Carnap's conception of analyticity with Frege's: 'Carnap held, and Frege before him, that the laws of logic held by virtue purely of language: by virtue of the meanings of the logical words' (Quine 1974, 78). For the sake of argument, we can agree with Quine that the two's views are similar in some important ways.²⁷ But they are also clearly not identical. For Frege logic was a substantive subject, charting the most general, universal, abstract truths about reality. He saw the principles of logic as 'boundary stones fixed in an eternal foundation that our thinking can overflow but never displace' (Frege, Quoted in Ricketts (2007), 203). But it is obvious that such a conception of logic

²⁷ The specific similarities and differences between Carnap and Frege's respective conceptions of analyticity is beyond the current scope.

is fundamentally incompatible with Carnap's principle of tolerance.²⁸ This is a fundamental divergence in their conceptions of logic. Again Quine, as with Leibniz above, associates Carnap's thinking with a far more robust conception of logic. This association again suggests a metaphysical reading.

Most importantly, Quine's phrasing suggests that Carnap's notion of analyticity is concerned with what *makes* logical statements true. This is made more explicit elsewhere:

What are we trying to get at when we call a sentence analytic, or true purely by virtue of the language? When we ask this question, our focus changes. Our focus shifts to the phrase 'true by virtue of'. How, given certain circumstances and a certain true sentence, might we hope to show that the sentence was true by virtue of those circumstances? If we could show that the sentence was logically implied by sentences describing those circumstances, could more be asked? But any sentence logically implies the logical truths. Trivially, then, the logical truths are true by virtue of any circumstances you care to name - language, the world, anything. (Quine 1986, 96)²⁹

In this passage, Quine appears to be making a distinction in terms of truth-making: in terms of what it is that generates the truth of a statement. Whether he means to appeal to truth-maker theory (which he opposes anyway) is perhaps arguable, but such a reading is natural for the metaphysical conception of analyticity that Quine evidently addresses here. The division drawn by Quine in this passage is between true-because-of-reality and true-because-of-language. It is not a distinction between knowable-via-language and knowable-via-facts, or a distinction between the constitutive and descriptive components of a language. Quine describes the metaphysical conception of analyticity. And Quine assumes it is this metaphysical notion that Carnap attempts to capture.³⁰

²⁸ As Ricketts suggests it is natural to imagine Frege, Carnap's logic teacher, as the embodiment of the orthodox conception of logic in the back of Carnap's mind against which the principle of tolerance is asserted (Ricketts 2007, 204–205). For more on Carnap's relationship to Frege, see Ricketts (2004) and Gabriel (2007).

²⁹ Although Quine is not talking specifically about Carnap in this passage, this is the same issue he raises for Carnap in "Two Dogmas", and the passage quoted is followed by an argument that parallels an argument from "Carnap and Logical Truth". Compare (Quine 1951a/1980, 42) with (Quine 1986, 96), and (Quine 1954/1966, 114) with (Quine 1986, 99). I think this is sufficient to show that it is in this same sense that Quine's argument against Carnap must be understood.

³⁰ Harman argues that a version of this argument applies whether analyticity is understood in terms of truth by virtue of meaning (metaphysical) or knowability in virtue of meaning (epistemological) (Harman 1967, 128–129). On the epistemological conception, how is it determined which knowledge is about the world, and which is about meaning? 'Why should it not count as knowing something about the world to know that copper is a metal or that copper is copper?' (Harman 1967, 129). Harman assumes the response will be to invoke conventionalism, and then deploys the supposed problem of Quine's observation, covered in the previous section. But this framing relies on an odd characterisation of epistemological conceptions of analyticity. They do not distinguish analytic and synthetic statements in terms of what they are about, but how we come to know them. This objection thus simply seems irrelevant. For current purposes, whether the epistemological conception of analyticity is defensible against Harman is academic since, unlike Boghossian, I attribute a semantic not epistemological conception of analyticity to Carnap.

Quine consistently argues as if Carnap were committed to this metaphysical notion, of analytic truths as objective truths of reason. This is simply not what Carnap meant. In fairness to Quine however, we must note that in *Meaning and Necessity*, Carnap says his notion of L-truth is intended as an explication of both Leibniz's necessary truth and Kant's analytic truth (Carnap 1956b/1988, 8). Quine's interpretation is not therefore groundless. However, what Quine underestimates is the flexibility Carnap gives his explications: the necessity of analytic propositions on his explication is language-relative, and not therefore directly equated to Leibniz's truth in all possible worlds. They are necessary-in-a-language, in as much as they cannot be false within that specific system. Carnap's explication intends to capture the spirit but not the specifics of Kant and Leibniz's conceptions. It seems particularly misplaced to describe Carnap's L-rules, as described above, as "truths of reason" at all.

It also has to be admitted that Carnap does sometimes use phrasing like "in virtue of" and "dependent upon" regarding analyticity. However, to read even these usages as adopting the metaphysical conception requires a misunderstanding of Carnap's entire approach to metaphysics. As early as the *Aufbau*, Carnap dismisses issues of this sort as involving a metaphysical notion of reality, which has no place in his construction system.³¹ Contra Boghossian however, this should not lead us to the conclusion that Carnap's conception of analyticity is an epistemological one. As we noted above, and is explained in more detail below, Carnap's conception of "analyticity" is semantic. His "in virtue of" talk should not be treated as definitive, being as it is an informal characterisation of the pre-systematic explicandum ("analyticity"), not the final explicatum (L-rules).

A metaphysical conception of analyticity is fundamentally incompatible with Carnap's general, and long-standing, anti-metaphysical outlook and approach.³² Carnap is not concerned with what *makes* sentences true.³³ Carnap's framing of the issue in his reply to "Two Dogmas" makes this clear in the first paragraph (Carnap 1952a/1988, 222):

Philosophers have often distinguished two kinds of truth: the truth of some statements is logical, necessary, based upon meaning, while that of other statements is empirical, contingent, dependent upon the facts of the world. (Carnap 1952a/1988, 222)

This may sound like the metaphysical conception. But Carnap immediately goes on to say that for truths of this first kind, 'it is sufficient to understand the statement in order to establish its truth; knowledge of (extra-linguistic) facts is not involved' (Carnap 1952a/1988, 222). The crucial sense in which synthetic statements are

³¹ See (Carnap 1928a/2005, §175–§178).

³² Friedman argues that Carnap's rejection of metaphysics pre-dates, and is therefore intellectually more fundamental, than his Vienna-era verificationism (Friedman 2007, §IV). In his *Intellectual Autobiography*, Carnap says that 'Even in the pre-Vienna period, most of the controversies in traditional metaphysics appeared to me sterile and useless' (Carnap 1963a, 44).

³³ This will be made even more apparent below in §6.5.1, where Carnap's deflationary approach to truth is discussed. Such truth-making talk involves a substantially more robust conception of truth than Carnap maintains.

truth-dependent is in terms of how we come to know this truth, not in the source or nature of the truth.

Even more important to bear in mind here is that, in these discussions of analyticity, Carnap is typically discussing the pre-existing notion of analyticity that he intends to provide an explication of. Every description of analyticity he gives cannot therefore be taken as exactly characterising his own conception. What can be taken as definitive is his informal characterisation of the explicandum “analytic”, and the criterion of adequacy for explicatum derived from it. And this concerns how we come to know these statements: ‘true in *S* in such a way that its truth can be established on the basis of the semantical rules’ (Carnap 1956b/1988, 10). The criteria of adequacy for his semantic conception is cashed out in epistemological terms, not metaphysical ones. And it could hardly be otherwise. Carnap’s criterion is derived from observations of how the natural language term “analytic” has typically been deployed by speakers. How one comes to know certain claims can be seen in linguistic behaviour. What makes a statement true seems beyond the scope of empirical investigation.

Attributing a metaphysical conception, even to the explicatum “analytic”, relies on a false (if not entirely unjustified) reading of Carnap. It is an even more significant error however to transpose this metaphysical conception onto Carnap’s fully explicated conception of analyticity as L-rules. The distinction between true-because-of-reality and true-because-of-language simply is not an accurate characterisation of the distinction as Carnap draws it. No matter how objectionable such a division is, it does not trouble Carnap.

5.3.5 *Holism, Revisability and Analyticity*

The second of the two dogmas of empiricism that Quine diagnoses and rejects, he calls reductionism. He defines it as the belief ‘[that to] each synthetic statement, there is associated a unique range of possible sensory events’ (Quine 1951a/1980, 40). This second dogma is related to the first. In fact, ‘[t]he two dogmas are, indeed, at root identical’ (Quine 1951a/1980, 41). After all, analytic statements are true in virtue of their meaning alone, and therefore seem capable of confirmation in isolation. In other words, they possess discrete meaning at the level of sentences. Quine argues that reductionism therefore gives credence to the idea of a sentence confirmed no matter what. Hylton interprets Quine’s argument in “Two Dogmas” as attempting to undermine reductionism on the basis that it is reductionism which gives the notion of analyticity its intuitive appeal (Hylton 2002, 13).³⁴ But, if Quine is right, holism is incompatible with analyticity, since reductionism takes the unit of empirical significance to be the sentence, when for the holist it is groups of sentences. This would make Carnap’s position untenable, since he maintains both holism and analyticity. Are these two positions contradictory?

³⁴ As Hylton also notes, this is obviously not true of Carnap (Hylton 2002, n.7).

Firstly, we must be careful to differentiate two interrelated but distinct forms of holism, both of which Quine endorses, one straightforwardly and the other under a particular interpretation. Meaning holism is the position that understanding a sentence is only possible through an understanding of the sentence's interrelations to a larger body of statements. Consequently, a sentence or term has no meaning in isolation. However, given Quine's extensionalist rejection of meanings, for Quine this amounts to no sentence or term being significant (i.e. meaningful) in isolation. Confirmation holism is the position that it is only in groups that sentences can be tested by experience, because sensory experiences do not uniquely correspond to individual sentences. Consequently, 'statements about the external world face the tribunal of sense experience not individually but only as a corporate body' (Quine 1951a/1980, 41). In "Two Dogmas", Quine takes this corporate body to be whole theories. In his later work however, he allows a mitigated form of holism, in which the unit is sufficiently large sets of sentences (Quine 1981b, 71). The two forms of holism can be complementary, but are separable. Although Quine typically refers simply to "holism", at issue here, as the crucial component of the argument of "Two Dogmas", is the thesis of confirmation holism.

As we saw in Chap. 1, Carnap accepted and embraced confirmation holism long before "Two Dogmas": 'it is, in general, impossible to test even a single hypothetical sentence. . . the test applies, at bottom, not to a single hypothesis but to the whole system of physics as a system of hypotheses' (Carnap 1934c, 318).³⁵ Carnap therefore explicitly disavows Quine's second dogma of empiricism.³⁶ But one may question whether he is in a position to do so. Is confirmation holism compatible with analyticity? For many Quineans, the answer is no. But I think there are good reasons to think that Carnap maintained both holism and analyticity simultaneously not because he had failed to recognise them as contradictory, but because as he understood them, they were compatible. After all, Carnap had been exposed to confirmation holism long before "Two Dogmas", both from reading French conventionalists like Duhem, and more directly from Neurath. It would be rash to simply dismiss his position as based on an oversight. And tellingly, Quine's arguments never shake Carnap's belief that his position is consistent:

With all this [pages 42–46 of "Two Dogmas"] I am entirely in agreement. But I cannot follow Quine when he infers from this fact that it becomes folly to seek a boundary between synthetic and analytic statements. (Carnap 1963a, 921)

Given this, I think a very plausible argument for Carnap's position and against Quine can be reconstructed on his behalf.

Like Neurath (as described in Chap. 2), Quine recognises that if experience is not uniquely designative of specific beliefs, and therefore only bodies of statements are confirmed or disconfirmed, then experience underdetermines the distribution of

³⁵ Carnap does not seem to have embraced semantic holism. This however is an issue beyond the current scope.

³⁶ Leitgeb and Carus note that Quine's criticism of reductionism is applicable to the *Aufbau* (Leitgeb and Carus 2020, §7).

truth-values within our web of beliefs. It is therefore a matter of choice as to which beliefs are revised, and how, in the face of recalcitrant experience. This extends even to those seemingly certain beliefs, maths and logic. Consequently, no statement is immune from revision, and yet any sentence can be maintained in the face of any evidence. The most important consequence of the combination of holism and underdetermination according to Quine is that:

it becomes folly to seek a boundary between synthetic statements, which hold contingently on experience, and analytic statements, which hold come what may. Any statement can be held true come what may, if we make drastic enough adjustments elsewhere in the system...Conversely, by the same token, no statement is immune from revision (Quine 1951a/1980, 43)³⁷

If every statement is equally open to revision or being held true come what may, then we cannot maintain the analytic-synthetic distinction.

The first thing to notice here is that Quine's characterisation of analytic and synthetic statements, as those "which hold come what may" or "which are immune from revision" and those "which hold contingently on experience" respectively, does not capture Carnap's notion of analyticity (Carnap 1963a, 921).³⁸ It is not clear whether this is something Quine thinks Carnap does accept or has to accept.³⁹ Either way, he is wrong. In his response to Quine, Carnap makes clear that (when worded more clearly), he is in agreement with Quine about everything except Quine's conclusion:

Quine has emphasized that in revising the total system of science no statement and no rule is immune or sacrosanct... Thus far I agree with Quine... However, I cannot agree with Quine when he concludes about this that there is no sharp boundary between physics and logic. In my view it is not a feature of the explicandum "analyticity" that these statements are sacrosanct, that they never should nor can be revoked in the revision of science... Since in changing the logical structure of language everything can be changed (Carnap 1952b/1990, 431–32)

In much the same way, changing certain rules of a game also changes the game we are playing (deciding to allow holding the ball in our hands means we stop playing football and begin playing rugby or netball), but this does not entail that any rules of any game are "held firm come what may". We just can't revise the rules that are constitutive of a game, without changing the game we are playing. But we can still play by any set of rules we want to, abandoning or accepting rules as we please.

The attribution of truth values to synthetic sentences changes continually, induced by new observations, even during a period in which the logical structure of language remains

³⁷ My reading of Quine's argument here agrees with Gary Ebbs', as opposed to what he calls the 'standard interpretation' (Ebbs 2016, 125).

³⁸ However, Quine himself uses the characterisation of those statements 'maintained independently of our observations of the world' as a potential non-arbitrary criterion for those truths made true by convention in 1936 (Quine 1936/1966, 95). Frost-Arnold suggests that it may be a result of the influence of Quine's teacher, C.I. Lewis, who characterised a priori statements in this way (Frost-Arnold 2013, 97).

³⁹ A charitable interpretation of Quine's argument here is given in Ebbs (2016).

unchanged. A revision of this sort is not possible for the analytic sentences. (Carnap 1952b/1990, 432)

Carnap's point is that there are senses in which analytic statements are both revisable and unrevisable.

In the broadest sense, all sentences are revisable. This is a direct consequence of the principle of tolerance. But within a specific language system, an analytic sentence is not revisable in the same way as a synthetic sentence.

Since the truth of an analytic sentence depends on the meaning, and is determined by the language rules and not the observed facts, then an analytic sentence is indeed "unrevisable" in another sense: it remains true and analytic as long as the language rules are not changed. (Carnap 1952b/1990, 432)

To put it another way, whilst analytic statements are not held true come what may and we can choose to drop them, it is also not possible for them to be rendered false in their respective language system in light of revisions within our body of beliefs. This is a subtle clarification, and one that could easily be overlooked.

Carnap was explicit about his belief in the revisability of all sentences, including the constitutive rules of a language: 'all rules are laid down with the reservation that they may be altered as soon as it seems expedient to do so' (Carnap 1934d/2001, 318). And if these rules are open to revision, so are all analytic statements too, since it is these rules that determine the domain of analytic truth for the language. It should be clear then, that analyticity and revisability are completely independent characteristics. What sets analytic statements apart is their constitutive role in the language's structure, not any epistemological characteristics concerning immunity to revision as Quine so frequently implies.

Confirmation holism says that no single sentence has associated with it a unique set of confirmatory experiences. But for Carnap, analytic statements are constitutive not substantive; they have no empirical content. They make no claims on experience. Analytic claims are not true because they are confirmed by all experiences. They don't require an appeal to experience at all. As constitutive components of a language system, the analytic statements cannot come out as false within that system. But if in light of disconfirmatory evidence the language system as a whole consistently fails to provide usable predictions, then it is reasonable to consider the analytic statements as candidates for revision. If the system as a whole is failing, then the system may need to be reconfigured from the ground up.

Perhaps it would be worthwhile to introduce a new term to distinguish between the senses in which analytic and synthetic terms are revisable. I suggest differentiating a sentence being *revisable* from a sentence being *surrenderable*. By "revisable", we mean much the same as Quine's usage: a sentence is revisable (in a particular language) if its truth-value can be freely changed. A sentence is surrenderable if changing its truth-value necessitates a change of language: changing it means giving up on the language too. By "surrenderable", we therefore mean that whilst the sentence is not "held true come what may", neither can its truth-value be changed within the current language. Clearly on this criteria, synthetic sentences for Carnap are revisable whilst analytic sentences are only surrenderable. I think this captures

the essence of Carnap's insistence that 'revision of this sort [i.e changing truth-value within the language] is not possible for the analytic sentences.' (Carnap 1952b/1990, 432). He doesn't want to say they are unrevisable but has no term for the sense in which they are open to change. Hopefully "surrenderable" does the job.

Whilst the analytic statements of a language system are not therefore revisable in exactly the same sense as the synthetic statements, they are surrenderable, and therefore also answerable to holistic revision in light of disconfirmatory empirical evidence. Far from ignoring the full consequences of confirmational holism, Carnap renders even the process of language change a holistic one, showing that the analytic/synthetic distinction is not incompatible with confirmation holism. The obvious response from a Quinean is that this response assumes and relies on the distinction between substantive and constitutive, and therefore upon the distinction between analytic and synthetic. This would be right. What this shows is that, given some sort of verificationism, semantic holism follows from conformational holism only if the analytic/synthetic distinction is rejected. So we find at the bottom of Quine's and Carnap's dispute a clash of opposing presuppositions.

5.3.6 *Changing Beliefs vs Changing Language*

What this distinction between surrenderable and revisable underlines is that, for Carnap, there is a difference between a change of language and a change of beliefs within a language. Carnap describes it as:

a distinction between two kinds of readjustment in the case of a conflict with experience, namely, between a change in the language, and a mere change in or addition of, a truth-value ascribed to an indeterminate statement, (i.e., a statement whose truth value is not fixed by the rules of language) (Carnap 1963a, 921)

Quine objects that this distinction is pointless, it does no philosophical work. The differences between the revisability of sentences is not one of principle, but of varying degrees. For Quine, sentences are differentiated in terms of our willingness to revise them, determined by how centrally they are located within our web of belief. Statements of maths and logic are less likely to be revised because they are central to our web of beliefs and they are consequently less 'germane' because of our 'natural tendency to disturb the total system as little as possible' (Quine 1951a/1980: 43–44). Revision of those more central beliefs entail more extensive redistributions of truth-values, but this is a difference in quantity not quality. The difference between revising "analytic" and "synthetic" statements is one of degree, not of kind. According to Quine's challenge, it becomes incumbent upon Carnap 'to find a principled distinction between two kinds of reasons for rejecting sentences' (Richardson 1997a, 148). One sort of reason for analytic sentences, and one for synthetic.

As presented in much secondary literature, Carnap does attempt to draw a principled and decisive distinction between the two types of revision:

As Quine reads Carnap's views, then, they depend upon an epistemological contrast: internal questions are to be settled by rule-governed procedures of justification, procedures that are obviously unavailable for external questions. (Hylton 2002, 19)⁴⁰

This interpretation stems from Carnap's 'strict and principled distinction between the theoretical and the practical' (Richardson 2007, 299). This distinction is a crucial one for Carnap, as it underlies his principle of tolerance. The choice of language is constitutive of the scheme of justification, and therefore justification prior to choice of language is not possible. The distinction between internal and external revisions becomes a distinction between rule-governed epistemological revisions and pragmatic considerations of the language-systems overall utility. In cases of internal revision, the rules of the language along with the adopted rules for confirmation and justification can determine whether or not a specific revision is justified. But this does not apply to external revisions, because they concern the adoption of a language. The two types of revision are distinguished in terms of the type of considerations relevant to them.

The first problem with this characterisation of the distinction, at least for a defender of Carnap, is that it cannot withstand Quine's arguments. As Quine understands Carnap, these internal revisions are practiced according to the confirmation relations of the language. But as Quine argues, these confirmation theories are not up to the task. This is not an argument that Carnap's own attempts thus far have all failed, but that all such attempts will necessarily fail because of holism. Confirmatory and falsifying experiences do not uniquely designate the specific belief (or beliefs) within the theory to be rejected. It is only ever the theory as a body that is confirmed or falsified (Quine 1951a/1980, 40–41). As a result, not only is it not possible for the justification-scheme of a language-system to determine the correct beliefs to retain or reject, but the choice between rival theories would seem to take on the same justificatory significance that Carnap attributes to internal revisions. If theories are the unit of significance, then we cannot distinguish between changes within and changes between theories.

Once we recognise holism (and underdetermination), it becomes clear that no system of justification can uniquely determine the right or wrong revision to make in all cases. We must therefore allow that pragmatic revision also occurs within theories. This blurs the line between the internal and the external: 'in whatever sense the framework for science is pragmatic, so is the rest of science' (Quine 1991, 272). In other words, whatever pragmatic considerations are relevant to the choice between frameworks are equally applicable to internal revisions. Consequently, Quine claims that '[Carnap's] pragmatism leaves off at the imagined boundary between the analytic and the synthetic. In repudiating such a boundary I espouse a more thorough pragmatism' (Quine 1951a/1980, 46). This "more thorough pragmatism" is a realisation imposed by the acceptance of holism and underdetermination, both of which Carnap accepts (although Quine doesn't acknowledge this). As a result, it seems that Carnap has no choice but to accept this "more thorough pragmatism" and reject the distinction between internal and external revisions.

⁴⁰ See also Hylton (2017, 2021).

The second problem with this characterisation of the internal-external distinction, and far more troubling for the Quinean, is that Carnap never makes it. I find no evidence of the strict overlap of the distinctions internal-theoretical and external-pragmatic in Carnap's work; that is, a seamless mapping of internal questions to theoretical issues and external questions to pragmatic decisions. So what is Carnap's actual position? How does Carnap defend the external-internal distinction, and how does he respond to Quine's "more thorough pragmatism"? These are two separable questions.

First, let's make clear that Carnap already accepts Quine's "more thorough going pragmatism", albeit in his own way. As we have seen, according to Carnap's principle of tolerance, the choice between alternative languages is always a pragmatic one, never a question of right or wrong. As Creath argues, this conventionalism is not limited to the logical, but includes every aspect of the language of science we adopt. Questions about the principles of induction and deduction, about holism, and the structure of scientific theory are all adopted on the basis of convention too (Creath 1992, 154–55). Again, there is no fact of the matter. There is no one objectively privileged language of science which we must discover. The choice between languages, and therefore the choice to change language, is always a practical decision. This much of Quine's interpretation is true.

But there is clear evidence that Carnap recognises that pragmatic considerations can also be relevant to internal questions, and are not relevant exclusively to choices between languages. Carnap discusses the situation in which a recalcitrant experience is incompatible with our theory and 'some change must be made' (Carnap 1934d/2001, 317). First Carnap makes clear that in these cases, it is possible that 'the P-rules [laws of physics] can be altered... the protocol-sentence can be taken as being non-valid; or again the L-rules [maths and logic]... can also be changed' (Carnap 1934d/2001, 317). In other words, everything is open to revision. This is Quine's position being advocated by Carnap 17 years before "Two Dogmas". Even more importantly:

There are no established rules for the kind of change which must be made. Further, it is not possible to lay down any set of rules as to how new primitive laws are to be established on the basis of actually stated protocol-sentences (Carnap 1934d/2001, 317)

Carnap is discussing the revision of analytic *and* synthetic statements. In other words, external *and* internal revisions. Although Carnap does not use the term, this means that pragmatic decisions are inevitable, in cases of both internal and external revision. Friedman also recognises this aspect of Carnap's thought (Friedman 2006, 48). It seems clear then, that Carnap's pragmatism is no less thorough than Quine's.

The question becomes how he can maintain the distinction between internal and external questions, whilst also admitting that pragmatic considerations are relevant to both. Or, to put it differently, what means does Carnap have to distinguish the internal from the external? Whilst we have rejected the rigid dichotomy, this does not render the two indistinguishable. The difficulty lies in specifying this difference without either contradicting scientific practice or incurring unwanted philosophical commitments.

It might be thought that the distinction can be drawn as follows. Whilst Carnap recognised that pragmatic reasons play a role in internal revisions, these revisions are not exclusively pragmatic. However, it is still the case that external revisions are exclusively pragmatic. Rather than strictly rule-determined revisions, internal revisions are guided by rules, but these rules are (as a consequence of holism and underdetermination) insufficient to eliminate the need for pragmatic considerations entirely. In some cases the rules may be sufficient to give a correct answer, but not in all cases. By contrast, external revisions are always and solely governed by pragmatic thinking. There is then a distinction, between the partially-pragmatic internal revisions and solely-pragmatic external, it is just not the strictly dichotomous one that has often been supposed.

But again, the Quinean can raise a problem. Kemp and Hylton utilise an argument made by Hilary Putnam to show that even a change of language can be theoretically motivated.⁴¹ This would undermine the partial- vs solely-pragmatic means of drawing the distinction, as it would be evidence of a theoretical (and therefore not solely pragmatic) external revision. This is incompatible with the supposed epistemological significance of the principle of tolerance; 'Synthetic sentences are answerable to evidence; analytic sentences are a matter of the choice of language, which does not require theoretical justification' (Hylton and Kemp 2019, §3.1).

Putnam argues that, whilst there are some analytic statements that will be accepted solely on the basis of decisions, these will be those trivial cases involving single-criterion concepts. For instance "All bachelors are unmarried" is analytic, as "bachelor" is a concept with a single criterion for application: being a single man. Such examples of analytic statements are of 'little or no systematic import. . . there could hardly be *theoretical* reasons for accepting or rejecting it' (Putnam 1962, 68). These statements lack further connections to the web of beliefs. Choosing to adopt or reject these concepts will have no repercussions for the rest of a system of beliefs, and similarly a rejection will never be prompted by changes elsewhere in the system.

But for any more complicated, non-trivial definitions this is not the case. Putnam argues that many important definitions in science are what he calls law-cluster concepts, constituted 'by a cluster of laws which, as it were, determine the identity of the concept' (Putnam 1962, 52). These concepts do have systematic import, and are enmeshed in a wider body of theory. Putnam uses the example of the pre-relativity definition of kinetic energy $e=1/2mv^2$. This appears to be a definition, and would therefore seem to be analytic. But this definition was rejected as part of the adoption of relativistic mechanics, and considered false. What this shows is that, at least in some cases, the reasons for changing language are 'of the same sort that lead us to make revisions to synthetic sentences' (Hylton and Kemp 2019, §3.1).

The crucial point here for the Quinean against Carnap is that these apparently analytic definitions are on an epistemological par with the synthetic statements of a theory. The reasoning, motivations, and evidence that would trigger and justify the

⁴¹ Hylton also utilises Putnam's argument in (Hylton 2007, 66–67). Putnam's argument is also approved by Quine (1960, 57n).

revision of an apparently analytic statement are of the exact same kind as would figure in the revision of a synthetic statement. The revision of analytic statements has been shown by Putnam to be triggered by the same theoretical justifications and answerable to the same body of evidence as the synthetic statements of the language. Where is the privilege that the principle of tolerance was supposed to provide analytic statements? It would seem then that there is no difference between internal and external revisions. Again, they argue, the distinction between internal and external breaks down.

So then, how are we to distinguish internal and external questions to avoid this collapse? As we saw, Carnap allows that pragmatic thinking can play a role in internal revisions. In fact, when holism and underdetermination combine, this is the only way. But pragmatic decision-making in such cases functions to break a deadlock of insufficient justification, not because the question to be answered has no right or wrong answer. Underdetermination and holism place limitations on our capacity to know, but not on truth. This is a fundamentally different situation from that we face when answering external questions. So whilst pragmatic factors are at play in both internal and external questions, they are at play for different reasons. In the case of external questions, this is a consequence of non-truth-valuability. In the case of internal questions, it is as a consequence of insufficient justification, which does not alter the truth-valuability of the questions themselves.

What has to be admitted in the face of Putnam's argument is that theoretical considerations likewise play a role in answering both internal and external questions. But this is an admission Carnap was happy to make:

The decision of accepting the thing language, although itself not of a cognitive nature, will nevertheless be influenced by theoretical knowledge, just like any other deliberate decision concerning the acceptance of linguistic or other rules. (Carnap 1950, 23)

For Carnap, the difference between internal and external revisions is not the reasons or motivations that may trigger such changes. Revisions within or changes between language systems are both potential consequences of a variety of factors. The important distinction between the two lies in the fact that unlike internal revisions, external revisions are not truth-valuable. They are not truth-valuable because they are not answers to factual questions concerned with what is the case, but instrumental value questions about what should be done or, as here, about what the point and purpose of a language under construction should be.

With the distinction between internal and external questions drawn thus thinly, it is of course highly questionable whether it can answer to the Quinean demand that an epistemological reason, turning on the nature of analyticity itself, be given for upholding the disputed categorical distinction between changes in belief and changes in language. How then is Quine's demand to be answered? The answer is that the demand is misplaced when it is directed at Carnap's notion of analyticity. One type of question has truth-valuable answers, the other does not. The reason for this lies in the different type of problems to which they respond, one factual, the other volitional.

To the objection that this does not enlighten us as to how analytic truths fulfil their epistemological office, it must be responded that this should occasion no surprise when it is realized that, contrary to the presumption with which Quine's demand was raised, analytic truths do not have an epistemological office.

5.3.7 *The Significance of the Analytic/Synthetic Distinction*

A surprising eventuality given the debate as detailed up to this point is that in 1974 Quine adopts a notion of analyticity in his own work, and retains it from then on.⁴² He allows that we can use “analytic” to characterise statements where ‘*everybody* learns it is true by learning its words’ (Quine 1974, 79), extending analyticity to all first-order logical truths and their logical consequences. And yet looking back on his argument in “Two Dogmas”, Quine says that ‘my reservations over analyticity are the same as ever’, and that his new notion of analyticity is ‘obvious and useful but epistemologically insignificant’ (Quine 1991, 271). This confirms that it was the supposed epistemological significance of Carnap's conception of analyticity that had been, and remained, the element with which Quine was fundamentally dissatisfied. This is how we should understand Quine's claim that Carnap ‘make[s] such heavy demands on’ the notion of analyticity that they cannot be met’ (Quine 1954, 122).

Sadly, Quine is never sufficiently explicit either about why Carnap's notion of analyticity is epistemologically insignificant, or what epistemological significance he takes Carnap to need. And Carnap never explicitly responds to these criticisms, or even asks Quine for sufficient clarification.⁴³ Despite the importance of the issue under discussion here, it is largely addressed only implicitly in both Carnap and Quine's actual work. It therefore requires a degree of reconstruction. To do so, there are three key questions to answer. First, what epistemological significance does Carnap's project require of a conception of analyticity? Second, what epistemological significance does Quine *think* it requires? Third, if Quine is mistaken (which I will argue he is) what is the cause of this mistake?

To understand why Carnap's distinction between analytic and synthetic must be distanced from questions of epistemological significance, we must consider the purposes of his project. Carnap refers to his project, which I have argued should be understood as a form of naturalism, as the logic of science. This, Carnap describes as the ‘analysis of the linguistic expressions of science’ (Carnap 1938/1991, 393). And the purposes of these analyses? Carnap says ‘the purposes for which semantic analyses are made... [are] the analysis, interpretation, clarification, or construction

⁴² As Hylton emphasizes, this adoption does not precipitate any fundamental revisions in Quine's thinking, either in broad structure or doctrinal specifics (Hylton 2021).

⁴³ Carnap never had the opportunity to discuss or interrogate Quine's conception of analyticity, as it was only published after Carnap's death in 1970.

of languages of communication, especially languages of science' (Carnap 1950, 39). And what sort of questions will the logician of science be concerned with answering?

Is such and such a theory, t_2 , compatible or incompatible with theory t_1 ?...Do the two concepts c_1 and c_2 (which differ in their definitions) have the same meaning?...Does p_2 follow from p_1 with logical necessity? Or at least with the necessity of the laws of nature? (Carnap 1934a/1987, 47)

What is clear from Carnap's description is that logic of science is linguistically oriented. It is also a formal enquiry; 'logic of science can progress without exception according to the formal method' (Carnap 1934c, 9).⁴⁴ It is concerned with language, specifically the language of science, not (at least primarily) problems of inference and justification. It can be useful to think of the logic of science as composed of three sets of tasks: defensive, analytical, and constructive. The defensive task involves 'unmasking' and eliminating meaningless metaphysical terminology (Neurath 1932c, 67). The analytical task involves the process of conceptual clarification, analysis of the terms and theories of science, as well as the formal analysis of systems of implication, consistencies between theories, and so on. The most important and ambitious part of the logic of science is the constructive, which involves providing proposals: outlining possibilities and explicating alternatives for the language of science. These could be proposals for single terms, rules of inference, or whole languages. Explications are one such form of proposal. Specifically, proposals for the potential usage of terms.

One terminological clarification should briefly be made here. In common usage, proposal can have a slightly stronger connotation than it does for Carnap; a proposal can seem like something we ought to adopt. Carnap means it more loosely. His proposals are possibilities for consideration. Whether they should be adopted or not is a pragmatic question, determined by our purposes. Even when Carnap thinks a proposal is a good one, it is still only ever presented as an option, to be adopted for instrumental purposes. The normative value of these proposals is conditional on the pursuit of specific goals.

The logician of science supplies explications as potential refinements to the language of science. More ambitiously, the logician of science could present an entire language as a proposal, to show how a certain selection of L-rules, along with specified inference rules, would provide a language of science with specific advantages and disadvantages.⁴⁵ With these advantages and disadvantages made explicit in this way, the choice of the language of science, although still pragmatic, would be a better informed one. Whilst still not truth-valuable, the choice is also clearly not arbitrary. The choice of tool is motivated by the task it is intended to perform. 'The acceptance or rejection of linguistic forms in any branch of science will finally be decided by their efficiency as instruments' (Carnap 1950, 40). The

⁴⁴ Formal here is meant broadly, since from 1936 onwards it is taken to include semantics.

⁴⁵ For an example of the way Carnap thinks languages can be designed with specific purposes in mind, see Carnap (1936–37, 4–6).

primary purpose of the logician of science is therefore the provision of a wide range of efficient linguistic tools to the sciences; the creation of the most useful terminology and language-systems for the practices of scientists.⁴⁶

Clearly, this is not the traditional project of epistemology as prior to the sciences. This is not the quasi-foundationalist quest motivated by the need for certainty that Carnap is sometimes interpreted as supplying. In a talk given at the 1935 Paris Congress, Carnap declared the task of contemporary scientific philosophy to be 'the transition from epistemology to the logic of science' (Carnap, Quoted in Richardson 1996, 309).⁴⁷ Rather it is a scientific activity conducted collaboratively with other scientists. Logic of science is a form of metatheory, science of the language of science, conducted within and with the assistance of the rest of science.

He who wishes to investigate the questions of the logic of science must, therefore, renounce the proud claims of a philosophy that sits enthroned above the special sciences, and must realize that he is working in exactly the same field, only with a somewhat different emphasis (Carnap 1934d/2001, 332)

Together, scientists and logicians of science will 'cooperatively explore the technical consequences of adopting [proposals]' (Creath 2007a, 323).

With this understanding of Carnap's project in mind, we are now in a position to ask why Carnap needs the notion of analyticity? What role does it play in the logic of science? Why is it so important for logic of science? Carnap's answer: 'I believe that the distinction between analytic and synthetic statements, expressed in whatever terms, is practically indispensable for methodological and philosophical discussions.' (Carnap 1963a, 921). We need it because of the way it allows us to propose and evaluate potential language systems. This is after all the point of the logic of science. The task of the logician is to experiment with the potential language-systems that scientists could use. And it is through sets of L-rules that these language-systems can be explicitly constructed, tested and communicated. This is why Carnap was unphased by Quine's worries about the lack of a principled dividing line resulting in a gerrymandering of the analytic-synthetic distinction. In the construction of a language system, the analytic/synthetic distinction is not intended to line-up with or adhere to some external or pre-existing cleavage. What looks to Quine like problematic gerrymandering is for Carnap simply the engineering of a distinction between the analytic and synthetic statements of a language that best suits the purposes for which the language is intended. That's just what it means to generate proposals.

Aside from generating proposals, Carnap's methods for formal language construction can also be put to use in the analysis of extant scientific theories. We

⁴⁶ As Leitgeb and Carus note, even if the Carnapian is successful in refuting Quine's arguments against the analytic/synthetic distinction for artificial languages, a line of attack remains open to the Quinean in questioning the value of these artificial constructions for the development of science (Leitgeb and Carus 2020, Supplement B). Although it is hoped that the current chapter demonstrates this utility.

⁴⁷ See also Uebel (2018).

noted above Carnap's willingness to allow the extension of linguistic frameworks via the introduction of so-called P-rules. The potential value of these introductions is exhibited in Carnap's remarkable analysis of the metrical fundamental tensor " $g_{\mu\nu}$ " which determines the metrical structure of physical space (Carnap 1934d/2001, 178). Carnap asks whether this term is a mathematical or a physical one. He then provides two possible languages, one in which the tensor is a logical symbol, and the result is a Euclidean homogenous space, and another in which the tensor is a descriptive term and consequently we have Einsteinian non-homogenous space (Carnap 1934d/2001, 178). What is distinguished here are the special and general theories of relativity; in the latter it is a physical term, in the former it is not. For scientists concerned with the non-Euclidean consequences of relativity theory, the latter language is to be adopted. Here we see how the semantic structure of scientific theories conditions their expressive power, and correlatingly how the choice of construction has direct and significant consequences for our understanding of scientific theories and, ultimately for the pursuit of science itself.

These are the kind of analyses that Carnap sees the logician of science as supplying. The provision of such analyses and proposals is the primary purpose of the logic of science, a purpose one cannot achieve without the use of the constitutive potential of analytic terms. This also explains the importance of the principle of tolerance for Carnap. If we cannot countenance the possibility of multiple languages and multiple systems of logic, then the extent to which logicians of science can provide proposals is severely limited. The extent to which the logician of science can contribute within Carnap's project is far more significant than it is within Quine's.

So what does Carnap require from his conception of analyticity to facilitate this project? Fundamentally, he needs to provide a semantic distinction between the substantive and the constitutive components of a language, which he does by separating the analytic and synthetic. This is all the significance Carnap needs, and is what his account supplies him. Analyticity therefore stands revealed as a 'semantical concept', not an epistemological or metaphysical one (Carnap 1963a, 917). It is not intended to capture, and is thus not burdened with, any ontological or epistemological baggage. The distinction itself is not being drawn on a principled or inherent epistemological basis, it makes no metaphysical claims about the nature of their truth, or the sentences in and of themselves. It also, importantly, has no bearing on the certainty or otherwise of our beliefs. But additionally, we can now see why Carnap advocates an analytic/synthetic distinction at all. Because its useful. Because it allows us to do logic of science. Carnap's primary motivation was not epistemological, but purely analytical.

I think here we should remember Stein's anecdote, that ultimately Carnap believed time would tell whose project was more successful. Carnap believed that the long-term utility of logic of science and its contribution to the practice of science would vindicate the assumptions on which it functions. In other words, the basis for the analytic/synthetic distinction is its practicality within the broader project of logic of science. And the same is true of the internal/external distinction. A more fundamental, philosophical defence is not forthcoming. Carnap's analytic/synthetic distinction does not need to latch onto or capture something pre-existing and more fundamental.

5.3.8 *Semantics and Epistemology*

So why does Quine insist that Carnap's account lacks sufficient epistemological significance? What significance is Quine looking for that Carnap fails to provide? I think we can reconstruct it, by looking at the basis of the misunderstandings and disagreements between Carnap and Quine. But first, it must briefly be made clear what the relationship between analyticity and epistemology is. To do so, we must make clear Carnap's division of the semantic and the pragmatic. Semantics proper, or 'pure semantics', involves 'the study of *constructed language systems* given by their rules' (Carnap 1955a, 33). This includes questions of meaning, of the deductive relationships between statements, and ultimately also the theory of truth (Carnap 1942/1948, 10). Semantics applies to language at a certain level of abstraction, as a formal system divorced from its users and uses. By contrast, pragmatics encompasses all enquiries concerned with the ways in which language is used by people. Pragmatics is then an empirical discipline.

In its narrow sense, pragmatics refers to investigations of how languages function in the wild: 'the empirical investigation of historically given *natural languages*' (Carnap 1955a, 33). This includes what Carnap calls descriptive semantics and descriptive syntax, empirical studies of how natural languages are actually deployed by their speakers. Carnap gives the hypothetical example of 'a certain Eskimo language not previously investigated' (Carnap 1942/1948, 12). Through observation, the facts about the typical usage of terms within the language can be established. Through observation of the use of the term "igloo", its syntax (that it is a predicate) and semantics (that it designates certain buildings) can be ascertained. Semantical investigations proper, "pure semantics", whilst possible in isolation, can also turn, as part of the explicatory processes, to the clarification of such natural language terms. In these cases, semantics can provide more exactly specified counterparts to pragmatical concepts:

Many of the concepts used today in pure semantics were indeed suggested by corresponding pragmatical concepts which had been used for natural languages by philosophers or linguists, though usually without exact definitions. Those semantical concepts were, in a sense, intended as explicata for the corresponding pragmatical concepts (Carnap 1955a, 34)

It is this situation that obtains between the semantic concept of analyticity, the L-Rules, and their pragmatic counter-parts.

But Carnap also deployed a wider conception of pragmatics than customary today.⁴⁸ According to this:

we may first distinguish between those problems - or components in complex problems - which are of a factual, empirical, rather than logical nature. They occur especially in the

⁴⁸ Uebel has previously characterised these two senses of pragmatics, the wide and the narrow, as 'methodological pragmatics' and 'linguistic pragmatics' (Uebel 2013a, 536). Uebel's account of Carnap's wider notion of pragmatics agrees with the account of Carnap's pragmatics found in Stein (1992), Carus (2007a), and Richardson (2007).

theory of knowledge and the philosophy of science. If construed as problems of semiotic, they belong to pragmatics. They have to do, for instance, with the activities of perception, observation, comparison, registration, confirmation, etc., as far as these activities lead to or refer to knowledge formulated in language (Carnap 1942/1948, 245)

Epistemological questions, concerning ‘the activities of gaining and communicating knowledge’, are not concerned with the formal properties of an abstracted language system (Carnap 1942/1948, 250). Discussions of confirmation, justification, belief, verification and so on require a reference to a speaker (Carnap 1942/1948, 8). They cannot be separated from the epistemic agent doing the knowing, believing or verifying. All epistemology therefore falls within the wider-scoped pragmatics.⁴⁹

It is therefore within the realm of pragmatics that much discussion of the traditional conception of analyticity occurs. Remember Carnap’s informal characterisation of the explicandum “analyticity” as “truth based on meaning” and “understanding being sufficient for the determination of truth”. These characterisations are focused on the means by which an epistemic agent arrives at knowledge of analytic statements. Similarly, Stein describes the basis of Carnap’s beliefs as being ‘that the classification of a sentence as analytic in some sense ‘explains’ how we know it’ (Stein 1992, 282). In dividing the pragmatic from the semantic, Carnap makes clear that these epistemological questions are a separate and additional set of worries, not intended to be addressed or resolved via the analytic/synthetic distinction.

The crucial point here is that the analytic/synthetic distinction as drawn by Carnap is first and foremost a semantic one. It is in the domain of semantics that the relevant distinctions can be drawn precisely. Inevitably, by contrast, the pragmatic counterpart of analyticity remains more or less vague, though explication of individual instances can provide some clarification. The motivation and means for drawing the analytic/synthetic distinction is not the assumption of a principled epistemological distinction, but rather the more instrumental and utilitarian goal of a distinction between the constitutive and substantive components of a language, to facilitate linguistic construction. But this is not how Quine interpreted it, as we will now see.

5.4 Divergent Conceptions of Empiricism?

As we have seen, Quine and Carnap, despite initially seeming to share very similar interests (both were scientifically minded empiricists, with a scepticism of much traditional philosophy, a tendency towards metaphysical deflationism, and a deep interest in formal logic) disagree consistently, over topics both small and large. For many recent philosophers looking back at the debate, there is a frustration that despite constantly acknowledged disagreement, Carnap and Quine never managed

⁴⁹ See also Carnap (1955b).

to resolve much of anything. That Carnap and Quine were to some extent talking past one another is now more or less a matter of consensus. They argued over numerous specific doctrinal issues, but never addressed the elephant in the room: there are foundational disagreements that are the cause of these specific arguments. Various recent philosophers have noticed this.⁵⁰ Drawing on this scholarship, we can produce a convincing account of what this basic disagreement is, an account that explains both the various doctrinal disputes, and the broader failure to communicate.

Richardson argues that a misunderstanding and disagreement about the nature of empiricism is fundamental to understanding the dispute. He argues that Quine consistently misinterpreted Carnap's motivations, and that as a result he continually argued against beliefs and positions that Carnap never adopted, holding Carnap to criteria that were not relevant to him. First, we will address Quine's conception of empiricism. Then, his interpretation of Carnap's empiricism. And finally, Carnap's own view of empiricism.

Quine's empiricism is one of his most basic commitments, 'it is a commitment to the claim that all significant discourse stands in confirmation relations with experience' (Richardson 2004, 69–70). This attitude underlies all of Quine's work. Although Quine never (so far as I know) defines precisely what he means by empiricism, he does give some hints. He for instance refers to empiricism as 'a theory of evidence' (Quine 1981a, 39). It is clear however that Quine considers it the correct theory; '[Quine] takes it for granted as well that an empiricism of some sort is correct' (Creath 2007a, 330). It is also clear that Quine sees his own empiricism as the most recent milestone in the historical development of empiricism, beginning with the British empiricists of the eighteenth century (Quine 1981b, 67–8). Quine's acceptance of the correctness of empiricism is a fundamental epistemological commitment.

Quine also sees Carnap within this empiricist tradition. And whilst Carnap clearly is an empiricist, he is not the empiricist Quine typically portrays him as. Quine believes that 'Carnap was driven to [the analytic-synthetic] distinction in an effort to explain the certainty of logical and mathematical truth and that he sought such an explanation due to a prior and independent commitment to empiricism' (Richardson 1997a, 155). Friedman similarly highlights how, whilst ostensibly discussing Carnap, Quine frequently discusses classical Kantian notions of analyticity (Friedman 2006, 39–40). In discussing the difference between analytic and synthetic statements, Quine describes the characteristic of the former as their 'inward necessity' (Quine 1934/1990, 37, Quine, 1936/1966, 95). Quine's lectures on Hume similarly opens with the claim that 'the theory of knowledge begins as a quest for certainty' (Quine 1946b/2008, 50). "Carnap and Logical Truth" claims that the empiricist attempt to answer the question '[h]ow is logical certainty possible?...reached its maturity in the work of Carnap' (Quine, 1954/1966: 100). And again, in his paper in Carnap's *Library of Living Philosophers* volume,

⁵⁰ See Stein (1992, 279), Richardson (1997a, 152), Friedman (2006, 39), and Leitgeb and Carus (2020, §7)).

Quine begins with the question '[h]ow is logical certainty possible?' (Quine 1963, 385). In Quine's framing, the pursuit of certainty is central to Carnap's empiricist epistemology. This would explain Quine's repeated assumption that analyticity requires unrevisability, since what is certain isn't revisable. The basic motivation, on Quine's reading of Carnap, for drawing the analytic/synthetic distinction is epistemological.

We can support this analysis whilst simultaneously being charitable to Quine, by noting that Quine's attribution of this concern with certainty to Carnap may have been accurate prior to 1932. In his symposium talk of 1930 for instance, Carnap accepts that 'any uncertainty in the foundations of [mathematics] is extremely disconcerting' (Carnap 1931/1983, 41). But with Carnap's embrace of the principle of tolerance in 1932 and abandonment of traditional epistemology around 1934, this concern ceases to be relevant to him. Quine's mistake is not necessarily in his original diagnosis, but his failure to recognise that this diagnosis no longer applies to Carnap's post-1934 thought.

Richardson argues that both of Quine's assumptions are false. He argues that Carnap only adopts empiricism in the early-30s, having been strongly influenced by neo-Kantianism at least until the *Aufbau*. Richardson therefore argues that the adoption of the analytic-synthetic distinction is prior to Carnap's adoption of empiricism, not epistemologically motivated by it. This depends on readings of the early Carnap beyond my brief here, so I set it aside. But Richardson also argues that Carnap's distinction must already be in place within a formal language before that language's 'empiricist credentials can be investigated' because for Carnap, a language is essentially defined by the way the analytic/synthetic distinction is drawn (Richardson 1997a, 157).

This interpretation is supported by Carnap's own statements on the nature of empiricism. 'It seems to me that it is preferable to formulate the principle of empiricism not in the form of an assertion—"all knowledge is empirical"...but rather in the form of a proposal or requirement' (Carnap 1936–37, 33). For Carnap, empiricism is a proposal, a convention for which sorts of languages to work with, adopted for pragmatic reasons. Carnap's empiricism has more in common with Neurath's stance than with Quine's doctrinal empiricism. Empiricism does not require the analytic-synthetic distinction for epistemological purposes. Rather, Carnap advocates, on pragmatic grounds, certain ways of drawing this distinction so as to provide the most useful means for constructing languages for conducting science.

The crucial contrast comes to this. For Quine, empiricism is an epistemological position on the nature of evidence. For Carnap, empiricism is a proposal that a certain subset of language-systems, with certain rules of inference and reduction sentences, will maximise the efficacy of human cognitive practices and therefore provide the most useful languages for the practice of science. By Van Fraassen's standards, as discussed in Chap. 2, Quine's empiricism is far more naïve than Carnap's. The traditionally empiricist epistemological position Quine attributes to Carnap brings with it commitments and motivations that Carnap's pragmatic position does not. Carnap's empiricism only makes sense within the context of his

adoption of the principle of tolerance and his project of the logic of science. And again, these are two topics which Quine does not touch on in detail.

Quine's references to the principle of tolerance are dismissive or disparaging, but never amount to real criticism or analysis. In correspondence with Carnap, Quine claims (one suspects not entirely sincerely) to 'fear your principle of tolerance may finally lead you even to tolerate Hitler' (Quine, 4/02/1938, in Creath 1990b, 241). Elsewhere, Quine asks exasperatedly: 'All this tolerance of language and waiving of commitments is reasonable enough, but is there no end to it?' (Quine 1951b, 66). In their correspondence, Isaacson recalls Quine calling the principle of tolerance 'the serpent in Carnap's garden' (Quoted in Verhaegh 2018, 140). This suggests that Quine failed to fully recognise the significance of the principle of tolerance, and the profound reorientation of philosophy into logic of science that it precipitates. As Richardson pithily puts it, 'Quine's Carnap is a very traditional philosopher—he is simply a technically competent empiricist. If Hume had logic, he would have been Carnap' (Richardson 2004, 69).

The epistemological significance that Quine expects Carnap to supply is that demanded by traditional empiricism. Quine's Carnap needs a version of the analytic-synthetic distinction that not only supplies epistemological certainty, but that supplies a metaphysical dividing line between two different ways in which statements are made true. This is a distinction that is inherent and objective, not conventional and language-relative. This worry exists in Quine's work from his very first engagement with Carnap. His concern over the absence of a principled distinction and the possibility of indefinite expansion develops into this concern over epistemological significance. If what truths count as analytic is simply a matter of convention, then what *principled* reason is there for deciding certain truths should be analytic whilst others are not? Specifically, what reason is there to decide to draw the analytic-synthetic distinction in the way the logical empiricists wanted to, so as to conform to Hume's fork? Why not draw it literally anywhere else? Like Buridan's ass, Quine can see no reason to choose one potential location for the division over any other, so he ends up choosing none. This concern becomes, in the mature Quine, the argument that the analytic/synthetic distinction has no epistemological significance. After all, if we can draw the line anywhere, what does the line do? But once we see Carnap's project for what it was, a version of naturalised scientific metatheory, it becomes clear that Quine viewed Carnap's project as a failure because he applied the criteria of a different project to it.

5.5 Conclusion

The purpose of this chapter was to demonstrate that Carnap can defend his position from Quine's criticisms, once we understand Carnap's logic of science as diverging from traditional epistemology. Specifically, I aimed to show that with a language-relative purely semantic notion of analyticity, the naturalistic

considerations like holism and underdetermination which Quine thinks undermines Carnap are consistent with and capable of integration into the logic of science. Most important however, was to show that many of Quine's criticisms are misdirected because Quine did not recognise what the purposes of the logic of science were. He interpreted Carnap as a more or less traditional empiricist epistemologist (along the lines of Schlick or Ayer), when Carnap's project is premised on a radical reorientation of philosophy into a formal metatheory of science within science. Without recognising the importance of this reorientation, it is also impossible to understand Carnap's more specific arguments. What the above demonstrates is that Quine's arguments only demonstrate the incompatibility of naturalism with a more robust or traditionalist conception of analyticity.

One final caveat for those who feel Quine has been done a disservice in the preceding chapter. I have defended Carnap from Quine in part by pointing to Quine's misunderstandings.⁵¹ I do however recognise that a parallel version of the preceding story may well be told from Quine's perspective, a story in which Quine's arguments are repeatedly met by replies from Carnap that fail to address Quine's worries. What is important is that whilst such a story may better represent Quine and his work, its possibility does not undermine the preceding demonstration of Quine's failure to invalidate Carnap's project.

⁵¹ I cannot claim to have exhaustively refuted every criticism that has been raised against the analytic/synthetic distinction here. The focus has been on the objections that I understand to be the most pressing, from a Quinean perspective in particular, for Carnap's semantic (and not the traditional conception) of analyticity.

Chapter 6

Challenges to the Bipartite Metatheory Interpretation



So far, I have argued for the compatibility and complementarity of Neurath and Carnap's mature work. But their relationship was never as easy as between some intellectual collaborators (Marx and Engels for instance). Despite their fruitful intellectual relationship and close friendship, Carnap and Neurath remained very different in temperament and personality. As their friend and colleague Hempel said, '[t]here could hardly have been a more striking contrast in philosophical style within the same school of thought than that between Carnap and Otto Neurath' (Hempel 1975, 6).¹ And it cannot be denied that Neurath and Carnap were engaged in ongoing disagreements. For some scholars, this has been taken as indicative of fundamental disagreement, which would be a severe challenge for the bipartite metatheory conception. In this chapter, I will show that all these apparent conceptual ruptures can be overcome, reinforcing the compatibility of their respective projects.

6.1 Analyticity

Carnap maintains a unique semantic conception of the analytic/synthetic distinction that I have argued is consistent with epistemological naturalism. But what about Neurath? It is occasionally claimed that Neurath, like Quine, rejects the notion of analyticity. Creath for example says Neurath maintains no analytic/synthetic distinction (Creath 1996, 163). More commonly it is noted that Neurath argues for the revisability of analytic statements, and the revisability of their status as analytic,

¹ This difference is apparent in how their colleagues describe them. Popper recalls Neurath as 'a man of tremendous vitality and drive, of a man who believed passionately in his social, political and philosophical theories... who would not look behind him or, when rushing ahead, care very much about whom his big stride might knock down' (Popper 1973, 52). By contrast Feigl remembers Carnap 'always as extremely rational, calm, well-balanced' (Feigl 1970, XIII).

although the status of the analytic/synthetic distinction itself is left somewhat ambiguous.² So let's be clear, Neurath does not reject the analytic/synthetic distinction:

We shall admit without reservation that analytic statements (cf. Carnap's proposal concerning the definition of 'analytic', 'contradictory', 'synthetic') are treated differently from content [i.e. synthetic] statements. . . *This distinction is essential.* (Neurath 1934/1983, 103)

He not only accepts the coherence of the distinction, but specifically cites Carnap's characterisation of it approvingly.

So why is this not widely recognised? One factor is the aforementioned assumption that naturalism requires a rejection of analyticity. Additionally, there is a tendency to read Neurath through the lens of Quine, Neurath being read as a precursor and Carnap as his foil (Koppelberg 1989). This misunderstanding is not innocuous. It not only deflates the originality and significance of Neurath's work, but it also maintains the illusion of a theoretical rift between Carnap and Neurath. Neurath and Carnap's conceptions of analyticity are one and the same. By equating Neurath's naturalism to Quine's, it can be easy to assume that he shared Quine's criticisms of Carnap. When Neurath is read on his own terms, this mistake is easily avoided.

6.2 Ballungen and Explication

Neurath repeatedly claims that language is inherently murky, and that it will never be rendered fully determinate. 'Imprecise "verbal clusters" [*Ballungen*] are somehow always part of the [encyclopedia]. If imprecision is diminished at one place, it may well re-appear at another place to a stronger degree' (Neurath 1932a/1983, 92). And as Hempel put it, Neurath 'looked at Carnap's reliance on precise model languages with respect but deep misgivings' (Hempel 1975, 6). Is this account of language incompatible with Carnap's methodology of explicationism? Is not explication an attempt to refine and clarify language in exactly the way Neurath denies is possible? Sometimes what Neurath says would suggest so:

The more one reflects upon this matter the more one is led to abandon the notion of *absolutely precise terms* even as limit of a refining process. The supposition that there are such limits seems to me to be another illustration of what may be called "Pseudo-rationalism" (Neurath 1937c/1983, 174)

Elsewhere, Neurath claims that in 'many cases a new shaping of language is certainly superfluous, indeed dangerous', questioning the practical expediency of retaining previously metaphysical terms with new interpretations (Neurath 1931c/1983, 65). In practice Neurath was linguistically cautious, with such terms typically avoided, added to what he jokingly called his *Index Verborum Prohibitorum*.

² See for instance Rutte (1982b/1991, 188), Zolo (1989, 41), and Ruytinx (1983, 43).

Mormann refers to Neurath's picture of language as a non-Cartesian one (Mormann 1996). By this, he means that Neurath rejects two crucial assumptions about language made by Descartes. Specifically, Descartes believed it possible to arrive at a final, completed language, built upon clear and distinct ideas. Neurath rejects this as a pseudo-rationalist fantasy. Instead, he argues that the very notion of a final language is delusional, that all language should be recognised as provisional, and that all languages will inevitably be permeated with a degree of uncertainty. This uncertainty comes partly from *Ballungen*. It 'is the job of Ballungen to preclude Carnap's recourse to such clean elements... note that Ballungen preclude definite, once-and-for-all fixed meanings' (Cartwright et al. 1996, 157). But along with this pragmatic reason for assuming the impurity of language, there is also an epistemological reason. For Neurath, as we saw, Ballungen are also a consequence of our inability to survey all our knowledge at once:

our experience in this is like that of a miner who at some spot of the mine raises his lamp and spreads light, while all the rest lies in total darkness. If an adjacent part is illuminated those parts vanish in the dark that were lit only just now. (Neurath 1921/1973, 198)

Even if we engage in a process of linguistic refinement, we cannot view our language in its totality. Even if we deliberately intervene in language, language still has a life of its own. Does this not render Carnapian explication a waste of time? Why would Neurath support the pursuit of a goal that can never be reached? Mormann contends that Carnap never fully understood or responded to Neurath's *Ballungen*; 'For him vagueness, opacity were defects of the actually existing language of science that "in principle" could be neglected' (Mormann 1999, 174).

This suggestion of a fundamental rift is however a significant exaggeration. For one, the characterisation of Carnap it relies on is wrong. Carnap doesn't neglect vagueness in natural language, he developed an entire methodology to alleviate it: explication. Neurath is concerned with natural languages, those he repeatedly reminds us we are born into and have little control over. We must 'consider the current plenitude of science as our base of operations' (Neurath 1937d/1987, 138). Ballungen are not features of constructed formal languages, but organic outgrowths of language at work in the world. But Carnap agrees with Neurath about the indeterminacy of natural language. That's his entire motivation for explicating things to begin with.

Let us step back and ask, what are *Ballungen*? The German term "Ballung" means 'congestion or concentration' (Cartwright and Cat 1996, 82). Neurath proposes 'clot' as a translation (Neurath 1944/1971, 18). Now look at Neurath's descriptions of them:

unprecise terms ("clusters") occur that I best characterise by the nature of their application (Neurath 1934/1983, 105)
complex (messy) statements of little cleanliness (Neurath 1935b/1983, 128)

Ballungen are non-deliberate assemblages, unhelpful organic linguistic conglomerations. These are not simply concepts held together by Wittgensteinian family-resemblances. Rather, these clusters seem to contain various ill-defined concepts,

blurred together. *Ballungen* are then imprecise, messy notions, best characterised by their use, which are given to us by natural language and which typically contain a multiplicity of meanings and uses. In other words, they are the perfect raw material for teasing apart and sharpening via Carnapian explication.

Neurath even concedes that his propensity for getting rid of terms rather than revising their meanings is a personal preference rather than a necessity; ‘I cannot pretend that no new definitions could be proposed which would avoid a particular danger, but I do not like to act as a terminological rope-dancer’ (Neurath 1941/1983, 217). Neurath’s disinterest in explicatory processes was a matter of instinctive preference, not an inevitable consequence of his philosophy. Given the utility of Carnap’s method, as demonstrated in Chap. 4, we can see that explication is not simply finicky pedantry as suggested by Neurath’s comparison to “rope-dancing”, but of great utility to the Metatheory of science. Neurath’s preferences were mistaken, he underestimated the potential of such linguistic engineering.

Carnapian explication does not propose to provide a complete and final clarification of natural language, and nor does it ignore or reject the possibility of the re-emergence of linguistic unclarity.³ What Carnap intends is the provision of linguistic tools of greater precision than those provided by natural language, for the purposes of certain scientific practice. Neurath describes his universal slang as the ‘purified everyday language’ (Neurath 1937c/1983, 180), ‘composed of ordinary terms of everyday language (certain dangerous terms omitted) and of certain added scientific terms’ (Neurath 1937b/1983, 202). These additional scientific terms include explications, the process of purification includes the process of explication. Far from incompatibility, there is a basis for agreement and collaboration between Neurath and Carnap here.

6.3 Conceptions of Protocol Statements

Carnap and Neurath were not always in complete agreement over the form and content of protocol statements; there wouldn’t have been a protocol sentence debate otherwise. But what is most important for current purposes is the compatibility of their mature, fully-developed notions of protocol statements. Given the significance of protocol statements for each of their respective projects, a fundamental disagreement over what protocol sentences are would pose a significant challenge to the compatibility of their respective projects, and therefore to Uebel’s bipartite metatheory interpretation.

Whilst Neurath’s proposals are not set in stone, his basic conception of protocol statements is stable from the early 30s onwards. This conception was explored at length in Chap. 3. For Neurath, protocol statements are physicalist statements

³ Carus claims the Carnapian can actually identify where *Ballungen* are most likely to appear (Carus 2007a, 282).

describing the perceptual states of an observer and information about the spatio-temporal context of observation. Protocol statements have no epistemic privilege, they are corrigible and revisable. It is only the very specific (and initially unintuitive) structure for protocol sentences that undergoes significant revision in Neurath's conception, moving from a triple- to a quadruple-embedded structure. It is the quadruple-embedded conception that I consider Neurath's mature conception. By contrast, Carnap's notion of a protocol sentence undergoes frequent and significant alterations. I will therefore start with a brief overview of Carnap's developing conception of protocol statements during the first half of the 1930s, before arriving at his mature conception in 1936 in "Testability and Meaning".

6.3.1 Carnap's Changing Conception

For current purposes, three main periods in the development of Carnap's conception of protocol statements can be identified.⁴

1. Early-Carnap: pre-1930
2. Middle-Carnap: 1930–35
3. Mature-Carnap: 1936 onwards

The early Carnap includes the *Aufbau* era, and pre-dates the protocol sentence debates. The middle is the period of intense fluctuation of his conception, including the crucial period of the protocol sentence debates. The mature period begins with "Testability and Meaning" in 1936, where Carnap presents a conception which is essentially retained from then on.

Prior to the protocol sentence debate, in the *Aufbau* Carnap utilised sentences in the auto-psychological language describing the immediate given (at bottom remembrances of similarity relations). In 1930, Carnap allows both methodological materialism and solipsism, treating physicalist and auto-psychological languages as equally basic, continuing to refer to statements as 'reducible to the given' (Carnap 1930/1959, 145). The auto-psychological protocol statements were understood as incorrigible. However, translation into the physicalist language was necessary for scientific purposes, because only a physicalist language 'makes inter-subjective knowledge possible' (Carnap 1930/1959, 144). Carnap therefore concludes that both constitution systems are equally 'correct and indispensable', the physicalist because it allows for the inter-subjective testability required by science, and the solipsistic for epistemological purposes because it allows justification via reduction to the given (Carnap 1930/1959, 144). There is however still an asymmetry of translation, with primacy given to the auto-psychological. Whilst physicalist statements are translatable into the auto-psychological language, the reverse is

⁴ This is a slight simplification. Uebel identifies five different positions maintained by Carnap between the *Aufbau* and "Testability and Meaning" (Uebel 2007a, 442–443).

not entirely possible; only the inter-subjectively verifiable component of an auto-psychological protocol is translatable, but a component of the meaning is private to the speaker.⁵

The major change in Carnap's conception of protocol statements, in which these prior commitments are abandoned, came in late 1932 in "On Protocol statements". There Carnap withdraws the epistemological primacy and privilege of the auto-psychological. Physicalistic statements can now serve epistemological purposes. With this epistemological privilege revoked, protocol statements also lost their certainty and incorrigibility. This renunciation of epistemological privilege is the most important change that Carnap's conception undergoes during the protocol sentence debate. The consequence is a total conventionalism about protocol statements; 'Every concrete sentence of the physicalistic system language can serve under certain circumstances as a protocol sentence' (Carnap 1932d/1987, 465). Here Carnap's logical tolerance first manifests, there is no longer a fact of the matter about what constitutes a protocol statement, simply proposals to be evaluated according to their practical utility.⁶ And the proposal Carnap advocates is one according to which any statement can be considered a protocol statement. The form of protocol statements is completely unrestricted. However, this completely conventional conception of protocol statements appears only in "On Protocol statements".⁷ But whilst Carnap walks back this total conventionalism, his rejection of epistemological privilege is definitive.

Carnap's mature position arrives in 1936 in "Testability and Meaning". Here, Carnap settles for physicalist statements in a 'thing language' about medium sized objects as the most practical form for protocol statements (Carnap 1936–37, 466).⁸ Protocol statements are still neither certain nor incorrigible, but status as a protocol is no longer purely a matter of decision. Carnap proposes the requirement that protocol statements contain only inter-subjectively observable predicates, by which Carnap means that for a predicate P, a person 'is able under suitable circumstances to come to a decision with the help of few observations' that P or $\neg P$ is confirmed (Carnap 1936–37, 455). What this conception amounts to is physicalist protocol statements formulated in ordinary language containing only observable predicates. With this mature conception of protocol statements in hand, Carnap no longer focuses on debating competing conceptions, but instead attempts to build on his conception by providing deductive relations between theoretical and observational terminology.

⁵ This is made clear in unpublished manuscripts from 1930. See Uebel (2007a, 191–200).

⁶ See Awodey and Carus (2007, 183–192).

⁷ That this is not Carnap's final conception is not always clear in secondary literature. See for example Creath (1990a, 412), Coffa (1991, 371), and Richardson (1997a, 211; 2000, S158).

⁸ Carnap himself highlights 1936 as a decisive turning point in his thinking about protocols (Carnap 1963g, 886).

6.3.2 *Mature Positions, Apparent Disagreement*

This somewhat simplified story sees Carnap conceding a certain amount to Neurath, abandoning methodological solipsism and phenomenistic auto-psychological protocol statements as practically impossible, and adopting a physicalist protocol language. Their initial disagreement about the need for a phenomenal protocol language seems to be overcome, and a broad agreement reached; protocol statements are physicalistic reports, open to revision and accepted by decision. But crucially, Carnap never adopts the specifics of Neurath's account. Two key elements of Neurath's proposal are never embraced by Carnap. Firstly, the multiply-embedded bracket structure, and secondly, the requirement of contextualising information contained within the protocol statement. These requirements are intimately connected for Neurath. The purpose of the structure he advocates is to exemplify the contextualisation, as demonstrated in Chap. 3. But Carnap rejects both. The multiply-embedded structure was rejected as impractical, and the restrictions Carnap placed on contextual information never required containment within the protocol itself.

This point of contention surfaces in the correspondence between Carnap and Neurath, a decade after their axis of the protocol sentence debate seemed to have concluded:

As far as your formulations are concerned (which unfortunately are not in agreement with your opinions) I have told you since many years that I cannot accept them and hence I agree with the criticism of these formulations by R., Schlick, and many others. In distinction to R., I know your actual conception from conversations; and I am in agreement with it... I agree with R. in his criticism of your triple-involved form of protocol sentences, (Carnap, 15/03/1943, in Cat and Tuboly 2019, 577)

Neurath replies:

You always tell me, you agree with Russell's and Schlick's remarks on my protocol statements, but my statements intended are different. Please, tell me first what you think how I should express my statements properly and then please tell me, why even then they are not in harmony with your opinion (Neurath, 25/09/1943, in Cat and Tuboly 2019, 598)

This exchange suggests a lingering disagreement that is sadly left unresolved. Despite claiming that he is in agreement with Neurath, Carnap goes on to reject his proposal for the form of protocol statements. So did they agree or not? Making sense of this interaction is essential for understanding the compatibility of their mature conceptions of protocol sentences. There are two potential disagreements between Carnap and Neurath, which broadly correspond to the two elements of Neurath's proposal that Carnap doesn't adopt. The first is a disagreement about structure and logical form. The second is the issue of what information needs to be contained within the protocol statement itself.

6.3.3 Carnap's Formal Objection

The above quotes appear in Carnap and Neurath's correspondence during a discussion of Russell's *An Inquiry Into Meaning and Truth*. The passage Carnap refers to is one we encountered in Chap. 3:

Thus according to Neurath the data of empirical science are all of the following form: "A certain person (who happens to be myself, but this, we are told, is irrelevant) is aware at a certain time that a little while ago he believed a phrase which asserted that a little while before that he had seen a table." (Russell 1940/1995, 146)

Russell's criticism concerns Neurath's multiply-embedded bracket structure. If Carnap agrees with Russell as he claims, then his problem must also be with the structural and formal-logical aspect of Neurath's protocol. Carnap had questioned the logical structure of Neurath's protocol statements before, criticising the impracticality of a protocol sentences with 'three nested components' (Carnap 1932d/1987, 465). But Carnap also admits that he agrees with what he understands as Neurath's conception of protocol statements (physicalistic, revisable, unprivileged). Carnap's concern therefore seems to be that the physicalistic conception of a protocol statement is in conflict with Neurath's proposed bracketed structure. The issue of the logic of Neurath protocol statements was discussed at length in Chap. 3. It was demonstrated that Neurath's proposal prioritises the presentation of the factors relevant to the acceptance procedure, in a way that maximises ease of understanding and practical utility. But whilst the motivations for Neurath's proposal (and his disinterest in utilising proper logical form) are clear, we must still demonstrate that Carnap and Neurath's conceptions cohere.

Uebel has previously argued that Carnap and Neurath's disagreement over protocol sentences shines a light on the differences between their two projects (or the two halves of their joint project) (Uebel 2007d, 396). He is right, but simple difference of emphasis is not the whole story. There is a subtle but significant difference in Carnap and Neurath's usage of "protocol": Carnap's protocol statements (physicalist statements about observable mid-size objects) are the object sentences of Neurath's protocol statements. Carnap gives the example protocol statement: 'a black round table' (Carnap 1936–37, 13). This is strikingly similar to the object sentence of Neurath's protocol 'Karl's protocol: Karl formulates: Karl sees: In the room is a round table' (Quoted in and Translated by Uebel 2007c, 386). Is this a strange coincidence? Or is this different usage indicative of a conceptual disagreement? No on both counts. This terminological divergence should not be understood as indicating a deeper theoretical rupture. But to see why, it is helpful to frame the issue a certain way.

According to the bipartite metatheory interpretation, scientific metatheory is separated into the pragmatics of science as practiced by Neurath and the logic of science as practiced by Carnap. So which part of the metatheory do protocol statements belong to? As Carnap understands them, they belong to the logic of science. For Neurath, they belong to the pragmatics of science. Again we have the appearance of theoretical disagreement, but really this difference is only

terminological. To see this, we must remember that the reception of protocol statements is not an event, but a process. Initially, an observation report is checked against Neurath's conditions *(i)–*(iv). If these conditions are met, the protocol is valid, and the extraction of the object sentence is licenced. This entire process falls under the pragmatics of science, as it concerns acceptance. The data-bank, the output of the process of acceptance, is the starting point for Carnap's logic of science. There is then a clear continuity between the pragmatics and logic of science here. But crucially, they address different aspects of science. Logic of science concerns the logic of observation reports within the language of science, establishing definitions, patterns of deduction and so on. Pragmatics is concerned with the acceptance conditions for observation reports in scientific practice. The latter concerns science as an activity, where the former concerns science as a body of knowledge. The starting point for the logic of science is the endpoint of the pragmatics of science. Neurath's object sentences are the only element of the protocol statement that makes it into the language of science, in what I have called the data bank, because only they are relevant to scientific theory. Issues of observer reliability and competency are crucial to understanding the practice of science, but not for understanding scientific theories. Carnap starts from the presupposition that the statements of the language of science are acceptable and accepted. His concern is how these data are utilised once they have been accepted. Both Neurath and Carnap use "protocol sentence" to refer to observation reports as utilised within their sub-field of the Metatheory. But both are referring to one half of a process. What Carnap and Neurath ultimately disagree about is at what stage in the process we apply the term "protocol sentence". But there is no substantive theoretical disagreement here.

Neither Carnap or Neurath would have entertained the question of what a protocol statement "really" is. Both recognised the debate as one of competing proposals. Now we know their proposals are not theoretically incompatible, but simply require a change in terminology from one party. And I think for the purposes of terminological clarity, Neurath's proposal ought to be adopted. Firstly, Neurath's proposal has brought about the Neurathian terminology utilised throughout the discussion above. But perhaps more importantly, Neurath's usage of "protocol sentence" adheres more closely to the original intention: those basic evidence statements of science. As we saw, significant processing is required to arrive at Carnap's protocol sentences. The need now is for a term to replace Carnap's use of protocol. A tentative proposal of a name for evidence statements as utilised within the logic of science is "data sentences".

6.3.4 Carnap's Context Objection

We can now turn to Carnap's rejection of Neurath's requirement that protocol statements contain the contextual information to answer the question 'When, where, and how?' (Neurath 1946a/1983, 233). Carnap specifically rejects Neurath's requirement for 'designations of actions of perception' (Carnap 1936–37, 13). He

agrees with Neurath that ‘a certain connection between the basic sentences and our perceptions is required’, but ‘it is sufficient that the biological designations of perceptive activity occur in the formulation of the methodological requirement concerning the basic sentences. . . and that they need not occur in the basic sentences themselves’ (Carnap 1936–37, 13). The methodological requirement Carnap refers to here is the stipulation that protocol sentences must be inter-subjectively observable. As far as Carnap is concerned, the requirement of observability is sufficient. So long as the protocol statements are observable, the inclusion of contextualising information is superfluous. And Carnap is not wrong here. For his purposes, such information is superfluous. But as we have seen, for Neurath’s purposes it is far from it.

Carnap’s requirement of observability is a reformulation of his commitment to empiricism. As such, the requirement of a “certain connection” is in place to exclude the possibility of metaphysical statements entering the logic of science. But as should be clear now, Neurath’s demand for contextual information is not simply to guarantee such a connection. Neurath’s requirement is not simply an empiricist one. The contextual information itself is of great significance for Neurath, since it is this information that allows decisions on acceptance. Such information however is not necessary for Carnap. As we already saw, his logic of science starts from the assumption that the statements with which he is working are valid. That is what his observability criteria does. Carnap therefore misunderstands the significance of Neurath’s demands for contextual information.

There is one further possible complication here. Neurath’s criteria and Carnap’s requirement of observability are not the same. Carnap argues against Neurath’s protocol statements for having the limitation of being ‘intersubjectively confirmable but only *subjectively observable*’ (Carnap 1936–37, 11).

As Uebel notes, Neurath’s conditions (ii) and (iii), about speech-thinking and perception, whilst inter-subjectively confirmable, are not observable by others (Uebel 2007d, 395). Nor are descriptions of psychological states as object sentences, statements like “I feel angry”. But Carnap preferred an observable thing-language, intersubjectively confirmable and observable. Carnap therefore rejects Neurath’s inclusion of psychological predicates in physicalist statements, which rules out reports of one’s own psychological states. Framed this way, we seem to have a technical but significant disagreement. Carnap insists that observation statements of science need to be inter-subjectively observable and confirmable. Neurath requires that the terms of the protocol language need to be inter-subjectively confirmable, but can be only subjectively observable. But since Carnap’s logic of science starts with the object sentences supplied by the pragmatics of science, some of the object sentences Neurath supplies would simply be rejected by Carnap for failing to meet his criteria.

But before such conclusions are reached, it needs to be emphasised that the requirements on observability are, as Carnap recognises, a matter of decision about the language best suited to our purposes. The disagreement here is not one over what observability is, but what requirements will be most useful to adopt. Carnap, in line with his principle of tolerance, explicitly frames his choice of observability

conditions as a decision (Carnap 1936–37, 9–13). What practical reasons are there? That some predicates are only subjectively observable ‘is a serious disadvantage and constitutes reason against their choice’ (Carnap 1936–37, 12). Exactly why this is so disadvantageous is not spelled out in detail by Carnap. We can accept Carnap’s worries here, but his reasons are far from decisive. The question is then whether Neurath has better reasons than this for embracing inter-subjectively confirmable but only subjective observable predicates.

For Neurath, protocol statements like “Karl’s protocol: Karl formulates: Karl is feeling: Karl is scared” must be potentially valid as protocol statements because of their obvious significance for social sciences like sociology, history and anthropology. ‘Historians of human social life are highly interested in descriptive terms, such as deal with the feeling-tone of persons, their devotion their fear and hopes’ (Neurath 1944/1971, 15). Our physicalist language needs to allow for reports of ‘the state of a person who hears Beethoven or looks at certain forms of architecture’ (Neurath 1944/1971, 15). If anything, he argues, we need a more extensive, fine-grained terminology for describing feeling-tones. On this point, Neurath is unarguably right. The disciplines of psychology, anthropology and sociology would be significantly poorer under Carnap’s limitations. And importantly, Carnap concedes as much. The language described in “Testability and Meaning” is designed for limited purposes, and he notes that ‘we would have to take them as primitive predicates in a language of the whole of science. . . because in such a language we require them in any case’ (Carnap 1936–37, 12). The whole of science includes the social sciences. Ultimately then, Carnap can be read as voicing caution rather than outright disagreement. What Carnap ultimately argues is that, if a language doesn’t require psychological predicates, then it is more practical to do without them, since this allows for inter-subjective observability of all predicates. We can accept his point, whilst also recognising that the realities of practicing the social sciences means these criteria are not generalisable to unified science as a whole.

Ultimately then, any apparent disagreement between Neurath and Carnap over protocol statements is, if not entirely illusory, overstated and, most importantly, can be overcome. Their slight differences in terminology and differences of orientation with regards to their respective fields within the metatheory occasionally generate the appearance of disagreement, but this appearance is easily overcome by careful analysis.

6.4 Physicalism and Private Languages

In the early days of the Vienna Circle, Neurath advocated for traditional materialism against idealism, but Carnap convinced Neurath that materialism was as much a pseudo-thesis as idealism (Carnap 1963d, 51). Yet Neurath maintained that the spirit of materialism contained something valuable, which he developed into physicalism. The argumentative roles then flipped, with Neurath imploring Carnap to drop his methodological solipsism and embrace physicalism. An overview of Carnap’s

development between 1929 and 1936 was given in the previous section. As Carnap put it, ‘we abandoned the phenomenological language recognizing its subjective limitation’ (Carnap 1936a, 12).

In his intellectual autobiography, Carnap again endorses physicalism. ‘In my view, one of the most important advantages of the physicalistic language is its intersubjectivity.’ (Carnap 1963d, 51–2). And he credits Neurath for bringing him to the position:

Neurath emphasized from the beginning that language phenomena are events within the world, not something that refers to the world from outside... Neurath emphasized these facts in order to reject the view that there is something “higher”, something mysterious, “spiritual”, in language, a view which was prominent in German philosophy (Carnap 1963d, 29)

This position is stated clearly in 1937, and Neurath’s influence is explicitly cited by Carnap there too (Carnap 1936–37, 12).

And yet, there are still reasons to question whether Carnap understood the full force of Neurath’s argument for physicalism. As we saw in Chap. 2, Neurath argued that only through inter-subjective accessibility could consistency of usage in language be maintained and changes in usage be identified. Inter-subjectivity allows for other people to act as guarantors of linguistic competency and continuity of usage. Privately doing so is not possible because our own cognitive capacities are too limited. A private language has no guarantees and is therefore practically impossible. Neurath’s argument makes clear that physicalist language is not simply preferable, but practically necessary. So it seems worrying that, despite still describing himself as a physicalist, late in his life Carnap claims that ‘both system forms [physicalism and phenomenism] are possible’ (Carnap 1963c, 945). And this is not an isolated instance of poor phrasing on Carnap’s part. In his introduction to the Second Edition of the *Aufbau*, Carnap again seems to return to talk of the auto-psychological and physical as rival possibilities for language construction (Carnap 1961/2005, vii–viii). Discussing how he would do it differently were he to rewrite the *Aufbau*, Carnap says he would have started from ‘concrete sense data’ (Carnap 1961/2005, vii). How are we to make sense of this?

Some have interpreted this as indicative of a rift. Mormann has argued that Carnap misunderstands Neurath, and that only because of this misunderstanding can he think they share a commitment to physicalism (Mormann 1999, 172). But we need not draw such damning conclusions. Uebel has argued that Carnap’s claims need not be understood as a reversion to methodological solipsism when read in their proper context. He argues that we ought to distinguish ‘exploratory language constructions’ from the application of constructed languages to epistemological questions or the explication of vague or contested concepts (Uebel 2018, 377). What this distinction amounts to is a distinction between pure logic for logics sake, and logic applied to concrete situations. In the case of the former, only logical possibility is a relevant constraint, whilst in the latter practical concerns become relevant. Carnap’s discussions of the possibility of auto-psychological language construction must be understood as cases of exploratory logical work. This is borne out by the

context in which the quotes are made. In both cases, Carnap is concerned solely with the logic as a formal endeavour, and not with its applications. In the former case, he is discussing the possibility of constructing a logical system of concepts. In the latter case, he is introducing his *Aufbau*, which by that point was primarily (if not exclusively) valued for its technical, not epistemological, achievements.

With this in mind, we can see that speculation about exploratory logic does not fall foul of Neurath's physicalist requirements, because it is not relevant to it. Neurath's argument does not, and does not attempt, to demonstrate the logical impossibility of a private language. Rather, it demonstrates the practical impossibility of a functional private language based on the fallibility caused by the limitations of human memory and cognition. Neurath's private language argument, as detailed in Chap. 2, is relevant for constructed formal languages only in so far as they are intended to be made use of for purposes of communication, which is not true in the case of purely exploratory study. It is this same constraint of practicality, 'the great difficulty of constructing a clearly interpreted, practicable, purely phenomenalist language', that Carnap recognises as the reason for adopting a physicalist inter-subjective language (Carnap 1963c, 945).⁹ Whilst his tone is far more conciliatory and diplomatic than Neurath's, their conclusions are the same. Whilst both language systems may be logically possible, only a physicalist language is practically realisable. And Neurath and Carnap's continuing admiration for the *Aufbau* supports Uebel's conclusion here. After all, they both continue to hold the *Aufbau* in high esteem as a work of pure logic (Neurath can never have really valued it as anything else). They continue to recognise its value as a work of exploratory logic long after both have rejected foundationalism and reduction to the given. They continued to appreciate it as a work of technical ingenuity, rather than as something epistemologically substantive.

6.5 Truth and Semantics

One issue over which there unquestionably was disagreement between Carnap and Neurath was over the possibility of semantics in general, and a semantic notion of truth in particular. Carnap confirms as much in his intellectual autobiography:

Neurath believed that the semantical concept of truth could not be reconciled with a strictly empiricist and anti-metaphysical point of view... I showed that these objections were based on a misunderstanding of the semantical concept of truth (Carnap 1963d, 61)

Frustratingly though, they didn't agree on the nature of their disagreement. Whilst Neurath believed he had genuine theoretical concerns with Carnap's semantics, Carnap believed that Neurath's problems were a consequence of misunderstandings.

⁹ The same point is made elsewhere in the same volume, where Carnap says how he 'prefer[s] those arguments against methodological phenomenism which show that a language constructed on a physicalist basis is more suitable for the purposes of science' (Carnap 1963e, 864).

Whether Carnap and Neurath were at odds on this issue is not in question. The important questions are what type of disagreement it was, and more importantly whether they should have disagreed.

Both Uebel and Carus have recently argued that the debate is, as Carnap attempted to convince Neurath, not indicative of a fundamental theoretical rupture (Carus 2019; Uebel 2001, 2015).¹⁰ Whilst Uebel has argued that the differences are at least in principle bridgeable, Carus has argued that the disagreement stems from a difference of emphasis more than any fundamental differences between them. Drawing on their work, I argue that Neurath's opposition to semantics springs primarily from a combination of misunderstandings (albeit not without good reasons), over-caution and anti-metaphysical zeal, greatly exacerbated by Carnap and Neurath's geographical separation. Ultimately, Neurath was mistaken. However, it is relatively simple to say what he should have done: accept Carnap's conclusion.

6.5.1 *Carnap on Truth and Semantics*

In the early 1930s 'the notion of truth was regarded with suspicion in the circles of logical positivists. It seemed either superfluous or metaphysical' (Hempel 1983a/2000, 191). The correspondence theory of truth in particular was viewed as outdated, with only Schlick defending the 'good old phrase "agreement with reality"' (Schlick 1934/1979, 374). Both Frank and Hahn advocated a pragmatic conception of truth (Hahn 1933/1987, 43; Frank 1929/1941, 74). Hempel defended a coherence theory (Hempel 1935). As we saw in Chap. 3 Neurath was critical of the correspondence theory in particular, and (as we will see below) truth-talk more generally. Carnap had, like the other members of the Circle, previously voiced doubts about the notion of truth, but when 'Tarski explained his semantical theory to him, Carnap embraced it without agonizing and without delay' (Creath 1999, 72). Neurath was not so easily convinced.

Prior to 1935 Carnap had, like the rest of the left-wing of the Circle, been suspicious of any theory of truth, seemingly on syntactacist grounds: because '*truth and falsehood are not proper syntactical properties*' (Carnap 1934d/2001, 216). In *The Logical Syntax of Language*, Carnap attempts to show that truth cannot be defined in his syntactical system. His argument involves a version of the liar paradox and shows that the concept of truth, when defined in the object language as is the '*customary usage of the terms "true" and "false" leads, however, to a contradiction*' (Carnap 1934d/2001, 214). Carnap here realises that truth can only be

¹⁰ This is not a consensus position. Zolo for instance takes the correspondence of the 40s as a demonstration of their divergent and incompatible philosophical project (Zolo 1989, 168). And Mormann's interpretation will be discussed in detail below.

defined without contradiction in a meta-language, but opts instead to reject truth.¹¹ How to interpret Carnap's argument is a matter of much debate.¹² There is however consensus on the argument's lack of success. Coffa calls it 'one of the worst Carnap ever advanced' (Coffa 1987, 566), while Creath claims it 'is so bad that it is plausible to assume that Carnap was antecedently prejudiced against the concept of truth' (Creath 1990a, 411).

It is notable therefore that within a year, Carnap adopted the semantic conception of truth. Having been converted by Tarski, Carnap attempted to dispel the confusion under which his colleagues remained, with one of his talks at the 1935 Paris conference, later reprinted as part of "Truth and Confirmation" (Carnap 1949). Carnap begins by sharply separating truth and confirmation. A failure to do so he argues is the cause of much misunderstanding of truth-talk (Carnap 1936/1949, 119). Truth is time-independent where confirmation is time-dependent, a statement's degree of confirmation being specific to a certain point in time (Carnap 1936/1949, 119). Carnap goes on to argue against the mistaken notion that truth requires certainty. He shows that rejecting the possibility of absolute certainty is insufficient grounds for rejecting truth, because such a position tacitly relies on the principle that we can only employ concepts for which we can be absolutely certain of their applicability (Carnap 1936/1949, 123). But no concepts meet this criteria, it's impossibly strict. Carnap uses the uncontroversial everyday example "alcohol". We can't be absolutely certain that the term is applicable in each case, and yet this is clearly no grounds to abandon the concept (Carnap 1936/1949, 123). The same applies to truth. Truth is therefore consistent with anti-foundationalism.

Carnap again raises the spectre of the liar sentence, but now draws the opposite conclusion from it. Rather than abandon the concept of truth, we simply need to impose the limitation that truth must be defined for an object language in a meta-language. Contradiction is avoided, and truth can be retained. So what changed? Why did Carnap draw such different conclusions from the liar paradox only a year after *Logical Syntax*? Primarily, Carnap came to recognise the value of Tarski's definition, specifically the fact it retains 'the sense in which the word "true" is mostly used both in everyday life and in science' (Carnap 1936/1949, 121). By contrast, the assimilation of truth to confirmation required a complete deviation from typical usage of the concept in both science and everyday life. For instance, it requires the rejection of the principle of excluded middle, since there are many statements which neither themselves or their negations are scientifically accepted (Carnap 1936/1949, 119). What Tarski's definition provided for Carnap was a

¹¹ The discussion of meta-languages does not appear in the original 1934 German edition, but is not new to the English edition of 1937. The relevant material was cut from the original edition of *Logical Syntax* to save space, and published as the article "*Die Antinomien und die Unvollständigkeit der Mathematik*" (Carnap 1934e).

¹² Coffa attributes Carnap's argument to 'verificationist prejudice' (Coffa 1991, 304). Oberdan attributes it to Carnap's syntacticism (Oberdan 1992, 239; Oberdan 1993, 89–91). Ricketts argues it is 'rooted more in technical than in philosophical considerations' (Ricketts 1996, 231). Creath concludes there is no 'reading of the relevant passages that is fully satisfactory' (Creath 1999, 72).

more successful explication of truth than the Vienna Circle had previously had available to them. It retains the spirit of the original explicandum, but without either metaphysical or logical baggage (correspondence theory and liar paradox respectively). The use of meta-languages is a small price to pay for such a successful explication. Additionally, Carnap's syntacticism had been superseded since he wrote *Logical Syntax*. As Carus argues, Carnap had come to realise that syntax was insufficient for logical analysis of empirical science (Carus 1999, 21). This openness left Carnap sufficiently receptive to see the advantages of semantic truth.

6.5.2 Neurath on Truth

As we saw in Chap. 3, Neurath rejected the correspondence theory of truth, and the notion of comparison that underpins it. Consequently, Neurath has frequently been interpreted as advocating a coherence theory of truth by contemporary critics (Schlick 1934/1979, von Juhos 1935, Weinberg 1936, Russell 1940/1995, Ayer 1959), contemporary defenders (Hempel 1935) and modern commentators (Rutte 1982a/1991; Wolenski 1996; Haller 1996; Grundmann 1996; Niiniluoto 1999; Young 2018). Neurath himself denied it in his published work (Neurath 1934/1987, 101; 1944, 47–48 n.). But he was arguably even more forceful with his denials in private.¹³ But one contemporary who recognised this was Carnap: 'Neurath vehemently rejected this interpretation... there cannot be any doubt that Neurath never held this conception' (Carnap 1963a, 864).¹⁴ More recent specialist scholarship has reached something of a consensus on this.¹⁵ But if the coherence interpretation is wrong, what was Neurath's approach to truth?

As demonstrated in Chap. 3, Neurath argued that the correspondence theory of truth is incoherent, that the notion of comparing reality and language makes no sense. This rejection of correspondence theory, which Neurath (plausibly) considered the default theory of truth, left him with a significant suspicion of all truth-talk. Yet Neurath does not immediately abandon using the term "true". In 1931, Neurath argues that agreement with our existing body of statements is all "true" can ever amount to; '*There can be no other concept of 'truth' for science*' (Neurath 1931b/1983, 53). But over time, he becomes increasingly suspicious of even this proposed usage of the term:

It has become evident that the use of the terms "true" and "false" easily leads to all kinds of difficulties. One can completely renounce the use of these terms, but one can also try to redefine them appropriately (Neurath 1936d/1983, 161)

¹³ See Mancosu (2008, 204–205).

¹⁴ Hempel later came to the same conclusion (Hempel 1983a/2000, 192).

¹⁵ See Mancosu (2008, 30, 193) and Uebel (2015).

He ultimately opts for the former option and abandons talk of truth entirely, in favour of referring to the acceptance and rejection of statements (Neurath 1941/1983, 221–222). He ultimately opts for the former option and abandons talk of truth entirely, in favour of referring to the acceptance and rejection of statements (Neurath 1941/1983, 221–222). Uebel rightly argues that Neurath ‘intend[ed] a radical reinterpretation of our truth talk—not as an interpretation of the concept of truth—in terms of statement acceptance’ (Uebel 2015, 30). Similarly, Hempel argues that what Neurath provides should not really be considered a theory of truth so much as a sociological description of the conditions for acceptance and acceptability in science (Hempel 1983a/2000, 192). What has frequently been interpreted as a coherence theory is not an attempt at an analysis of “truth”, but an explication, a deflationary terminological replacement. He does not provide an analysis of the concept, of what “truth” really is, but provides an alternative basis for deploying our terminology to re-orient talk of truth conditions toward talk about acceptance conditions. This is not only far more radical than providing a coherence theory, but has obvious parallels with Carnap’s explicationist methodology generally, and Carnap’s notion of truth in particular.

But what motivated Neurath’s continued rejection of truth talk? Why did he not embrace Tarski as Carnap did? This is a more complicated issue. None of the Carnap-Neurath semantics debate appeared in published articles. It happened at conferences, in conversations and private correspondence. The correspondence is particularly frustrating. There is no developed argumentation, only a counter-positioning of accusations and opposed ideas; Neurath accuses Carnap of metaphysics, Carnap says Neurath has misunderstood him, repeat ad nauseum. Carnap seems particularly tired of the debate. He laments having had the same argument with Neurath since 1935 (Carnap, 11/05/1943, in Cat and Tuboly 2019, 581). Despite this, the ideas underlying the disagreement can be uncovered with a little work.

Fundamentally, Neurath was worried by the potentially metaphysical consequences of legitimising semantics. This worry is ever-present from 1935 onwards and is grounded in his fear of absolutism (Neurath 1941/1983, 221–222). Neurath saw in Tarski’s account a revival of the correspondence theory.¹⁶ Neurath had long maintained that ‘a rigorous empiricism reject[s] anything that smacks of the “absolute”’ (Neurath 1931a/1983, 48). And he repeatedly worries that the term “true” by itself implies ‘a kind of Aristotelian absolutism’ (Neurath 1946d, 80). This obsession has both philosophical and political motivations. Philosophically, absolutism was the hallmark of Neurath’s hated pseudo-rationalism. But there are also more concrete, real-world considerations which help to explain Neurath’s adamance. The correspondence took place during the last years of Nazism, and

¹⁶ Neurath was not alone in this interpretation. Popper similarly understood Tarski as ‘ha[ving] finally rehabilitated the much maligned correspondence theory of truth’ (Popper 1974/2002, 112). And Tarski certainly suggested as much when he claimed his ‘concept of truth... is to be included here, at least in its classical interpretation, according to which ‘true’ signifies the same as ‘corresponding with reality’ (Tarski 1936/1956, 401).

Neurath was concerned with understanding the intellectual underpinnings of totalitarianism. Absolutistic thought, Neurath argues, leads to absolutistic politics. He sees Nazism as an outgrowth of an absolutist tendency that dates back as far as Plato, and argues ‘this merciless habit in history very often is connected with absolutism in metaphysics and faith’ (Neurath, 25/09/1943, in Cat and Tuboly 2019, 592). This understanding of the connection between absolutist metaphysics and pluralist anti-metaphysics is shared by Frank (1950, 112–122). It also has obvious continuities with the concerns of the Vienna Circle in its early days.

Neurath repeatedly warns that despite Tarski’s theory being proposed for formal languages, it will inevitably be extended beyond its intended remit, leading to metaphysical speculation. But, as Mancosu notes, he ‘did not object formally to the theory nor to its application within formalized languages’ (Mancosu 2008, 215). In fact, Neurath concedes that Tarski’s theory of truth may be applicable to formal language, and therefore that Tarski and Carnap’s semantics could be formally useful, but insists that they are inapplicable to the ‘empiricist field’ (Neurath 1944/1971, 13). When he sees semantics step beyond this limited logical role, he declares himself ‘depressed to see here all the Aristotelian metaphysics in full glint and glamour, bewitching my dear friend Carnap’ (Neurath, 15/01/1943, in Cat and Tuboly 2019, 570).¹⁷ It is hard to interpret Neurath’s argument here as anything more than a slippery slope fallacy. If he concedes that semantics is formally acceptable, then rejecting it simply on the possibility of it being misunderstood is a drastic overreaction. Mathematics is capable of Platonist, metaphysical interpretation, but that would be a preposterous reason for abandoning it. And Carnap evidently felt so too: ‘You reject certain statements because they might perhaps be meant in a metaphysical way’ (Carnap, 7/11/1942, in Cat and Tuboly 2019, 562). Without a genuine argument, Neurath relies on ‘emotional reactions, namely your dislike of the term “truth” and your vague fear that this way would finally lead us back to old metaphysics’ (Carnap, 11/05/1943, in Cat and Tuboly 2019, 581). Neurath’s anti-metaphysical stance is here approaching paranoia.

There is little to say in defence of Neurath’s argument here. Neurath was wrong, Carnap was right. Carnap was right both that Neurath’s arguments don’t amount to much more than personal distaste, private worries and slippery slope arguments, and that they are premised on a misunderstanding of Carnap’s notion of truth. But if this is the case, what should Neurath have said about truth? Put simply, he should have followed Carnap and adopted a deflationist account of truth. And any attempt to pursue the Bipartite Metatheory in the same spirit ought to adopt the same position.

¹⁷ Neurath’s concerns about the danger of Aristotelianism were reinforced by Frank, who expressed concerns to Neurath about the creeping resurgence of scholasticism amongst his students (Reisch 2005, 209–210).

6.5.3 Deflationary Truth

The success of Tarski's definition has been hotly debated, with many arguing that it fails to penetrate or provide the intuitive notion of truth. But as Carus rightly points out, Carnap has no such ambitions (Carus 1999, 24). Carnap told Neurath, the 'concept of truth is not metaphysical but scientific...[and] is not at all meant in the sense of 'absolutely certain', 'indubitable' or anything like that' (Carnap, 4/02/1944, Cat and Tuboly 2019, 609). Carnap's notion of truth involves providing rules for a specific language. As Carnap makes clear, what he provides is not a definition of the concept "true", but a decision on the rules to govern the term (Carnap 1942/1948, 26). There is no 'language-transcendent' notion of truth at play here (Carus 2019, 349). What Carnap desired was an explication, and that's exactly what Tarski provided.

Carnap's account of truth is a deflationary one, according to which '*to assert that a sentence is true means the same as to assert the sentence itself*' (Carnap 1942/1948, 26). This is almost the dictionary definition of deflationism.¹⁸ On such an account there is no need for analysis of the nature of truth, since it does not designate a unique property or relation, "truth" adds nothing substantive. As Carnap emphasises, two sentences "x" and "the sentence 'x' is true" are 'different formulations for the same factual content' (Carnap 1936/1949, 121).¹⁹ As Rickett's notes, truth is always 'disquotationally eliminable' (Ricketts 1994, 192). All that's required from a deflationary theory is an account of its logical and linguistic functions. It was the realisation of the deflationary potential of Tarski's definition that explains Carnap's sudden change of heart:

Carnap used to tell his students a story about the first time Tarski explained to him his ideas on truth. They were at a coffeehouse, and Carnap challenged Tarski to explain how truth was defined for an empirical sentence such as "This table is black". Tarski answered that "This table is black" is true iff this table is black; and then, Carnap explained, "the scales fell from my eyes." (Coffa 1991, 304)

Carnap realised how little Tarski's theory entails. It brings with it no metaphysics, no relationship of correspondence, no property or relation of truth. Tarski's definition is (or can be) as minimal as any anti-metaphysician could want.²⁰ Uebel has previously argued that Neurath should have simply adopted a minimal, disquotational notion of truth as a supplement to his pragmatic theory of acceptance, and that doing so would not violate the anti-metaphysical spirit of his philosophy (Uebel 2001, 2007d, 2015). I agree with Carnap and Uebel's conclusion.

¹⁸ Compare this definition from the SEP: 'According to the deflationary theory of truth, to assert that a statement is true is just to assert the statement itself' (Stoljar and Damnjanovic 2010).

¹⁹ This is a clear statement of what is sometimes called the Equivalence Schema. See (Armour-Garb et al. 2022, §1.1).

²⁰ The questions of the correctness of Carnap's interpretation of Tarski and the details of his theory of truth must be left for another time.

Importantly, even if Tarskian truth has sometimes been interpreted (perhaps even by Tarski himself) as a correspondence theory, it wasn't by Carnap. Carnap even makes very similar arguments to Neurath's against the notion of comparing statement with fact on which the correspondence theory hinges. Carnap argues that a comparison is made in virtue of shared property (like colour or size), but that such a relationship doesn't capture the relation of world to language. He suggests this is better characterised as a confrontation, and urges us to recognise that confrontation is a two-stage process, 'the formulation of an observation and the confrontation of statements with each other' (Carnap 1936/1949), 127). This doesn't sound like correspondence theory, it sounds like Neurath. Carnap implores Neurath to realise that:

in criticizing a concept used by an author one should not criticize the term he uses or the bad things other people have said with the same term, but the meaning which the author gives to the term (Carnap, 11/05/1943, in Cat and Tuboly 2019, 581)

Carnap provides a deflationary explication, nothing more substantial than the formal account that Neurath has already conceded is acceptable.

Neurath argues that truth is redundant, that by the addition of the concept "true" to the Encyclopedia, the body of statements 'is not affected', it simply 'surround[s] it by a layer of formulations' (Neurath 1937e/1983, 206–207). But this (not utter) redundancy is consistent with a deflationist account. Uebel has called his belief that we can do without a notion of truth at all Neurath's 'blind-spot' (Uebel 2007d, 429). As Carnap shows, "true" has its uses. Carnap obviously makes particular use of truth-talk in his formal work, but "true" also allows for what Quine called semantic ascent, the move from talking about the world to talking about language. This allows us to make statements like "Everything you have said is true". Carnap clearly sees his conception of truth as consistent with the spirit of Neurath's thinking here.

In so far as he was willing to use the term, what Neurath provided was not a new substantive theory of truth, but an alternative usage that the term "truth" could be put to, which is what Carnap's explication will provide. This, we can now see, is perfectly in line with Carnap's own deflationary motivations. And, as recent scholarship on Carnap's syntax has shown, it is now being widely recognised that there is significant continuity between Carnap's pre- and post-semantics thought. Much of what Carnap circa-1934 (and therefore Neurath) understood as syntax is now considered semantics.²¹ Neurath had already, unknowingly, embraced semantics.

Unfortunately, Carnap was never able to correct Neurath's misapprehension, and this is in large part a consequence of external non-philosophical factors. I have previously mentioned the importance of the political and historical context for understanding Neurath's anti-absolutism. But perhaps more importantly, much of the semantics debate also occurred in a context of intense personal acrimony. After Carnap asked that his name be removed as editor from Neurath's *Foundations of the*

²¹ See Coffa (1976), Coffa (1987), Creath (1990a), Creath (1999), Ricketts (1996), and Awodey (2007)). For Neurath's reception of Logical Syntax, see Uebel (2009b).

Social Sciences, Neurath was deeply hurt, and his bitterness from two decades of sleights (real or imagined) and his feelings of underappreciation bubbled to the surface.²² Additionally, Neurath was concerned with what he saw as the increasingly formalistic and metaphysical direction of the Logical Empiricist movement.²³ The subsequent correspondence is laced with barely suppressed anger from Neurath, and frustration from Carnap. Their friendship suffered, so much so that Carnap's wife felt it necessary to write to Neurath, trying to heal the 'sore feelings on all sides' (Ina Carnap, 24/08/1945, in Cat and Tuboly 2019, 649). Their dispute by the mid-40s was as much personal as theoretical. And this was only exacerbated by the enforced impersonality of communication by letter. Sadly, Neurath's death prevented any reconciliation, and left the debate unresolved. Very little changed between the beginnings of the debate in 1935 and Neurath's death in 1945. Ultimately then, an understanding of the semantics debate requires more than simply philosophical analysis, but an understanding of the personal history of Neurath and Carnap.

6.5.4 *Mormann's Defence of Neurath*

What has not yet been addressed in the above interpretation is the most substantive attempt to demonstrate a theoretical basis for Neurath's rejection of semantics. Mormann argues for an interpretation according to which there is in fact a substantive theoretical basis for Neurath's opposition to semantics. On this account, Carnap and Neurath's disagreement is rooted in their differing conceptions of language. In demonstrating the struggles of this interpretation, the plausibility of the above interpretation is reinforced.

According to Mormann, there are two key aspects that are essential to understanding Neurath's opposition to semantics. The first is Neurath's anti-Cartesianism, according to which scientific knowledge is always provisional, resting on incomplete premises and expressed in an unstable and language of unclear concepts. There is no 'final knowledge based on clear and distinct ideas' (Mormann 1996, 90). The second key idea is the distinction between language as universal medium (LUM) and language as calculus (LAC). This distinction is Hintikka's, itself a development of Van Heijenoorts distinction between logic as language and logic as calculus (Van Heijenoort 1967). Hintikka distinguishes these two ways of understanding the relationship of logic and language to reality. On the LUM view 'you are a prisoner (as it were) of your language. You cannot step outside it, you cannot re-interpret it in a large scale, and you cannot even express its semantics in the language' (Hintikka 1988, 1). There is only one language and we're stuck in it. This assumption, he claims, is made in some form by 'Frege, early Russell, Wittgenstein and the Vienna Circle in its heyday' as well as Quine (Hintikka 1990, 160). On such a view truth

²² For details, see Reisch (2003) and Tuboly (2019, 108–109).

²³ See Tuboly (2017) and Reisch (2005).

ends up ineffable, and language is not open to interpretation, it can only speak about the actual world. For LUM adherents semantics is ruled out, and with it the concept of truth, 'for truth pertains to the relations of our sentences to the world' (Mormann 1999, 170). By contrast, LAC takes language to be reinterpretable like a calculus, and assumes that we can talk about language 'from an external vantage point', typically a meta-language (Hintikka 1990, 160). For LAC, the tell-tale characteristic is that 'there are many languages the semantics of which can freely be chosen. It is easy to construct new languages' (Mormann 1999, 171).

When viewed in terms of these two key ideas, Mormann claims, Neurath's opposition to Carnap's adoption of Tarskian semantics becomes recognisable as both principled and correct. According to Mormann, the Vienna Circle all initially adhered to an LUM view. Whilst Neurath developed an anti-Cartesian form of LUM, Carnap's was Cartesian. In 1935, with his adoption of Tarskian semantics, Carnap makes a fundamental change from an LUM to an LAC view, but persisting in his Cartesianism (Mormann 1999, 171). After all, 'the L[A]C-faction belongs to the Cartesian tradition: if a language is a calculus its terms have to correspond to clear and distinct ideas' (Mormann 1996, 94). Neurath recognised this change, and rightly criticised it. Carnap's adoption of a clearly Cartesian language was incompatible with physicalism, as Neurath tried to show him. Initially at least, this characterisation seems somewhat plausible. For Neurath there is only one language, the physicalist (Mormann 1999, 170). And the LUM view that we are stuck inside our language seems to capture Neurath's anti-Wittgenstein position that 'we cannot, so to speak, take up a position outside the making of statements' (Neurath 1931c/1983, 61). Similarly, Carnap's emphasis on the interpretability of language and his use of meta-languages seems to place him squarely in the LAC camp.

Uebel argues against Mormann, that even given Mormann's picture, there are still good reasons to see a compatibility between Neurath and Carnap (Uebel 2007c, 431–32). Uebel argues that what we really see is a difference of emphasis, that 'Carnap's thought centred on the conception of logic as calculus, Neurath's on that of natural language as a universal medium' (Uebel 2007d, 432). But equally, Neurath's acceptance of Carnap's syntax suggests he was also a calculist about logic, whilst Carnap's disinterest in natural language leaves no good reason to ascribe natural language calculism to Carnap (Uebel 2007d, 431–432). What we have therefore is not the dichotomy Mormann suggests, but a difference of emphasis, Neurath towards natural and Carnap toward formal languages. Here, Uebel rightly raises the point that Mormann's approach fails to separate out differences in approach to natural and formal language, which is a significant oversight. But even so, Uebel is still too generous to Mormann.

Firstly, Mormann's categorisations are problematic. Whilst Neurath seemingly meets Mormann's criteria, it is far less clear that Carnap does. For one thing, a transition from an LUM to an LAC view doesn't do justice to Carnap's developing views on logic. The crucial move for Carnap is the adoption of the principle of tolerance and logical pluralism, which is less a change in conception of logic than

a rejection of the singularity of logic. But Carnap's categorisation as a Cartesian is even more dubious. Mormann gives three commitments of Cartesianism:

- (1) Infallibilist foundationalism: Starting from indubitable true statements an all-embracing final system of philosophical and scientific knowledge can be erected.
- (2) Deductivism: Science is a deductive system of general statements ensured by proofs.
- (3) Transparentism: The system of (scientific and philosophical) knowledge can be expressed in a language whose concepts refer to clear and distinct ideas. (Mormann 1999, 166)

Carnap in his mature work, is clearly not an infallibilist foundationalist, as has been shown clearly in previous chapters. As for (2) and (3), as Carus says these are rejected by Carnap regarding the currently existing language of science, yet he retains them as aspirations (Carus 2019, 354). Whilst Carnap thought the goal of a deductive system of science was worth pursuing, he not only recognised it as currently unrealised, but understood it as potentially realisable in a variety of ways.²⁴ As Uebel argued above, the issue here is Mormann's failure to distinguish between approaches to natural and formal language. Carnap aimed for clear concepts in his constructed formal languages, but never denied the presence of vague concepts in natural language. But then why does Mormann accuse Carnap of Cartesianism? It is unclear. Mormann seems to greatly misunderstand or misrepresent Carnap's post-protocol sentence debate work.²⁵

More fundamentally, the LAC/LUM distinction simply is not effective as Mormann deploys it. As a heuristic device the distinction is useful. Hintikka introduces it as a 'contrast between ways of looking at the relation of language and its logic to reality' (Hintikka 1988, 1). He only intends to highlight two different traditions or approaches. But to diagnose the existence of a theoretical rift, as Mormann intends, the distinction has to be more than that. It has to be a sharp dichotomy between opposed camps. And the LAC/LUM distinction simply does not live up to this requirement. There are an abundance of cases that don't neatly fit on one side or the other of this division. Quine, rejects being placed in either camp (Quine 1990b, 176). Nobody can decide where Tarski falls; Niiniluoto locates Tarski's 1933 semantics between the LUM and LAC views (Niiniluoto 1999, 18), while Hintikka says Tarski's place in the dichotomy is 'ambivalent' (Hintikka 1990, 160). But dichotomies don't have middle ground. And if we cannot interpret LAC and LUM as dichotomous, then the different emphases exhibited by Carnap and Neurath are no longer problematic. After all, Carnap and Neurath are in no more significant disagreement with each other than Quine and Tarski are with themselves.

The same is true of Mormann's Cartesian/Anti-Cartesian dichotomy. He again suggests it can do a lot:

the evolving disagreements of the Vienna circle philosophers concerning physicalism, the form of protocol sentences, and the status of semantics may be interpreted as results of the ongoing dissociation of Cartesian and anti-Cartesian currents (Mormann 1999, 171)

²⁴ For more see Carus (2019, 354–55).

²⁵ As Carus notes, this is odd given Mormann's previous work on Carnap (Carus 2019, 354).

But by itself this distinction explains none of these things. Again, perhaps this distinction is useful as a general heuristic or means of categorisation, but it isn't strong enough to function as the backbone of an argument. Even if we accept Mormann's story, and Neurath and Carnap's respective categorisations as anti-Cartesian LUM and Cartesian LAC, this doesn't explain why 1935 was a breaking point for Neurath and Carnap. If Carnap's LUM view had been Cartesian too, why was Carnap's adoption of LAC in particular the cause of an argument that would last the next ten years? What made Cartesian LAC so much less tolerable for Neurath than Cartesian LUM? Presumably, the answer is Neurath's opposition to truth, theories of truth being impossible on an LUM view but not on LAC views. But then Cartesianism vs. non-Cartesianism or LAC vs LUM has no explanatory role in this story. Rather, the explanation is given by the account provided above of Neurath and Carnap's complicated attitudes to truth-talk.

6.6 Conceptions of Unity

The two most obvious interpretations of the unity of science are the ontological and nomological interpretations. Ontological unity is the claim that the entities (and therefore the theories) of the less fundamental sciences are ultimately reducible to those of the most basic science which is (almost always) physics. Nomological unity is the claim that the laws of all sciences are ultimately reducible to the laws of the most basic science. A famous example that encapsulates both the ontological and nomological senses of unity comes from Oppenheim and Putnam, who attempt to make precise how 'all science may one day be reduced to micro-physics' (Oppenheim and Putnam 1958, 27). They present a six-level hierarchy of entities, to each of which there corresponds a science. Each science, and its respective objects, are reducible to the level below via the deduction of its laws (Oppenheim and Putnam 1958, 27). The resulting hierarchy:

Level	Science	Object
6.	Sociology	Social-groups
5.	Psychology	Living things
4.	Biology	Cells
3.	Chemistry	Molecules
2.	Physics	Atoms
1.	Micro-physics	Particles

In this way, all of science is ultimately reducible to the laws and entities of micro-physics. Micro-physics thereby becomes the universal scientific theory, the theory from which all other science is derived, and upon which all theories and predictions

are ultimately based. Micro-physics therefore subsumes and renders redundant the other sub-disciplines of science.

6.6.1 *Linguistic Unity of Science*

Both Neurath and Carnap reject the ontological understanding of reduction. It is widely recognised that Carnap and Neurath instead advocated a linguistic conception of the unity of science. But this consensus masks an ambiguity and potential confusion about exactly what this entails. Firstly, the linguistic conception grew out of a more basic conception of unity, what Uebel calls ‘methodological unity’ (Uebel 2007d, 16). This is the claim that all of science is of the same kind and practiced according to the same methodology.

This thesis must be understood primarily as a rejection of the prevailing view in German contemporary philosophy that there is a fundamental difference between the natural sciences and the *Geisteswissenschaften*. . . literally “spiritual sciences” (Carnap 1963d, 52)

Defenders of this distinction argued that natural science with its empirical methodology was of a different kind to the intuition guided, a priori practice of the spiritual sciences. The unity of science thesis began as a rejection of this division, but out of this initial opposition a more substantive conception of unity developed.

The linguistic thesis of unity claims that all scientific knowledge is expressible in a shared physicalistic language, the universal language of unified science. This universal language is capable of simultaneously expressing statements from each of the sub-fields of unified science and combining them. Initially, both Carnap and Neurath refer to the physical language and the language of physics as the universal language. Frustratingly, “physical language” and “language of physics” were initially used interchangeably, which makes it unclear whether they considered the specific technical language of physics to be the universal language. In Neurath’s case, this seems to have been clumsy use of terminology. He says for example, ‘In a sense unified science is physics in its largest aspect, a tissue of laws expressing space-time linkages’ (Neurath 1931a/1983, 49). “In a sense” is doing some heavy lifting here. What Neurath describes is simply a physicalist, spatio-temporal language i.e. the universal jargon. But Carnap does seem to have taken the technical language of physics as the universal language: ‘We wish however to interpret the term “physical language” so widely as to include not the special linguistic forms of the present merely but also such linguistic forms as physics may use in any future stage’ (Carnap 1932c/1995, 54). If not the language of our current physics, perhaps the language of some future physics.

Their mature attitudes however are clearer, and there neither Carnap nor Neurath take physics to provide the universal language. Due to the unclarity of their earlier work however, the misunderstanding still occurs.²⁶ Neurath maintains that

²⁶ See for example Glock (2008, 38) and Bunge (2011, 148).

the universal language is the previously introduced universal slang. As Neurath frequently emphasised, unified science does not require a unique technical jargon, ‘we can use the everyday language which we use when we talk of cows and calves throughout our empiricist discussions’ (Neurath 1946a/1983, 233). Making a change with his earlier views, Carnap comes to a similar position, concluding that ‘what [he] really had in mind’ with his earlier accounts of the physical language was in fact the thing-language (Carnap 1936–37, 467).²⁷ It is the thing-language that is the universal language for the mature Carnap. He characterises the thing-language as ‘that language which we use in every-day life in speaking about the perceptible things surrounding us’ (Carnap 1936–37, 466). Carnap and Neurath both reached an understanding of the universal language as a physicalist refinement of everyday language.

What may initially sound like a grand ontological reductionist vision of the nature and structure of science is in actuality a measured position on the requirements necessary for communication and collaboration between different sub-disciplines of science, via a common medium-sized spatio-temporal thing language. Unity is not merely a doctrine about how language is, but a proposal for what language must be like for science to be practiced successfully. For Neurath and Carnap, unity was not simply a truth about the language of science, but a proposal to be enacted. Further unification of the sciences was a goal to be pursued.

Whether or not the specifics of their conceptions are compatible has however been a matter of contention. Whilst acknowledging that both provide linguistic conceptions of unity, Creath claims ‘their conceptions... are so very different that it hardly makes sense to speak of the unity of science as if it were some one doctrine’ (Creath 1996, 166). Uebel has opposed Carnap’s ‘hierarchy of reductively related theories’ to Neurath’s ‘looser’ notion of unity (Uebel 2001, 213–214).²⁸ Frost-Arnold however has argued that Uebel’s view (and by extension Creath’s) relies on a caricatured reading of Carnap (Frost-Arnold 2013, 118–119 n.4). Specifically, Frost-Arnold rejects the claim that Carnap’s conception of unity involves theoretical reduction of theories. To resolve this dispute, we first need to look at what Neurath and Carnap say about unity in more detail.

6.6.2 *Carnap and Neurath on the Unity of Science*

In Carnap’s mature work, unity is ‘a problem of the logic of science, not of ontology... concerning the logical relationships between the terms and laws of the various branches of science’ (Carnap 1938/1991, 397). This means establishing relations of reduction. And there are two components, terms and laws, each of

²⁷ This perhaps suggests that he never did take the technical language of physics as the universal language. But we will not dwell on this small interpretive point.

²⁸ Reisch has similarly read Neurath as opposing Carnap (Reisch 1997, 446–447).

which we shall deal with in turn. For terms, the unity thesis means that the physical thing-language is ‘a common reduction basis for the terms of all branches of science’ (Carnap 1938/1991, 404).²⁹ Prima facie, this sounds like an ontological project comparable to Putnam and Oppenheim’s. But Carnap’s notion of reduction is not reduction as typically understood by philosophers.³⁰ “Reduction” intuitively suggests the elimination and replacement of certain elements with a smaller group of elements, or establishing one element as constituted of another. But neither of these typical uses captures Carnap’s. Carnap’s reductions involve reduction sentences, sentences which introduce predicates by establishing their conditions of application through specification of their logical relationships to other pre-existing predicates (Carnap 1936–37, 441). They are not exclusive, one concept may have multiple reduction bases, and ‘mutual reducibility’, that set x is reducible to set y and vice versa, is possible (Carnap 1938/1991, 397). ‘A *definition* is the simplest form of a reduction statement’ (Carnap 1938/1991, 397). What the unity thesis says then is that there is a set of predicates which are sufficient for establishing the application conditions of all the terms of all the sciences, and this sufficient set is the physical thing-language. Remember that Carnap’s thing-language is a physicalistic language composed only of observable predicates. So Carnap’s unity thesis boils down to the claim that the entire language of unified science can be related back to observable thing-predicates. Scientific terms need not be reduced directly to observable thing-predicates, the process of reduction can be iterated through other further reducible terms. When Carnap’s notion of reduction is recognised for what it is, it becomes clear that Carnap’s conception of the unity of science is fundamentally a logical one. And importantly reducibility does not entail theoretical redundancy. Carnap is not advocating abandoning the terminology of science in favour of exclusive use of the physical thing-language. Additionally, in 1939 Carnap further loosens the conditions on purely theoretical terms defined by implicit definition, requiring only that some so defined term be related back to an observational basis via reduction sentences Carnap (1939/1971).

Carnap also makes clear that establishing the reducibility of terms does not entail the reducibility of laws (Carnap 1938/1991, 403). Carnap recognises that at ‘the present state of the development of science’ there is no nomological reducibility, but argues that ‘no scientific reason is known for the assumption that such a derivation should be in principle and forever impossible’ (Carnap 1938/1991, 403). He maintains that the attempt to demonstrate the unity of laws is potentially fruitful, and worth pursuing even though we ‘do not, of course, know whether it will ever be reached’ (Carnap 1938/1991, 403). This honest admission of ignorance, and a refusal to narrow our options without sufficient evidence seems both moderate and sensible. His position on nomological unity could be characterised as an optimistic agnosticism.

²⁹ Carnap prefers “term” to “concept” because it emphasises the linguistic rather over the psychological (Carnap 1938/1991, 397).

³⁰ One may question why he chose the term “reduction”, but since he did, that’s the terminology we must use.

Carnap considers the unity of science thesis to be a logical one, but for Neurath it is ‘a *historical fact* in a sociological sense’ (Neurath 1935a/1983, 115). The possibility of ‘collective work and communication’ between different branches of science is proof of the unity of science (Neurath 1935a/1983, 115). Let’s return to Neurath’s example of the forest fire:

In order to formulate the individual prediction: “This forest fire will soon be extinguished”, we combine biological statements (concerning tress, etc.), chemical statements (concerning fire, etc.), sociological statements (concerning fire service, etc. and statements of other disciplines (Neurath 1936a/1983, 132)

That we can make such cross-connections in practice is sufficient proof of the thesis. But we cannot simply rest on our laurels; ‘unification and systematization of science are permanent activities’ (Neurath 1937c/1983, 174). Unification is not new, but it can be facilitated and furthered through our practice, through the creation of norms and institutions. As we saw, this is one of the key motivations behind Neurath’s creation of the *Encyclopedia of Unified Science*.

Notably absent from Neurath’s discussion is any reference to reduction, whether of laws, theories or terms. Neurath does make reference to scientific laws, but with practical concerns in mind, that ‘[u]nder certain circumstances it must be possible to link the laws of all sciences with each other to make *one* definite prediction’ (Neurath 1931b/1983, 54). Rather than establishing any basis for reduction, he advocates ‘treat[ing] all statements and all sciences as coordinated and to abandon for good the traditional hierachy’ (Neurath 1944/1971, 8). Neurath presents his encyclopedism as an alternative to Pyramidism, ‘which intends to build a symmetrical and complete edifice of the sciences’ (Neurath 1937b/1983, 203). By contrast, his encyclopedism establishes only a ‘*mosaic pattern*’ of horizontal integration (Neurath 1937b/1983, 204). Strictly defined logical relationships play no role in Neurath’s characterisation of unity. We only need a taxonomy of science as ‘a rough bibliographic order’ (Neurath 1937b/1983, 204).

6.6.3 *Different Conceptions or Difference of Emphasis?*

It’s easy to see how these two positions can be interpreted as distinct and even opposed. Formal hierarchies of logical reduction are not the same as loose horizontal integrations for practical purposes. Uebel is certainly right to emphasise Neurath’s opposition to rigid reductionism. And, pace Frost-Arnold, Uebel is right that, in his earlier work at least, Carnap exhibits signs of the Pyramidism Neurath rejected. Carnap’s ‘genealogical tree of concepts’, where sociological concepts are reducible via the psychological and the physical, to the given is strikingly similar to the hierarchy of reduction given by Putnam and Oppenheim (Carnap 1930/1959, 144). But, I also think Carnap’s mature conception of reduction moves away from this hierarchical view, towards a position that Neurath not only could, but did, embrace. In the same article as his critique of Pyramidism, Neurath acknowledges

the importance of the ‘question [of] how to reduce statements and laws of all the sciences to a certain type of statement and law’ (Neurath 1937b/1983, 203). Carnap isn’t named, but is clearly invoked. Elsewhere, Neurath explicitly talks approvingly about Carnap’s project:

the unification of scientific language is a special and technical task... Carnap has proposed certain special devices, which make possible the introduction of terms so that they become dependent upon physical terms without, however, being ‘defined’ by the latter in the strict sense (Neurath 1937c/1983, 176)

If Carnap’s reductionism were an example of the Pyramidism he was critiquing, Neurath would not have endorsed it.

Uebel interprets Neurath as highly sceptical of nomological reduction (Uebel 2007b, 261). Uebel is right that ‘the encyclopedic unity envisaged by Neurath did not require a reductive unity of laws’ (Uebel 2007b, 264). And he is right that Neurath stresses the autonomy of sociological laws:

The development of physicalist sociology does not mean the transfer of laws of physics to living things and their groups, as some have thought possible. Comprehensive sociological laws can be found... without the need... to build up these sociological laws from physical ones (Neurath 1931c/1983, 75)

But all this being accepted, Neurath never claims that nomological reduction is impossible. In his correspondence with Dewey, Neurath contends that the possibility of reduction cannot be dismissed as ‘doomed in advance to defeat’ (Quoted in Reisch 2005, 89). What he rejects are pre-conceptions concerning which laws will reduce to which others as ‘a premature presumption’ (Neurath 1937b/1983, 202). But Carnap does not make such presumptions. And nor does Carnap suggest that establishing nomological unity would render the reducible laws redundant. Assuming that the unity of laws is established in the same way as the unity of terms, any reduction of sociological laws to those of physics would not render them redundant or dependent. It would simply establish clear logical connections. One could perhaps accuse Carnap of retaining something of the spirit of the traditional hierarchy, a lingering intuition that psychology reduces via biology down to physics, but this presumption is not built into Carnap’s thesis of unity. Both Neurath and Carnap are essentially agnostic about the possibility of nomological unity. The difference between them is simply in their attitudes. Whilst Carnap is optimistic and enthusiastic about the possibility of nomological reduction, Neurath is more sceptical and disinterested. Neurath’s scepticism may prove to be justified, or Carnap’s optimism may be vindicated, but either outcome is compatible with both Neurath and Carnap’s accounts of the unity thesis. There is no substantive dispute here.

And whilst Carnap’s focus is the logico-linguistic, he retains the same sociological understanding of unity as Neurath. As discussed earlier, his example of automobile construction closely parallels Neurath’s forest fire (Carnap 1938/1991, 404). Carnap sees the logico-linguistic as underpinning the practical; ‘the unity of the language of science is the basis for the practical application of theoretical knowledge’ (Carnap 1938/1991, 404). Carnap does not provide his logical reductions as

an alternative to the account of unity in practice. Rather he sees his logical work as supplemental to the conception of unity provided by Neurath.

Both Carnap and Neurath accept linguistic unity, but this classification by itself clearly doesn't sufficiently explain their individual approaches to the unity of science. I want to suggest a distinction between practical-linguistic and logico-linguistic senses of unity. The former concerns the possibility of utilising the language of science for collaborative coordination towards a specific goal, the latter concerns deductive relationships between the terms and laws in the language of science. These two senses of unity approach the laws and terms of science from different angles. For Neurath's practical-linguistic approach, statements are treated pragmatically, as generators of expectations for the purposes of predicting and controlling worldly phenomena. For Carnap's logico-linguistic approach, they are treated as statements of a calculus, whose rules for translation, derivation and deduction need to be clearly established. What is most important to note here is that neither sense opposes or contradicts the other. All statements possess both logical and practical aspects. Again, we see the bipartite metatheory at play: the logico-linguistic approach treats the language of science, whilst the practical-linguistic deals with the practice of science. Carnap and Neurath simply focused on the facet of language relevant to their respective part of the metatheory. The terminological reduction that Carnap provides does not establish the rigid hierarchy of theories and entities of the Pyramidist. Carnap's reduction does not contradict 'the fact that the vast mass of groups of statements are, as it were, *in one plane*' (Neurath 1937b/1983, 204). And importantly, they each recognise and acknowledge the significance of the other's area of emphasis. Creath is therefore right that the unity of science is not a single doctrine, but not because Carnap and Neurath provide incompatible accounts. Rather, "unity of science" is an umbrella term for a variety of inter-connected theses about the language of science and its application to the world, all of which cohere with one another.

6.7 Conclusions

The above should be sufficient to demonstrate that there is a fundamental philosophical agreement between Neurath and Carnap which underpins the Bipartite Metatheory of science. Things which initially appear to be theoretical ruptures between them are, upon closer inspection, simply variations in terminological usage or differences of emphasis resulting from the division of labour necessitated by the metatheory. The only exception to this is the case of truth and semantics, where there is some substantive disagreement. But as we have seen this is easily and organically overcome. We are now in position, in the final chapter, to explore the broader purposes of this metatheory.

Chapter 7

Bipartite Metatheory in Application



The previous chapter demonstrated the compatibility of Neurath and Carnap's projects, but only in the negative sense of eliminating lingering worries about potential theoretical incompatibilities between them. In what follows, I will show how the two halves of the metatheory are relevant to one another, and the consequences of this collaboration.

7.1 The Bipartite Dialectic

The Bipartite Metatheory envisions a metatheoretical approach to the study of science in which the two parts are in constant contact and cooperation with one another. The question for defenders of the bipartite metatheory interpretation is how Carnap's explicationist methodology for the logic of science interacts with Neurath's empirical pragmatics of science. Carus describes Carnap's explication project, as an 'implicitly dialectical conception' (Carus 2007a, 20). I want to suggest that this also true of the relationship between the two parts of the metatheory. The aim is to provide an explicitly dialectical conception of the bipartite metatheory as a discourse between formal and informal methods, between logicians and methodologists, in which dialogue creates mutual feedback, both epistemologically and normatively, with both logician and methodologist taking a constructive role in the creation of methodological and epistemological norms.

Carus provides an account of the methodology of explication applicable to Carnap's purposes, and consistent with a broader naturalistic epistemology. He argues that the method of explication is intended to increase the degree of constructedness in the language we use. Importantly, the distinction between natural and artificial language is one of degree, not a sharp dichotomy.

'the conception of a sharp separation, perhaps even a gap, between everyday concepts and scientific concepts. I see here no sharp boundary line but a continuous transition. The

process of the acquisition of knowledge begins with common sense knowledge; gradually the methods become more refined and systematic, and thus more scientific' (Carnap 1963b, 934)

Through explication, and increasing the constructedness of our language, the clarity and utility of that language is slowly increased. Carus describes the relation between constructed and natural languages, as 'piecemeal exchange within the context of a dynamic of a mutual feedback relation' (Carus 2007a, 278). This description also captures the nature of the relationship between the two aspects of the metatheory, between the logic of science and the pragmatics of science. Carus's picture of Carnap is not wrong, but can be supplemented. Carus understands the purpose of Carnap's method of explication, but potentially overstates its significance as an all-encompassing project in and of itself. Within the logic of science, explicationism is a crucial part of both analysis and the generation of proposals. But, what has to be emphasised is that the formal aspect of explicationism alone cannot fulfil the role of the metatheoretical project, because it needs external input lest it float disconnectedly. Now Carus would likely agree, and may argue that this external element is incorporated into the explicationist methodology. But I suggest that the provision of such information is the domain of the pragmatics of science, and that it is here that Neurath's empirical project plays its role.

To see this, we have to remember that explication is an ongoing and iterative process of conceptual refinement. The end of one explication may simply be the basis for another. What was once the explicatum may in future be the explicandum. Attempts to introduce the explicandum may reveal that further clarification of the explicatum is necessary. Alternatively, disambiguating the explicatum may be pursued via tentative attempts at the introduction of an explicandum and reflection on their failures and successes. Importantly, we should not conceive the process as one of rigid beginning to end movement. It is the potential for iteration, revision and reversal that gives the method of explication such broad scope and constructive power. However, in isolation, explication is powerless. As we saw before, Carnap's logical work is necessarily connected to the outside, to non-logical factors. Remember, Carnap sees his role as the provision of tools. The iterative process of refining these tools requires knowledge of their intended purpose, the context in which they are to be used, the methods with which these tools are applied, and finally evaluations of their success or failure in practice. These evaluations are not possible via a formal process, but must be provided by empirical study. What is needed is the sort of information provided by Neurath's pragmatics of science: historical, psychological, sociological and methodological studies of scientific practice. Without this empirical data to work from, the explicationist logician is flying blind.

On this point, Carnap himself was clear:

'A proposal for the syntactical formulation of the language of science is, when seen as a principle, a proposal for a freely choosable convention; but what induces us to prefer certain forms of language to others is the recourse to the empirical material which scientific investigation formulates... From this it follows that the task of the philosophy of science can only be pursued in a close cooperation between logicians and empirical investigators' (Carnap 1934c, 8)

The logic of science generates proposals for concepts (or even languages) to be utilised by science. The appraisal of the success of these tools, relative to the goals of science, is facilitated by empirical study of scientific methods, and the sociology and psychology of science. An understanding of how science has operated historically and how scientists behave allows the evaluation of the proposals provided by the logician. From such studies, we learn which methods or concepts are most likely to provide reliable predictions and interventions. This provides the basis for the appraisal of the proposals provided by the logic of science. Only in conjunction do the two parts of the metatheory fulfil their full function.

An example of exactly this process is present in Neurath's proposed protocol sentence schema, as described in Chap. 3 and Sect. 6.3. The schema, as we saw, is a reconstruction of the concept of scientific evidence. The constraints on this proposal were given by Neurath's understanding of the epistemology of science (anti-foundationalism, holism) but also by his empirical studies of the behaviour of scientists, specifically in their response to empirical falsification. As we saw, reflection on scientific practice demonstrated that contradictory evidence was rarely (if ever) sufficient for the rejection of a theory (as in Neurath's Ehrenhaft example). As a result of these empirical studies, the role of decision-making was integrated into Neurath's proposal for protocol sentences, this being an explication of the notion of scientific evidence.

But this relationship is not unidirectional. The logician of science, given the principle of tolerance, is free to generate whatever proposals they can, and in so doing it is possible that they provide proposals with widescale consequences that had not been considered before. The logician is free to pursue new ideas. In so doing, the logician may allow for a reappraisal of those very criteria by which its output is judged; the goals and methods of science. The new possibilities revealed by logical explorations may allow for the re-evaluation of the goals and the methods of science.

In his final use of the ship metaphor, Neurath envisions the future development of scientific methods: 'The whole business will go on in a way we cannot even anticipate today' (Neurath 1944/1971). As scientific theories develop, they shape the norms and methods of science, and these in turn shape subsequent theories. Not only is there a process of mutual feedback between the two halves of the metatheory, but this leads to a further feedback process between the theoretical and the methodological. Importantly though, the consequences of the interaction of logic and pragmatics of science is never definitive. The provision of a new concept never necessitates a specific revision of either theory or method. Just as in the case of scientific decision-making, in metatheoretical decision-making there is no calculus to follow. Whether a new concept is adopted, and how the consequences of this adoption are to be absorbed are both a matter of decision, to be taken by the scientific community. Once again, conventionalism and decisionism are pervasive. The bipartite dialectic itself embodies the epistemological consequences of rejecting pseudo-rationalism.

This organic process, governed by mediation and negotiation rather than strict rules, has parallels to the process of reflective equilibrium.¹ Nelson Goodman describes how:

'the rules and particular inferences alike are justified by being brought into agreement with each other. *A rule is amended if it yields an inference we are unwilling to accept; an inference is rejected if it violates a rule we are unwilling to amend.* The process of justification is the delicate one of making mutual adjustments between rules and accepted inferences; and in the agreement achieved lies the only justification needed for either' (Goodman 1983, 64)

This gradual process captures what was recognised in principle in Chap. 3, that Neurath's boat must be recognised as applying to the methodological aspects of science as well as the theoretical. The gradual, piecemeal approach is exemplified in Neurath's first description of the boat in 1913:

'The new ship emerges from the old through a process of continuous transformation' (Neurath, Quoted in Cartwright et al. 1996, 130–131)

And it recurs in the second use of the boat in 1921:

'the ship can be shaped entirely anew, but only by gradual reconstruction' (Neurath 1921/1973, 199)

This picture of the bipartite dialectic as a process of slow mutual readjustment is a fulfilment of Neurath's boat metaphor for naturalised epistemology.

One further consequence of this discussion is that a decisive separation of the metatheory into two discrete and discontinuous halves is not possible. We cannot cleanly separate the works of pragmatics from the works of logic. Historical and sociological studies will often require some conceptual analysis, whilst studies of language and logic will often make reference to real world examples. The division of labour between logic and pragmatics needs to be seen as a sliding scale, rather than a strict dichotomy. Carnap seems to have been aware of this, noting that 'no strict classification of the investigations themselves and the treatises in which they are set forth is possible' (Carnap 1934d/2001, 331). This is by no means a rejection of the Bipartite Metatheory interpretation, so much as a slight clarification of the exact way that the bifurcation of metatheoretical work takes place in practice. Whilst the clean separation is coherent in theory, any example involves an inevitable mingling of the two.

Stein has suggested that Carnap's explicationist methodology entails what he calls a 'blurring', but is maybe more accurately characterised as simply an interaction and exchange, 'between the purely cognitive, or theoretical, and the practical' (Stein 1992, 291). Carus calls it 'an interplay or mutual feedback relation' (Carus 2017, 173). It's in this dialectical interaction between the logic and pragmatics of science that the full consequences of methodological constructivism can be cashed out.

¹ For more on the connection between Carnap's explication and Goodman's reflective equilibrium, see Brun (2017).

7.2 Decisionism, Conventionalism and Constructivism

7.2.1 *The Pragmatic and the Formal*

What is absolutely crucial to the relationship between the two halves of the metatheory is the relationship between the practical and the conceptual, a point emphasised repeatedly by both Carnap and Neurath:

‘Our thesis that the logic of science is syntax must therefore not be misunderstood to mean that the task of the logic of science could be carried out independently of empirical science and without regard to its empirical results... the language of science is not given to us in a syntactically established form; whoever desires to investigate it must accordingly take into consideration the language which is used in practice in the special sciences, and only lay down rules on the basis of this... a convention can only be useful and productive in practice if it has regard to the available empirical findings of scientific investigation’ (Carnap 1934d/2001, 332)

This understanding of the practicality of all theoretical work is present as early as the manifesto; ‘we have to fashion intellectual tools for everyday life, for the daily life of the scholar but also for the daily life of all those who in some way join in working at the conscious re-shaping of life’ (Carnap et al. 1929/1973, 305). These technical tasks are of relevance both theoretically and practically. And practical factors play a role in determining the purposes toward which technical work is oriented. We see this expressed by other members of the left-wing. In a discussion of logicism, Hahn agrees with ‘Frege and Russell, [who] have indicated how numbers can be constructed’, on the grounds that ‘they do, in fact, do all the work that numbers in daily life and in science are supposed to do, and this is all we need’ (Hahn 1930b/1980, 15). Here again, we see formal projects intimately connected to practical purposes.

But what exactly is this connection between conceptual and practical? Here, the metaphor of the tool is of central significance:

‘In my view, a language, whether natural or artificial, is an instrument that may be replaced or modified according to our needs... A natural language is like a crude, primitive pocketknife, very useful for a hundred different purposes. But for certain specific purposes, special tools are more efficient’ (Carnap 1963b, 938)

This conception of languages as tools is an expression of Carnap’s conventionalism, itself a consequence of the principle of tolerance, and the instrumentalism of language that results from it. The principle of tolerance amounts to a rejection of logical monism (that there is one correct logic), in favour of logical pluralism (that we are free to use any system of logic for which we can specify the rules). The choice between the plurality of available logics is an external question, made on the basis of pragmatic factors and facilitated by an analysis of the consequences of adopting the system. There is no single correct choice. A proof that one logic ought to be adopted is impossible, since such a proof presupposes a consequence relation, and therefore a system of logic (Ricketts 1994, 183). The choice between logics is therefore epistemically arbitrary. But practically, the choice is far from arbitrary.

These choices are always concrete and contextual, made on the basis of our current goals.

Carnap's conventionalism here has clear parallels to Neurath's decisionism, as discussed in Chap. 2. There, we saw that Neurath recognised the social embeddedness of scientific theory choice, and the role of factors external to science. His critique of pseudo-rationalism began from a criticism of Descartes's attempt to separate the theoretical from the practical. Against this, he argued that thought and action are always intertwined, the theoretical and practical cannot be separated from one another. This unity of theoretical and practical also animated his *Encyclopedia*, which he intended to 'express the situation of a living being and not of a phantom', science 'as a being of flesh and blood' (Neurath 1938/1971, 25–26).

Neurath and Carnap's convergence on these positions is no mere coincidence. Neurath's decisionism and Carnap's linguistic conventionalism share motivations. As Carnap noted:

'The choice of a language form is a practical decision, [Neurath] argued, just as the choice of a route for a railroad or that of a constitution for a government. He emphasized that all practical decisions are interconnected and should therefore be made from the point of view of a general goal. The decisive criterion would be how well a certain language form, or a railroad, or a constitution, could be expected to serve the community which intended to use it. His emphasis on the interdependence of all decisions, including those in theoretical fields, and his warning against isolating the deliberation of any practical question, even that of the choice of a language form, made a strong impression upon my own thinking' (Carnap 1963d, 51)

Both Carnap and Neurath recognise the inescapable link between theory and practice that is a consequence of their naturalistic outlook. They both emphasise the inescapable embeddedness and context-dependence that surrounds our decisions. And both recognise that, with the rejection of the pseudo-rationalist notion of a singular decision-calculus, broader external goals can (and should) play a crucial role in all theoretical decision-making.

Just as Neurath's epistemic agent is inherently active, so is Carnap's linguistic agent. In much the same way as the Neurathian scientist constructs the norms of science, the Carnapian linguist actively constructs the norms of his proposed language.² This is brought out best in comparison with Quine, for whom we are left 'acquiescing in our mother tongue' (Quine 1968/1969, 49).³ For Carnap however, we play a decisive and deliberate role in the reconstruction that our language undergoes (even whilst allowing that natural languages will also undergo organic changes). At the core of his explicationist methodology is a crucial notion: the right concepts are not pre-existing entities to be uncovered through analysis, but to be constructed by us with specific purposes in mind. Carnap's view of concepts as tools for the provision of tools positions the logician as craftsman, not detective. Concepts aren't there to be found, we're here to make them.

² In a similar spirit, Carus has previously used the metaphor of Neurath's boat for explication (Carus 1999, 18; Carus 2007a, 276).

³ This comparison is made by Carus too (Carus 1999, 25).

As Creath points out, the principle of tolerance gives way to a form of linguistic conventionalism far more radical than simple underdetermination (Creath 1990a, 408). Duhemian underdetermination (which we have seen Neurath and Carnap readily embraced) posits that there will always be multiple theories which satisfy the existing body of evidence. But simple Duhemian underdetermination, Creath emphasises, takes for granted that what counts as evidence is already settled. ‘Carnap’s new conventionalism, by contrast, ultimately says that what counts as evidence and what the appropriate logical relations are (even what the logical consequence relation itself is) *are all up for grabs*’ (Creath 1990a, 408). Importantly for the bipartite metatheory interpretation, this consequence of logical pluralism is embraced by Neurath too, as is clear from his proposed protocol sentence schema. As we saw in Chap. 3, his conception of a protocol statement is a proposed reconstruction of the concept of scientific evidence. The criteria for the adequacy of his proposal are not abstract or pre-existing notions of the nature of evidence. Rather, his proposal is intended to meet the practical requirements learned from scientific practice and history: specifically, inter-subjective accessibility (via the requirement of empirical controllability), historical stability (via deliberate preservation of a protocol’s content) and practical utility (via the multiply-embedded structure). His schema is as much a proposal for what we should require of scientific evidence as much as it is an attempt to meet these criteria.

Pluralism is a natural outgrowth of recognising the underdetermination of theory choice, and therefore an inevitable outcome of the decisionism that is central to Neurath’s epistemology of science. If decisions are never unequivocally determined, there never being one singularly correct choice, then allowing a variety of intellectual positions utilising different approaches is an obvious consequence. For science, theoretical and methodological pluralism, allowing scientists to pursue different research programmes utilising differing approaches, is an obvious outcome of underdetermination. This means acknowledging the value of exploring a variety of different options. But this plurality also needs to be able to communicate and collaborate, they cannot be allowed to simply splinter off from each other. ‘It is the problem of any democracy, which any actual scientific research organization has also to solve: on the one hand the non-conformists must have sufficient support; on the other hand, scientific research needs some cooperation’ (Neurath 1946a/1983, 236). As we will see below, for Neurath epistemological pluralism is inseparable from political pluralism.⁴ But Neurath also believed that the International Encyclopedia could fulfil a unifying function as a ‘platform on which all kind[s] of

⁴ For more, see O’Neill (2003) and O’Neill and Uebel (2004).

discussions may take place' (Neurath 1946d, 82). The Encyclopedia allowed for discussion, disagreement and dissent, recognising these as inevitable features of a broader, non-monolithic collaborative process, whilst drawing people from different disciplines and countries together and encouraging inter-disciplinary collaboration.

7.2.2 *Methodological Constructivism*

As should be clear now, both Carnap and Neurath accept both underdetermination and conventionalism. We decide what constitutes evidence, and are still faced with the decisions necessitated by Duhemian underdetermination. They both recognise that convention and decision permeate unified science at all levels. Importantly, these parallels are not simply an interesting coincidence, but a demonstration of a fundamental agreement of purpose and orientation. And most significantly, they both respond to this conclusion in the same way: by embracing methodological constructivism.

I introduced the notion of methodological constructivism in Chap. 3 to describe Neurath's response to decisionism in the pragmatics of science. Methodological constructivism is not so much a thesis as it is an orientation or attitude to decisionism itself. It is the flipside of the rejection of pseudo-rationalism, the positive to its negative. In response to anti-foundationalism, underdetermination, and the resulting decisionism, Neurath reacted with optimism. Where other philosophers saw a terrifying absence of absolutes, Neurath found this liberating. Decisionism enforces a responsibility on the epistemic agent, for both the content of science and its methodology. But again, for Neurath this is a positive. Recognition of our epistemic responsibility requires honesty, transparency, self-criticism and, more radically, the democratisation of knowledge. These consequences will be discussed in more detail below. The recognition of the fact of Neurath's boat as the best model for knowledge should not bring us pessimism or nihilism, but should embolden us with the recognition of our own constructive capacity. The outcome is a metatheoretical constructivism.⁵

A parallel conclusion is reached by Carnap in the logic of science. And here, the key insight from which this constructivism grows is the principle of tolerance. The principle of tolerance gives decision-making a central role in the logic of science too. The choice of language is an external one, guided by the requirements of explicit specification and practical utility. There is no calculus that will decide the optimal language for a given context. We are left in a position necessitating decision, but without the possibility of a decisive right or wrong answer. What I think has been overlooked is that this crucial moment in Carnap's philosophical

⁵ This term has previously been used by Uebel to describe Neurath's position (Uebel 1996).

development should also be recognised as a rejection of pseudo-rationalism. To be clear, I am not attributing the Principle of Tolerance to Neurath's influence.⁶ Carnap arrived at the position independently. What I am suggesting is a convergence of Neurath and Carnap on the same meta-philosophical realisation. Whereas Neurath's rejection of pseudo-rationalism is apparent in his earliest writing, Carnap arrives at this conclusion during the process of writing *Logical Syntax*. He concluded that it is not possible to encompass logic and arithmetic in one language that could then be counted as universally basic for all others, and that in the absence of such "grounding" there is a plurality of ways of making do, even conceptually.

The principle of tolerance is a rejection of the pseudo-rationalist impulse to find *the* logic where there is no single system. It is a rejection of the absolutist, unitary assumption that had pervaded philosophy for hundreds of years, and still dominated the thinking of great logicians like Russell and Frege. This is clear in Carnap's own framing of his adoption of the Principle of Tolerance. He argues that previous attempts to move beyond the classical forms 'were hampered by the striving after "correctness". Now, however, that impediment has been overcome, and before us lies the boundless ocean of unlimited possibilities' (Carnap 1934d/2001, xv). And this freedom brings a reorientation away from capturing the pre-existing, towards constructive decision-making:

'Once the fact is realized that all the pros and cons of the Intuitionist discussions are concerned with the forms of a calculus, questions will no longer be put in the form: "What *is* this or that like?" but instead we shall ask: "How *do we wish to arrange* this or that in the language to be constructed?"... On this view the dogmatic attitude which renders so many discussions unfruitful disappears' (Carnap 1934d/2001, 46–47)

As Ricketts puts it, the principle of tolerance 'saps foundational disputes' (Ricketts 2007, 218). The "dogmatic attitude" that leads to such unfruitful philosophical discussion is pseudo-rationalism.

Like Neurath, Carnap takes the anti-absolutist rejection of pseudo-rationalism to be liberating, his conclusions are optimistic. This is an embrace of methodological constructivism in the logic of science. This attitude is perfectly captured by Jeffrey, Carnap's student, who describes Carnap's voluntarism:

'His persistent, central idea was: "It's high time we took charge of our own mental lives" - time to engineer our own conceptual scheme (language, theories) as best we can to serve our own purposes; time to take it back from tradition, time to dismiss Descartes's God as a distracting myth, time to accept the fact that there's nobody out there but us, to choose our purposes and concepts to serve those purposes, if indeed we are to choose those things and not simply suffer them' (Jeffrey 1992, 28)

⁶ Although Carnap did acknowledge a shared approach. As Awodey and Carus have shown, the first example of Carnap having adopted the Principle of Tolerance is found in "On Protocol Sentences" (Awodey and Carus 2007, 2009). In that paper, Carnap presents his approach as a rejection of absolutism, and credits Neurath as 'the first to turn decisively against this absolutism' (Carnap 1932d/1987, 469).

Carus recognises this too (Carus 2013, 265–268; Carus 2017, 180). Something of this attitude has therefore been recognised by scholars before, but the affinity with Neurath has not been. Or at least, has not been sufficiently emphasised. Carnap's key philosophical insight converges organically with Neurath's, because they are expressions of the same methodological constructivism in their respective fields. There is a fundamental agreement. This methodological constructivism involves an act of radical honesty: recognising that your decisions are exactly that, that there is no calculus to appeal to, and consequently recognising the relevance and role of pragmatic factors like goals and values in the process. In the absence of inherent pre-existing answers to be discovered and adhered to, volition is not just licenced, but required. The whole of science is in a position that parallels that of Neurath's epistemic agent. And the consequence is best expressed in the above quote from the manifesto: 'the conscious re-shaping of life' (Carnap et al. 1929/1973, 305).⁷

7.3 Self-Reflexivity and Self-Direction

7.3.1 *Self-Reflexivity and Pseudo-Rationalism*

The possibility of unified science playing the role of deliberately reshaping the world according to human goals is facilitated by another consequence of the rejection of pseudo-rationalism: perpetual self-reflection and self-criticism. One further consequence of Neurath and Carnap's fallibilism is the need for constant vigilance in self-criticism. In the absence of epistemic foundations, no part of the encyclopedia should ever be taken for granted indefinitely. Frank describes the process within science of old concepts falling into misuse, and consequently being revised and overthrown, only for these new concepts to fall into misuse later. This process 'takes place in eternal circles' (Frank 1917/1949, 78). Reluctance to allow such perpetual reinvention is the source of intellectual calcification and philosophical mystification. Frank describes this reflexive process of science's conceptual self-revision as the 'restless spirit of enlightenment that keeps science from petrifying into a new scholasticism' (Frank 1917/1949, 78). Similarly, Neurath considers how 'It will be stimulating when we new critics of our language will be criticized by means of the procedures we proposed' (Neurath 1941/1983, 217). Without this perpetual reflection and revision, the possibility of dogmatism arises. It is by preventing doubt, by ascribing authority to the traditional, by allowing ideas to remain unquestioned, that absolutism and dogmatism, both intellectual and political, is allowed to grow.

⁷ This broader social vision of the role for Unified Science was lost in post-war Anglo-American Logical Empiricism, which was more in line with Schlick's project. For the socio-historical factors in this change, see Richardson (2003), Reisch (2005, 2007), and Howard (2003, 2009).

This self-reflexivity is most explicit in the bipartite dialectic, and in Carnap and Neurath's methodological constructivism. As we have seen, the bipartite metatheory's dialectical process of mutual readjustment is a deliberately self-reflexive approach to the piecemeal improvement of science. As Neurath notes, this process cannot help but be self-reflective:

'What comes from an "experiment" with a modified scientific language will be analysed by a man who is modified by this "experiment", which is more than an experiment: it performs a kind of self-education' (Neurath 1941/1983, 92)

The central role of reflexivity has been noted by some contemporary scholars of Neurath.⁸ It is particularly emphasised by Zolo, who rightly notices the importance of reflexivity (routed in underdetermination) to Neurath's epistemology, and something of its relationship to overcoming pseudo-rationalism (Zolo 1989, Chapter 3). But what Zolo does not recognise (or at least doesn't emphasise) is the role of this reflexivity in allowing science, via the metatheory, to fulfil this purpose by and for itself. As metatheoris, the logic and pragmatics of science allow science to fulfil this moderating role for itself. By allowing the appraisal and adjustment of not only the concepts and methods, but also the norms and goals of science, unified science can both reflect on and guide itself without the need for a recourse to some external arbiter. Here, the fundamentally naturalistic nature of unified science is clearest. There is no need for traditional philosophy as the queen of sciences, outside and above scientific work in the earthly plane. Whatever role philosophy aspired to can be fulfilled by science itself. 'Whatever is asserted in science can be criticized from a more comprehensive scientific point of view' (Neurath 2011, 20).

Science is not a monolithic entity to be appraised from the outside. Neurath's naturalism denies any such position because science is not some external other. Science is a process and activity, in which we participate even whilst we engage with it. It is the means through which we are able to understand and control the world. So reflection on the process of science only occurs within and by science itself. As we saw in the previous section, the scientific metatheory is capable of reflection on all aspects of the scientific process: theory, methods, and concepts. Science is a part of the world, and as such open to scientific study as is any other earthly thing. Scientific metatheory enables us 'to talk as scientifically about science as one talks about plants, animal [sic] or humans in the special sciences themselves' (Neurath 2011, 23). The metatheory of science is just one more sub-discipline of unified science. There is then no need for an autonomous, *a priori* philosophy since the metatheory can fulfil the same reflexive appraisal whilst remaining naturalistic and empiricist.

⁸ See for example Wartofsky (1982, 82), Ibarra and Mormann (2003, 237), Uebel (2004a, 52–55), and Vrahimis (2020, 584)).

7.3.2 *Neurath's Anti-Totalitarianism: Mosaics and Orchestration*

Despite the insistence on the centrality of self-criticism, this vision of an activist science has not been uncontroversial. At the 1939 Congress for the Unity of Science, Horace Kallen accused the movement, and Neurath in particular, of totalitarianism.⁹ Against this, Neurath insisted that the Unified Science project was 'antitotalitarian through and through' (Neurath 1946a/1983, 242). In a similar vein to Kallen, Hayek argued against the 'scientistic' mindset which he saw in Neurath's work (Hayek 1941/1952, 105). Hayek argued that the outcome of scientism would be a technocracy, working from the presumption of a complete body of knowledge to institute centralised planning in pursuit of a 'purely technical optimum of universal validity' (Hayek 1942–44/1952, 96). This caused Neurath to attempt to outline how he believed the Unified Science movement could function in a non-authoritarian way. Sadly, many of these ideas are left in embryonic form. But, there is sufficient material for us to develop these ideas on his behalf.

The worries expressed by Neurath's critics are common ones. Similar arguments have been raised about various forms of scientism and Enlightenment thought. The concern is that allowing science to play an active and constructive social role will lead to a form of technocratic authoritarianism, the attribution of authority and power to experts who dictate to and impose upon the majority of the population; the elevation of a new scientist caste, a Soviet bureaucracy with scientists instead of commissars. Hayek describes the engineering mindset of 'the man whose supreme ambition is to turn the world round him into an enormous machine, every part of which, on his pressing a button, moves according to his design' (Hayek 1942–44/1952, 101–102). This is a mindset that cultivates top-down monopolisation. Kallen makes this accusation explicitly:

'The proponents of the [unified science] program, further, would have to establish themselves in a position to exact conformity and to control education. They would need, in the twentieth century much the same powers and privileges which certain schoolmen enjoyed in the thirteenth' (Kallen 1946, 494)

This accusation involves a variety of misunderstandings and mistakes. It mistakes coordination for control, unification for centralisation, and education for indoctrination. More generally, it misrepresents Neurath's understanding of knowledge, and it misrepresents the institutions of Unified Science that he advocated.

A lot can be learned initially by simply looking at the metaphors that Neurath chose for science: the mosaic and the encyclopedia. We discussed in Chap. 2 how Neurath used the encyclopedia as the model for scientific knowledge, but some details are worth re-emphasising. Firstly, an encyclopedia is a collaborative project, which brings together expertise from a vast range of fields. Secondly, an encyclopedia is non-dogmatic and pluralistic. There is no official line or ideology that must be conformed to, nor is there a requirement that all information be

⁹ For more on the historical context of Kallen's criticisms, see Reisch (2005, Chapter 9).

completely systematised. For the historical development of science, Neurath used the metaphor of a slowly evolving mosaic, ‘changing stones for others and varying the whole pattern’ (Neurath et al. 1938/1971, 3). A mosaic is a whole, composed of many smaller units. Crucially though, for a mosaic there is no key piece, no single essential foundational component, only an inter-connection of parts. And a mosaic is flat; no stratification, no hierarchies, no gradation. These key features of pluralist collaboration, horizontal integration, and limited systemisation must be kept in mind.

What Kallen and Hayek overlook is that technocracy is deeply at odds with the rejection of pseudo-rationalism (and therefore the left-Vienna Circle’s conception of unified science). For technocracy to function, a single optimal outcome or system must be discoverable and implementable by experts. The authority of the technocrat in imposing their will relies on their having the “right” answer. But this is fundamentally incompatible with Carnap and Neurath’s anti-absolutism. ‘*“The” system is the great scientific lie*. Not even as an anticipated goal is it a useful guiding thought’ (Neurath 1935a/1983, 116). From the very start then, the accusation of technocracy is misguided when we understand Neurath’s epistemology. In a previous section, we discussed Neurath’s recognition of the inevitability of intellectual pluralism within science. This reasoning extends into the social, political and institutional too:

‘acknowledging a kind of primary “pluralism” in our scientific approach has also its consequences for our daily life. If science enables us to make more than one sound prediction, how may we use science as a means of action? We can never avoid a “decision”, because no account would be able to show us one action as “the best”, no computation would present us with any “optimum”, wherever actions have to be discussed. Therefore “decision” plays its part in any kind of scientific research as well as in our daily life.’ (Neurath 1946d, 80)

So how are such decisions to be made?

One may still argue that even given the rejection of a univocal decision calculus, technocracy may still have a role. After all, experts could play the role of decision makers. But again, Neurath rejects the privileging of scientists with special decision-making authority:

‘logical empiricism does not see why scientists, trained to discover as many alternatives as possible, should be particularly able to select one alternative only (one that never can be based on calculation) by making a decision or performing an action for other people with different desires and attitudes’ (Neurath 1946a/1983, 239)

Here it becomes clear that the accusation of technocracy not only misrepresents Neurath’s understanding of scientific knowledge, but also the role of expert authority in science. We have noted on numerous occasions the importance of decisionism in Neurath’s philosophy. But in none of those cases does he resort to appeals to experts or authority. In fact, Neurath’s understanding of the organisation of science is explicitly anti-authoritarian:

‘the encyclopedism of logical empiricism challenges any intellectual authority which pretends to preach the truth (whether it identifies the truth with the leader’s intuition, with the interest of the state or any other human group, with the decision of a deity, or with anything else) it is out of the question that it should not challenge any attempt to misuse any kind of distorted empiricism for creating a similar authority’ (Neurath 1946a/1983, 238)

Not only does Neurath reject placing scientists in positions of institutional power, but he rejects the suggestion that experts alone play the role of executive decision-makers. Neurath's entire epistemology undermines the notion of scientists as those with unique or privileged access to truth which could licence granting scientists such executive authority.

Whilst the intellectual authority of scientific expertise inevitably remains, this is not the pernicious and oppressive institutionalised authority that Neurath's critics assumed. A recognition of the value of expert opinion is compatible with a rejection of centralised institutional power and scepticism of authority, especially if combined with a broader democratisation of knowledge which allows a well-informed public to make reasoned decisions when confronted with competing plans for dealing with specific problems given available resources.

Science is a communal enterprise directed at broad societal goals, and as such the decisions that are central to it ought to be taken communally. Participation was not to be exclusive; 'if someone who is standing outside a particular branch of knowledge has something scientifically testable to contribute, he is only welcome' (Neurath 2011, 21). The increased dissemination of knowledge allows for increased participation, with the goal of the democratisation of knowledge. This was the purpose of Neurath's involvement in various educational projects.¹⁰ This is aided by the unification of language, which diminishes the possibility of elite cabals. 'If priests and rulers have a language of their own, they become separated from the ruled masses, and it is just the unification of language that is a step forward to some democratic possibilities' (Neurath 1946a/1983, 236). As Neurath emphasised, linguistic unity is not reduction to some technical language, comprehensible only by the trained expert (Neurath 1946a/1983, 237). Linguistic unity involves expressibility in our everyday spatio-temporal language. Again, Neurath's imagery is illustrative: he dubs it the "universal slang", with connotations of universal accessibility, informality and ease of use. Linguistic unity means comprehensibility to the majority. And the greater the dissemination of information, the more people are capable of participation in the process of critical self-reflection. The left-wing's unity of science thesis therefore stands in direct and deliberate opposition to the possibility of technocracy. Science is both communal process and guided by decisions, so can only benefit from the involvement of as many well-informed (and it is taken for granted, good-faith) participants as possible. Neurath always saw the process as a democratic one, and 'democracy implies the rejection of experts in making decision [sic]' (Neurath 1945/1996, 251). We have already discussed the importance of pluralism, and deference to authority has been rejected. What meaningful sense of hierarchy or technocracy have we left?¹¹

When people read "Unified", the assumption tends to be "centralised". Kallen for instance explicitly equates unification with centralisation and totalitarianism (Kallen

¹⁰ For Neurath's educational work see Stadler (1989/1991), Stadler (2015, 318–320), Dvorak (1982/1991), and Cat et al. (1996). For ISOTYPE, see Müller (1991) and Nemeth (2019).

¹¹ For Neurath's parallel argument against Hayek regarding political economy, in favor of pluralistic socialist planning, see O'Neill (2007).

1946, 493). It was to counter this association that Neurath adopted Kallen's own term "orchestration" for the non-coercive coordinative efforts that Unified science would attempt (Neurath 1941/1983, 214). With the ongoing development of science, and research projects of ever greater scale, further development is only achievable by widescale collaborative work.¹² Isolated scientists simply cannot do these things on their own. And this collaboration benefits from, perhaps even requires, coordination. What it does not need is the installation of a central committee to dictate from the centre outwards. Centralisation is irrelevant, even antithetical, to Neurath's goals. We must remember Neurath's understanding of the linguistic unity of science: unity for practical purposes. The same applies here. He intended unification as means to facilitate collaboration and discussion. Collaboration does not need control, 'we stressed the point that actual cooperation in fruitful discussion should demonstrate how much unity of action can result, without any kind of authoritative integration' (Neurath 1946a/1983, 230).

Also important to note is that, when unification is distinguished from centralisation, it does not imply homogeneity or uniformity. As Neurath highlights, cooperation does not require theoretical agreement (Neurath 2011, 18). What it does require is linguistic unity and an institutional or organisational apparatus that facilitates conversation in a maximally accessible way. But beyond the basic requirement of mutual intelligibility and the desiderata of public availability (speaking a common language and sharing data), there are no other theoretical requirements for collaboration.¹³ Unified science can accommodate scientists who disagree on theories, on the interpretation of experimental results, and even on the appropriate methods for a given question. Neurath's Ehrenhaft example is a case in point. In fact, Neurath argued that more extensive sharing, discussion and collaborative work may actually increase diversity:

'thinking in terms of comprehensive planning in the social field, does not imply the exclusion of variety, on the contrary, one can imagine that people acknowledge pluralism as an element of life. Arguing on these lines, one may realize a greater variety of ways of life within our society by means of comprehensive planning than people' (Neurath 1946d, 81)

Again, "planning" conjures worrying Soviet connotations, but this is misleading. Planning need not be associated with hierarchy and technocracy. It is crucial to remember that in Neurath's proposals, these are not plans imposed from the outside or above, these are plans decided on democratically within the sciences for the sciences. It isn't external control, but self-direction 'to place [science] within the reach of its practitioners' (Nemeth 1982a/1991, 282). Planning and democracy are not incompatible.

¹² Obvious examples here are international scientific projects of the last fifty years, like the International Space Station or CERN. What Neurath envisioned is in some ways much closer now than it was during his lifetime.

¹³ Neurath also makes the interesting point that protocol sentences will naturally exhibit a centrifugal pull, providing scientists with a natural cooperative baseline (Neurath 1946a/1983, 241).

The question we are left with is what form the institutions or organisations of unified science are to take. Instead of centralised planning in the soviet style, Neurath envisaged a ‘democracy of cooperation’ for the orchestration of the sciences (Neurath 1946a/1983, 236). Whilst specifics are missing, the spirit is clear: democratic, non-hierarchical, de-centralised, and allowing local autonomy within a broader framework of collaboration. These may take the form of forums, meetings, federations, or associations that allow collaboration and criticism without monopolisation by a central body. For a comparison, the closest thing to what Neurath hints at would seem to be the anti-hierarchical horizontal networks used by anarchists.¹⁴ Local groups are part of a broader association, but with sufficient local autonomy to follow their own goals and projects. Sufficient agreement on basic methods and goals co-exist with the ongoing revision of both goals and methods. ‘No system from above, but systematisation from below’ (Neurath 1936c/1983, 153). The other obvious example here is the Unified Science movement itself. It brought people together and disseminated ideas through conferences and journals. It allowed (even encouraged) debate and dissent, and had no official line that could not be questioned. And the capacity for de-centralised collaboration has only been made easier with the development of technology in the time since Neurath’s death.

It is the responsibility of the scientific community to work towards the creation of organisational forms that best suit the goals of the metatheoretical project. ‘Its evolution would be based on conventions which could never be definite or authoritative as far as the aspirations of conscientious logical empiricists are concerned’ (Neurath 1946a/1983, 242). Methodological constructivism applies at the level of unified science as a whole. Science has the same responsibilities collectively as each epistemic agent has individually. Science as a collective endeavour has a responsibility to self-regulate, to be self-critical, and to engage in the bipartite dialectic of piecemeal mutual readjustment. With no foundations and decisions to be made, the path of science is not a predetermined one to be discovered, but one to be chosen and created. But as Neurath constantly emphasised, science is always a collaborative process, and as such these decisions needed to be made on the basis of transparent public discussion and this required a democratisation of knowledge and broadening of education. This would allow for decisions to be made by the community at large, and for co-ordinated action to be orchestrated in the aftermath. Fundamentally, methodological constructivism is incompatible with technocracy, because technocracy is a renunciation of epistemic responsibility on behalf of the vast majority of the epistemic community.

The radical honesty described above as central to methodological constructivism also plays an additional role:

‘There is no judge in a chair who decides who is nearer the truth. There is no way of “impartiality” or “scientific objectivity”, there is no point outside our life, from which we may finally decide what is “impartial” or “scientifically objective”’ (Neurath 1946b/1983, 243)

¹⁴ This also suggests that Neurath has a better claim than Feyerabend to the name “anarchist epistemology”.

In so far as science can be objective, it is not objectivity as external.¹⁵ Transparency about the motivations and methodology behind decisions allow for further discussion, opening each of these factors up for debate too. In their response to conventionalism and decisionism, both Neurath and Carnap demonstrate an honestly optimistic response. They have abandoned the pseudo-rationalist pursuit of absolutes, but are still able to retain a hope in the human capacity for improvement through intellectual progress that echoes the Enlightenment. Institutional or organisational level methodological constructivism maximises the self-reflexivity of science as a whole.

7.4 Conclusion

What the foregoing should make clear is that, when fully fleshed out, the Bipartite Metatheory that Neurath and Carnap present goes far beyond a narrow traditional empiricist epistemology. The naturalistic rejection of pseudo-rationalism and the reorientation of philosophy it entails opens the door to a project that is much broader and much more radical than the caricatured picture of the Vienna Circle's Logical Positivism. The integration of the theoretical and the practical and the embrace of decisionism, necessitates science taking a more active and self-conscious role in its own development, and how this development effects the world beyond it. This combines with the self-reflexivity inherent in the rejection of pseudo-rationalism, to provide a pluralistic, democratic and non-hierarchical unified science (and within it the bipartite metatheory), a project free from paternalistic authoritarianism, and instead embracing a more humble and egalitarian vision of the role of scientific progress in shaping the future. The unified science project of scientific metatheory fulfils the vision first expressed in the Vienna Circle's manifesto:

'the spirit of the scientific world-conception penetrating in growing measure the forms of personal and public life, in education, upbringing, architecture, and the shaping of economic and social life according to rational principles' (Carnap et al. 1929/1973, 317–318)

¹⁵ This conclusion has important similarities to Heather Douglas's recent work on values and objectivity in science Douglas (2007). How the dissimilarities are to be dealt with is to be investigated elsewhere.

References

- Antony, Louise M. 1993. "Quine as Feminist: The Radical Import of Naturalized Epistemology." In *A Mind of One's Own: Feminist Essays on Reason and Objectivity*, ed. Louise. M. Antony and Charlotte E. Witt, 185–225. Boulder: Westview Press.
- Armour-Garb, Bradley, Daniel Stoljar, and James Woodbridge. 2022. "Deflationism About Truth." In *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/entries/truth-deflationary/>. Accessed December 2022
- Awodey, Steve. 2007. "Carnap's Quest for Analyticity: The Studies in Semantics." In *The Cambridge Companion to Carnap*, ed. Micheal Friedman and Richard Creath, 226–247. Cambridge: Cambridge University Press.
- Awodey, Steve, and André Carus. 2007. "The Turning Point and the Revolution: Philosophy of Mathematics in Logical Empiricism from Tractatus to Logical Syntax." In *The Cambridge Companion to Logical Empiricism*, ed. Alan Richardson and Thomas Uebel, 165–192. Cambridge: Cambridge University Press.
- . 2009. "From Wittgenstein's Prison to the Boundless Ocean: Carnap's Dream of Logical Syntax." In *Carnap's Ideal of Explication and Naturalism*, ed. Pierre Wagner, 79–106. New York: Palgrave Macmillan.
- Ayer, A. J. 1936. "Verification and Experience." *Proceedings of the Aristotelian Society* 37(1): 137–156.
- . 1936b/2001. *Language, Truth and Logic*. London: Penguin. Repr.
- . 1956. *The Problem of Knowledge*. Middlesex: Penguin Books.
- . ed. 1959. *Logical Positivism*. New York: Free Press.
- . 1963. "Carnap's Treatment of the Problem of Other Minds." In *The Philosophy of Rudolf Carnap*, ed. Paul Arthur Schilpp, 269–282. Illinois: Open Court.
- . 1977. *Part of My Life*. London: Collins.
- . 1982. *Philosophy in the Twentieth Century*. London: Unwin Paperbacks.
- Bar-Hillel, Yehoshua, and Rudolf Carnap. 1953. "Semantic Information." *The British Journal for the Philosophy of Science* 4(14): 147–157.
- Ben-Menahem, Yemima. 2005. "Black, White and Gray: Quine on Convention." *Synthese* 146: 245–282.
- Bentley, Joseph. 2022. "Protocol statements, physicalism, and metadata: Otto Neurath on scientific evidence." *Studies in History and Philosophy of Science Part A* 96(C):125–134.
- Bentley, Joseph and Thomas Uebel. (2023). "On Quine's Epistemological Objection to Carnap's Analyticity." In *Interpreting Carnap: Critical Essays*, ed. Alan Richardson and Adam Tuboly. Cambridge: Cambridge University Press.

- Biddle, Justin. 2013. "State of the Field: Transient Underdetermination and Values in Science." *Studies in History and Philosophy of Science* 44(1): 124–133.
- Bird, Alexander. 2012. "Kuhn, Naturalism, and the Social Study of Science." In *Kuhn's the Structure of Scientific Revolutions Revisited*, ed. Vasso Kindi and Theodore Arabatzis, 205–230. New York: Routledge.
- Boghossian, Paul A. 1996. "Analyticity Reconsidered." *Noûs* 30(3): 360–391.
- . 2003. "Epistemic Analyticity: A Defense." *Grazer Philosophische Studien* 66(1): 15–35.
- Bonioli, G. 2003. "Kant's Explication and Carnap's Explication: The Redde Rationem." *International Philosophical Quarterly* 43(3): 289–298.
- Brun, Georg. 2016. "Explication as a Method of Conceptual Re-Engineering." *Erkenntnis* 81(6): 1211–1241.
- . 2017. "Conceptual Re-Engineering: From Explication to Reflective Equilibrium." *Synthese* 197: 925–945.
- Bunge, M. 2011. "Two Unification Strategies: Analysis or Reduction, and Synthesis or Integration." In *Otto Neurath and the Unity of Science*, ed. John Symons, Olga Pombo, and Juan Manuel Torres, 145–157. New York: Springer Dordrecht.
- Butts, Robert E. 1999. "The Reception of German Scientific Philosophy in North America: 1930–1962." In *Experience, Reality, And Scientific Explanation: Essays In Honor Of Merrilee And Wesley Salmon*, ed. Maria Carla Galavotti and Alessandro Pagnini, 1–14. Dordrecht: Kluwer.
- Campbell, Richmond. 1998. *Illusions of Paradox*. Boulder: Rowman and Littlefield Publishers.
- Carnap, Rudolf. 1928a/2005. *The Logical Structure of the World and Pseudoproblems in Philosophy*. Chicago: Open Court. Translated by Rolf A. George. Repr. 2005.
- . 1928b/2005. "Pseudoproblems in Philosophy." In *The Logical Structure of the World and Pseudoproblems in Philosophy*. Chicago: Open Court. Translated by Rolf A. George, 323–343. Repr.
- . 1930/1959. "The Old and New Logic." In *Logical Positivism*, ed. A.J. Ayer, 133–146. New York: Free Press. Translated by Isaac Levi
- . 1931/1983. "The Logicist Foundations of Mathematics." Translated by Paul Benacerraf and Hilary Putnam In *The Philosophy of Mathematics: Selected Readings*, ed. Paul Benacerraf and Hilary Putnam, 2nd ed., 31–41. Cambridge: Cambridge University Press.
- . 1932a/1959. "The Elimination of Metaphysics Through Logical Analysis of Language." In *Logical Positivism*, ed. A.J. Ayer, 60–81. New York: Free Press. Translated by Arthur Pap.
- . 1932b/1959. "Psychology in Physical Language." In *Logical Positivism*, ed. A.J. Ayer, 165–198. New York: Free Press. Translated by George Schick.
- . 1932c/1995. *The Unity of Science*. Bristol: Thoemmes Press. Translated by Max Black. Repr.
- . 1932d/1987. "On Protocol Sentences." *Noûs* 21(4): 457–70. Translated by Richard Creath and Richard Nollan.
- . 1934a/1987. "The Task of the Logic of Science." In *Unified Science*, ed. Robert S. Cohen, 46–66. Dordrecht: D. Reidel.
- . 1934c. "On the Character of Philosophic Problems." *Philosophy of Science* 1(1): 5–19.
- . 1934d/2001. *The Logical Syntax of Language*. London: Routledge. Translated by Amethe Smeaton. Repr.
- . 1934e. "Die Antinomien und die Unvollständigkeit der Mathematik." *Monatsh. f. Mathematik und Physik* 41: 263–284.
- . 1935. *Philosophy and Logical Syntax*. London: Kegan Paul, Trench, Trubner and Co.
- . (1936/1949). "Wahrheit und Bewährung." In *Actes du Congrès international de philosophie scientifique, Sorbonne, Paris 1935, fasc. 4. Unité de la science (Paris: Hermann)*, 18–23. *Readings in Philosophical Analysis*, ed. Herbert Feigl and Wilfred Sellars, 119–127. New York: Appleton-Century-Crofts. Modified and translated as "Truth and Confirmation".
- . 1936–37. "Testability and Meaning." *Philosophy of Science* 3(4): 419–471, 4(1): 1–40.
- . 1938/1991. "Logical Foundations of the Unity of Science." In *The Philosophy of Science*, ed. Richard Boyd, Philip Gasper, and J.D. Trout, 393–404. Cambridge: MIT Press.
- . 1939/1971. "Foundations of Logic and Mathematics." *International Encyclopedia of Unified Science* 1(3). Chicago: University of Chicago Press.

- 1942/1948. *Introduction to Semantics*. Cambridge: Harvard University Press. Repr.
- 1945. "The Two Concepts of Probability." *Philosophy and Phenomenological Research* 5: 513–532.
- 1949. "Truth and Confirmation." In *Readings in Philosophical Analysis*, ed. Herbert Feigl and Wilfred Sellars, 119–127. New York: Appleton-Century-Crofts.
- 1950. "Empiricism, Semantics, and Ontology." *Revue Internationale De Philosophie* 4(11): 20–40.
- 1952a/1988. "Meaning Postulates." In *Meaning and Necessity: A Study in Semantics and Modal Logic*, 2nd ed., 222–229. Chicago: University of Chicago Press. Repr.
- 1952b/1990. "Quine on Analyticity." In *Dear Carnap, Dear Van: The Quine-Carnap Correspondence and Related Works*, ed. Richard Creath, 427–432. Berkeley: University of California Press.
- 1955a. "Meaning and Synonymy in Natural Languages." *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 6(3): 33–47.
- 1955b. "On Some Concepts of Pragmatics." *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 6(6): 89–91.
- 1956a. "The Methodological Character of Theoretical Concepts." In *The Foundations of Science and the Concepts of Psychology and Psychoanalysis*, ed. Herbert Feigl and Michael Scriven, 38–76. Minneapolis: University of Minnesota Press.
- 1956b/1988. *Meaning and Necessity: A Study in Semantics and Modal Logic*, 2nd ed. Chicago: University of Chicago Press. Repr.
- 1961/2005. "Preface to the Second Edition." In *The Logical Structure of the World and Pseudoproblems in Philosophy*, v–xi. Chicago: Open Court. Repr. 2005.
- 1962. *Logical Foundations of Probability*, 2nd ed. Chicago: University of Chicago Press.
- 1963a. "W.V. Quine on Logical Truth." In *The Philosophy of Rudolf Carnap*, ed. Paul Arthur Schilpp, 915–922. Illinois: Open Court.
- 1963b. "P.F. Strawson on Linguistic Naturalism." In *The Philosophy of Rudolf Carnap*, ed. Paul Arthur Schilpp, 933–939. Illinois: Open Court.
- 1963c. "Nelson Goodman on Der Logische Aufbau der Welt." In *The Philosophy of Rudolf Carnap*, ed. Paul Arthur Schilpp, 944–946. Illinois: Open Court.
- 1963d. "Intellectual Autobiography." In *The Philosophy of Rudolf Carnap*, ed. Paul Arthur Schilpp, 3–86. Illinois: Open Court.
- 1963e. "Robert S. Cohen on Dialectical Materialism vs. Empiricism." In *The Philosophy of Rudolf Carnap*, ed. Paul Arthur Schilpp, 863–867. Illinois: Open Court.
- 1963f. "Nelson Goodman on Der logische Aufbau der Welt." In *The Philosophy of Rudolf Carnap*, ed. Paul Arthur Schilpp, 944–946. Illinois: Open Court.
- 1963g. "Ayer on Other Minds." In *The Philosophy of Rudolf Carnap*, ed. Paul Arthur Schilpp, 886–889. Illinois: Open Court.
- 1963h. "Abraham Kaplan on Value Judgements." In *The Philosophy of Rudolf Carnap*, ed. Paul Arthur Schilpp, 999–1016. Illinois: Open Court.
- Carnap, Rudolf, Hans Hahn and Otto Neurath. 1929/1973. "The Scientific Conception of the World: The Vienna Circle." In *Empiricism and Sociology*, ed. Marie Neurath and Robert S. Cohen, 299–318. Dordrecht: D. Reidel.
- Cartwright, Nancy, and Jordi Cat. 1996. "Neurath against Method." In *Origins of Logical Empiricism*, ed. Ronald N. Giere and Alan W. Richardson, 80–90. University of Minnesota Press: Minneapolis.
- Cartwright, Nancy, Jordi Cat, Lola Fleck, and Thomas Uebel. 1996. *Otto Neurath: Philosophy Between Science and Politics*. Cambridge: Cambridge University Press.
- Carus, André W. 1999. "Carnap, Syntax, and Truth*." In *Truth and Its Nature (If Any)*, ed. Jaroslav Peregrin, 15–35. Dordrecht: Springer Science+Business Media.
- . 2007a. *Carnap and Twentieth-Century Thought*. Cambridge: Cambridge University Press.
- . 2007b. "Carnap's Intellectual development." In *The Cambridge Companion to Carnap*, ed. Micheal Friedman and Richard Creath, 19–42. Cambridge: Cambridge University Press.

- . 2013. "History and the Future of Logical Empiricism." In *The Historical Turn in Analytic Philosophy*, ed. Eric Reck, 261–293. New York: Palgrave-Macmillan.
- . 2017. "Carnapian Rationality." *Synthese* 194: 163–184.
- . 2019. "Neurath and Carnap on Semantics." In *Neurath Reconsidered: New Sources and Perspectives*, ed. Jordi Cat and Adam T. Tuboly, 339–361. Cham: Springer Nature.
- Cat, Jordi. 1995. "The Popper-Neurath Debate and Neurath's Attack on Scientific Method." *Studies in History and Philosophy of Science* 26(2): 219–250.
- . 2019. "Otto Neurath." *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta, Winter 2017 ed. <https://plato.stanford.edu/archives/fall2022/entries/neurath/>. Accessed December 2022.
- Cat, Jordi and Adam T. Tuboly, eds. 2019. *Neurath Reconsidered: New Sources and Perspectives*. Cham: Springer Nature.
- Cat, Jordi, Nancy Cartwright, and Hasok Chang. 1996. "Otto Neurath: Politics and the Unity of Science." In *The Disunity of Science: Boundaries, Contexts, and Power*, ed. Peter Galison and David J. Stump, 347–369. Stanford: Stanford University Press.
- Coffa, J. Alberto. 1976. "Carnap's Sprachanschauung Circa 1932." *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* 2: 205–241.
- . 1987. "Carnap, Tarski and the Search for Truth." *Noûs* 21(4): 547–572.
- . 1991. *The Semantic Tradition from Kant to Carnap*. Cambridge: Cambridge University Press.
- Creath, Richard. 1987a. "The Initial Reception of Carnap's Doctrine of Analyticity." *Noûs* 21(4): 477–499.
- . 1987b. "Some Remarks on 'Protocol Sentences'." *Noûs* 21(4): 471–475.
- . 1990a. "The Unimportance of Semantics." *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association, 1990*: 405–416.
- , ed. 1990b. *Dear Carnap, Dear Van: The Quine-Carnap Correspondence and Related Works*. Berkeley: University of California Press.
- . 1991. "Every Dogma Has Its Day." *Erkenntnis* 35(1/3): 347–389.
- . 1992. "Carnap's Conventionalism." *Synthese* 93(1/2): 141–165.
- . 1996. "The Unity of Science: Carnap, Neurath, and Beyond." In *The Disunity of Science: Boundaries, Contexts, and Power*, ed. Peter Galison and David J. Stump, 158–169. Stanford: Stanford University Press.
- . 1999. "Carnap's Move to Semantics: Gains and Losses." In *Alfred Tarski and the Vienna Circle*, ed. Jan Woleński and Eckehart Köhler, 65–76. Dordrecht: Kluwer Academic Publications.
- . 2004. "Quine on the Intelligibility and Relevance of Analyticity." In *The Cambridge Companion to Quine*, ed. Roger Gibson, 47–64. Cambridge: Cambridge University Press.
- . 2007a. "Quine's challenge to Carnap." In *The Cambridge Companion to Carnap*, ed. Micheal Friedman and Richard Creath, 316–335. Cambridge: Cambridge University Press.
- . 2007b. "Vienna, the City of Quine's Dreams." In *The Cambridge Companion to Logical Empiricism*, ed. Alan Richardson and Thomas Uebel, 332–345. Cambridge: Cambridge University Press.
- . 2012. "Before Explication." In *Carnap's Ideal of Explication and Naturalism*, ed. Pierre Wagner, 161–175. New York: Palgrave Macmillan.
- Dahms, Hans-Joachim. 1996. "Vienna Circle and French Enlightenment – A Comparison of Diderot's Encyclopédie with Neurath's International Encyclopedia of Unified Science." In *Encyclopedia and Utopia: The Life and Work of Otto Neurath (1882–1945)*, ed. Elisabeth Nemeth and Friedrich Stadler, 53–62. Dordrecht: Kluwer Academic Publishers.
- . 2004. "Neue Sachlichkeit in the Architecture and Philosophy of the 1920s." In *Carnap Brought Home: The View from Jena*, ed. Steve Awodey and Carsten Klein, 357–375. Chicago: Open Court.
- Dewey, John. 1930. *The Quest for Certainty: A Study of the Relation of Knowledge and Action*. London: George Allen and Unwin.

- Douglas, Heather. 2007. "Rejecting the Ideal of Value-Free Science." In *Value Free Science? Ideals and Illusions*, ed. Harold Kincaid, John Dupré, and Alison Wylie, 120–142. Oxford: Oxford University Press.
- Dutilh Novaes, Catarina. 2020. "Carnapian Explication and Ameliorative Analysis: A Systematic Comparison." *Synthese* 197: 1011–1034.
- Dutilh Novaes, Catarina, and Eric Reck. 2017. "Carnapian Explication, Formalisms as Cognitive Tools, and the Paradox of Adequate Formalization." *Synthese* 194: 195–215.
- Dvorak, Johann. (1982/1991). "Otto Neurath and Adult Education – Unity of Science, Materialism, and Comprehensive Enlightenment." In *In Rediscovering the Forgotten Vienna Circle: Austrian Studies on Otto Neurath and the Vienna Circle*, ed. Thomas Uebel, 265–274. Dordrecht: Kluwer Academic. Translated by Thomas Uebel.
- Eagle, Antony. 2004. "Twenty-One Arguments against Propensity Analyses of Probability." *Erkenntnis* 60(3): 371–416.
- Ebbs, Gary. 2011. "Carnap and Quine on Truth by Convention." *Mind* 120(478): 193–237.
- . 2014. "Quine's Naturalistic Explication of Carnap's Logic of Science." In *A Companion to W.V.O. Quine*, ed. Gilbert Harman and Ernie Lepore, 465–482. Chichester: John Wiley and Sons.
- . 2016. "Reading Quine's Claim That No Statement Is Immune to Revision." In *Quine and His Place in History*, ed. Frederique Janssen-Lauret and Gary Kemp, 123–145. New York: Palgrave Macmillan.
- . 2017. *Carnap, Quine, and Putnam on Methods of Inquiry*. Cambridge: Cambridge University Press.
- Feigl, Herbert. 1970. "Homage to Rudolf Carnap." In *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association 1970*, XI–XVI.
- Feldman, Richard. 1999. "Methodological Naturalism in Epistemology." In *The Blackwell Guide to Epistemology*, ed. John Greco and Ernest Sosa, 170–186. Oxford: Blackwell Publishing.
- . 2012. "Naturalized Epistemology." *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/archives/sum2012/entries/epistemology-naturalized/>. Accessed December 2022.
- Frank, Philipp. 1917/1949. "The Importance for Our Times of Ernst Mach's Philosophy of Science." In *Modern Science and Its Philosophy*, 61–78. Cambridge: Harvard University Press.
- . 1929/1941. "Physical Theories of the Twentieth Century and School Philosophy." In *Between Physics and Philosophy*, 55–103. Cambridge: Harvard University Press.
- . 1934/1941. "Is There a Trend Today Toward Idealism in Physics?" In *Between Physics and Philosophy*, 104–126. Cambridge: Harvard University Press.
- . 1935/1941. "The Positivist and the Metaphysical Conception of Physics." In *Between Physics and Philosophy*, 127–138. Cambridge: Harvard University Press.
- . 1936a/1987. "The Fall of Mechanistic Physics." In *Unified Science*, ed. Robert S. Cohen, 110–129. Dordrecht: D. Reidel.
- . 1936b/1941. "Philosophical Misinterpretations of the Quantum Theory." In *Between Physics and Philosophy*, 151–171. Cambridge: Harvard University Press.
- . 1936c/1949. "Logical Empiricism and the Philosophy of the Soviet Union." In *Modern Science and Its Philosophy*, 198–206. Cambridge: Harvard University Press.
- . 1938/1949. "Ernst Mach and the Unity of Science." In *Modern Science and Its Philosophy*, 79–89. Cambridge: Harvard University Press.
- . 1941a. *Between Physics and Philosophy*. Cambridge Massachusetts: Harvard University Press.
- . 1941b/1949. "Why do Scientists and Philosophers So Often Disagree About the Merits of a New Theory?" In *Modern Science and Its Philosophy*, 207–215. Cambridge: Harvard University Press.
- . 1947. "The Institute for the Unity of Science." *Synthese* 6(3/4): 160–167.
- . 1949a. *Modern Science and Its Philosophy*. Cambridge: Harvard University Press.
- . 1949b. "Introduction: Historical Background." In *Modern Science and Its Philosophy*, 13–61. Cambridge: Harvard University Press.
- . 1950. *Relativity: A Richer Truth*. Boston: The Beacon Press.

- 1951. “The Logical and Sociological Aspects of Science.” *Proceedings of the American Academy of Arts and Sciences* 80(1): 16–30.
- 1954. “The Variety of Reasons for the Acceptance of Scientific Theories.” *The Scientific Monthly* 79(3): 139–145.
- 1955. “Dogmatism as an Element of Acceptance of Theory.” *The Scientific Monthly* 80(4): 278–279.
- 1957/2004. *Philosophy of Science: The Link Between Science and Philosophy*. New York: Dover. Repr.
- 2021. *The Humanistic Background of Science*, ed. George Reisch and Adam Tamas Tuboly. New York: SUNY Press.
- Friedman, Michael. 1992. “Epistemology in the ‘Aufbau.’” *Synthese* 93(1/2): 15–57.
- 1999. *Reconsidering Logical Positivism*. Cambridge: Cambridge University Press.
- 2001. *Dynamics of Reason*. Stanford: CSLI Publications.
- 2006. “Carnap and Quine: Twentieth-Century Echoes of Kant and Hume.” *Philosophical Topics* 34(1/2): 35–58.
- 2007. “The Aufbau and the rejection of metaphysics.” In *The Cambridge Companion to Carnap*, ed. Micheal Friedman and Richard Creath, 129–152. Cambridge: Cambridge University Press.
- Frost-Arnold, Greg. 2013. *Carnap, Tarski, and Quine at Harvard: Conversations on Logic, Mathematics, and Science*. Chicago: Open Court Press.
- Gabriel, Gottfried. 2007. “Carnap and Frege.” In *The Cambridge Companion to Carnap*, ed. Micheal Friedman and Richard Creath, 65–80. Cambridge: Cambridge University Press.
- Galison, Peter. 1990. “Aufbau/Bauhaus: Logical Positivism and Architectural Modernism.” *Critical Inquiry* 16(4): 709–752.
- Gamow, George. 1949. “SUPERNOVAE.” *Scientific American* 181(6): 18–21.
- Glock, Hans-Johann. 2008. *What Is Analytic Philosophy?* Cambridge: Cambridge University Press.
- Goldman, Alvin. 1994. “Naturalistic Epistemology and Reliabilism.” *Midwest Studies In Philosophy* 19(1): 301–320.
- 1999. “A Priori Warrant and Naturalistic Epistemology.” *Philosophical Perspectives* 13: 1–28.
- Goodman, Nelson. 1983. *Fact, Fiction, and Forecast*, 4th ed. Cambridge: Harvard University Press.
- Grice, H. Paul, and P.F. Strawson. 1956. “In Defense of a Dogma.” *The Philosophical Review* 65(2): 141–158.
- Groß, Angélique. 2019. “Generating Cognitive Tools: Neurath’s Educational Ideal and the Concept of ISOTYPE.” In *Neurath Reconsidered: New Sources and Perspectives*, ed. Jordi Cat and Adam T. Tuboly, 177–195. Cham: Springer Nature.
- Gustafsson, Martin. 2014. “Quine’s Conceptio of Explication – and Why It Isn’t Carnap’s.” In *A Companion to W.V.O. Quine*, ed. Gilbert Harman and Ernie Lepore, 508–525. Chichester: John Wiley and Sons.
- Grundmann, Thomas. 1996. “Can Science be Likened to a Well-Written Fairy Tale? A Contemporary Reply to Schlick’s Objections to Neurath’s Coherence Theory.” In *Encyclopedia and Utopia: The Life and Work of Otto Neurath (1882–1945)*, ed. Elisabeth Nemeth and Friedrich Stadler, 127–134. Dordrecht: Kluwer Academic Publishers.
- Haack, Susan. 1990. “Rebuilding the Ship while Sailing on the Water.” In *Perspectives on Quine*, ed. Robert Barrett and Roger Gibson, 111–127. Cambridge: Basil Blackwell.
- 1993. “The Two Faces of Quine’s Naturalism.” *Synthese* 94(3): 335–356.
- Hacking, Ian. 1983. *Representing and Intervening: Introductory Topics in the Philosophy of Natural Science*. Cambridge: Cambridge University Press.
- Hahn, H. 1929/1980. “Empiricism, Mathematics, and Logic.” In *Empiricism, Logic and Mathematics: Philosophical Papers*, ed. Brian McGuinness, 39–42. Dordrecht: D. Reidel.
- 1930a. “The Significance of the Scientific World View, Especially for Mathematics and Physics.” In *Empiricism, Logic and Mathematics: Philosophical Papers*, ed. Brian McGuinness, 20–42. Dordrecht: D. Reidel.

- 1930b/1980. "Superfluous Entities, or Occam's Razor." In *Empiricism, Logic and Mathematics: Philosophical Papers*, ed. Brian McGuinness, 1–19. Dordrecht: D. Reidel.
- 1933/1987. "Logic, Mathematics, and Knowledge of Nature." In *Unified Science*, ed. Brian McGuinness, 24–45. Dordrecht: D. Reidel.
- Haller, Rudolf, ed. 1982. *Schlick und Neurath—ein Symposium*. Grazer Philosophische Studien 16/17.
- Haller, R. 1982a/1991. "The Neurath Principle: Its Grounds and Consequences." In *Rediscovering the Forgotten Vienna Circle: Austrian Studies on Otto Neurath and the Vienna Circle*, ed. Thomas Uebel, 117–30. Dordrecht: Kluwer Academic. Translated by Thomas Uebel
- 1982b. "New Light On The Vienna Circle." *The Monist* 65(1): 25–37.
- 1982c/1991. "Two Ways of Experiential Justification." In *Rediscovering the Forgotten Vienna Circle: Austrian Studies on Otto Neurath and the Vienna Circle*, ed. Thomas Uebel, 191–202. Dordrecht: Kluwer Academic. Translated by Thomas Uebel
- 1985/1991. "The First Vienna Circle." In *Rediscovering the Forgotten Vienna Circle: Austrian Studies on Otto Neurath and the Vienna Circle*, ed. Thomas Uebel, 95–108. Dordrecht: Kluwer Academic. Translated by Thomas Uebel
- . 1996. "Otto Neurath – For and Against." In *Encyclopedia and Utopia: The Life and Work of Otto Neurath (1882–1945)*, ed. Elisabeth Nemeth and Friedrich Stadler, 29–38. Dordrecht: Kluwer Academic Publishers.
- Hanfling, Otto. 1996. "Logical Positivism." In *Routledge History of Philosophy Volume IX: Philosophy of Science, Logic and Mathematics in the Twentieth Century*, ed. Stuart G. Shanker, 193–213. London: Routledge.
- Hanna, Joseph F. 1968. "An Explication of 'Explication'." *Philosophy of Science* 35(1): 28–44.
- Harman, Gilbert. 1967. "Quine on Meaning and Existence, I. The Death of Meaning." *The Review of Metaphysics* 21(1): 124–151.
- . 1996. "Analyticity Regained?" *Noûs* 30(3): 392–400.
- Haslanger, Sally. 2007. "What Good are Our Intuitions?" *Aristotelian Society Supplementary Volume* 80(1): 89–118.
- Hayek, F.A. 1941/1952. "The Counter-Revolution in Science." In *The Counter-Revolution of Science: Studies in the Abuse of Reason*, 103–206. Glencoe, Illinois: Free Press.
- 1942–44/1952. "Scientism and the Study of Society." In *The Counter-Revolution in Science: Studies in the Abuse of Reason*, 11–102. Glencoe, Illinois: Free Press.
- Hempel, Carl. 1935. "On the Logical Positivists' Theory of Truth." *Analysis* 2(4): 49–59.
- 1950/1959. "The Empiricist Criterion of Meaning." In *Logical Positivism*, ed. A.J. Ayer, 108–129. New York: Free Press.
- 1975. "Rudolf Carnap, Logical Empiricist." In *Rudolf Carnap, Logical Empiricist: Materials and Perspectives*, ed. Jaakko Hintikka, 1–14. Dordrecht: D. Reidel.
- 1983a/2000. "Schlick and Neurath: Foundation vs. Coherence in Scientific Knowledge." In *Selected Philosophical Essays*, ed. Richard Jeffrey, 181–98. Cambridge: Cambridge University Press. Translated by Christian Piller,
- 1983b. "Valuation and Objectivity in Science." In *Physics, Philosophy and Psychoanalysis: Essays in Honor of Adolf Grünbaum*, ed. Robert S. Cohen and Larry Laudan, 73–100. Dordrecht: Kluwer.
- Hintikka, Jaakko. 1988. 'On the Development of the Model-Theoretic Viewpoint in Logical Theory.' *Synthese* 77(1): 1–36.
- 1990. "Quine as a Member of the Tradition of the Universality of Language." In *Perspectives on Quine*, ed. Robert Barrett and Roger Gibson, 159–175. Cambridge: Basil Blackwell.
- Holdcroft, David. 1983. "Schlick and the Verification Theory of Meaning." *Revue Internationale de Philosophie* 37(144–145): 47–68.
- Holton, Gerald. 1978. "Subelectrons, Presuppositions, and the Millikan-Ehrenhaft Dispute." *Historical Studies in the Physical Sciences* 9: 161–224.
- Howard, Don. 2003. "Two Left Turns Make a Right: On the Curious Political Career of North American Philosophy of Science at Midcentury." In *Logical Empiricism in North America*, ed.

- Gary L. Hardcastle and Alan W. Richardson, 25–93. Minneapolis: University of Minnesota Press.
- . 2009. “Better Red Than Dead - Putting an End to the Social Irrelevance of Post-War Philosophy of Science.” *Science & Education* 18(2): 199–220.
- . 2019. “Otto Neurath: The Philosopher in the Cave.” In *Neurath Rediscovered: New Sources and Perspectives*, ed. Jordi Cat and Adam T. Tuboly, 45–65. Cham: Springer Nature.
- Hung, Tscha. 1985. “Remarks on Affirmations (‘Konstatierungen’).” *Synthese* 64(3): 297–306.
- Hylton, Peter. 2001. “‘The Defensible Province of Philosophy’: Quine’s 1934 Lectures on Carnap.” In *Future Pasts: The Analytic Tradition in Twentieth-Century Philosophy*, ed. Juliet Floyd and Sanford Shieh, 257–276. New York: Oxford University Press.
- . 2002. “Analyticity and Holism in Quine’s Thought.” *The Harvard Review of Philosophy* 10(1): 11–26.
- . 2007. *Quine*. New York: Routledge.
- . 2014. “Quine’s Naturalism Revisited.” In *A Companion to W.V.O. Quine*, ed. Gilbert Harman and Ernie Lepore, 148–162. Chichester: John Wiley and Sons.
- . 2017. “Carnap and Quine on the Nature of Evidence (and the Nature of Philosophy).” *The Monist* 100(2): 211–227.
- . 2021. “Carnap and Quine on analyticity: The nature of the disagreement.” *Noûs* 55(2): 445–462.
- Hylton, Peter and Gary Kemp. 2019. “Willard Van Orman Quine.” *Stanford Encyclopedia of Philosophy* <https://plato.stanford.edu/entries/quine/>. Accessed September 2022
- Ibarra, Andoni, and Thomas Mormann. 2003. “‘Engaged scientific philosophy in the Vienna Circle: The case of Otto Neurath.’” *Technology in Society* 25(2): 235–247.
- Isaac, Joel. 2011. “Missing Links: W.V. Quine, the making of ‘Two Dogmas’, and the analytic roots of post-analytic philosophy.” *History of European Ideas* 37(3): 267–279.
- James, William. 1907/2000. “Pragmatism: A New Name for Some Old Ways of Thinking.” In *Pragmatism and Other Writings*, ed. Giles Gunn, 1–132. New York: Penguin.
- Janssen-Lauret, Frederique. 2018. “Willard Van Orman Quine’s Philosophical Development in the 1930s and 1940s.” In *The Significance of the New Logic*, ed. Walter Carnielli, Frederique Janssen-Lauret, and William Pickering, iv–xlviii. Cambridge: Cambridge University Press.
- Jardine, Nick. 2008. “Explanatory Genealogies and Historical Testimony.” *Episteme* 5(2): 160–179.
- Jeffrey, Richard. 1992. *Probability and the art of judgement*. Cambridge: Cambridge University Press.
- Justus, James. 2012. “Carnap on Concept Determination: Methodology for Philosophy of Science.” *European Journal for Philosophy of Science* 2: 161–179.
- . 2014. “Carnap’s Forgotten Criterion of Empirical Significance.” *Mind* 123(490): 415–436.
- Kallen, Horace. 1946. “The Meanings of ‘Unity.’ Among the Sciences, Once More.” *Philosophy and Phenomenological Research* 6(4): 493–496.
- Keas, Micheal N. 2018. “Systematizing the theoretical virtues.” *Synthese* 95: 2761–2793.
- Kemp, Gary. 2014. “Quine’s Relationship with Analytic Philosophy.” In *A Companion to W.V.O. Quine*, ed. Gilbert Harman and Ernie Lepore, 69–88. Chichester: John Wiley and Sons.
- Kim, Jaegwon. 1988. “What Is ‘Naturalized Epistemology’?” *Philosophical Perspectives* 2: 381–405.
- Kitchener, Richard F. 2004. “Logical Positivism, Naturalistic Epistemology, and the Foundations of Psychology.” *Behavior and Philosophy* 32(1): 37–54.
- Kitcher, Philip. 1992. “The Naturalists Return.” *The Philosophical Review* 101(1): 53–114.
- . 2001. *Science, Truth, and Democracy*. Oxford: Oxford University Press.
- . 2008. “Carnap and the Caterpillar.” *Philosophical Topics* 36(1): 111–127.
- Koppelberg, Dirk. 1989. “Quine and the Analytic Tradition: A Reply to Professor Agassi.” *Journal for General Philosophy of Science* 20(2): 367–372.
- . 1990. “Why and How to Naturalize Epistemology.” In *Perspectives on Quine*, ed. Robert Barrett and Roger Gibson, 200–211. Cambridge: Basil Blackwell.

- Kornblith, Hilary. 1995. "Naturalistic Epistemology and Its Critics." *Philosophical Topics* 23(1): 237–255.
- . 1999. "In Defense of a Naturalized Epistemology." In *The Blackwell Guide to Epistemology*, ed. John Greco and Ernest Sosa, 158–169. Oxford: Blackwell Publishing.
- Kraft, Viktor. 1950/1953. *The Vienna Circle: The Origins of Neo-Positivism*. New York: Philosophical Library. Translated by Arthur Pap.
- Kuhn, Thomas. (1962/2012). *The Structure of Scientific Revolutions*, 4th ed. Chicago: University of Chicago Press.
- . 1977/1998. "Objectivity, Value Judgement, and Theory Choice." In *Philosophy of Science: The Central Issues*, ed. Martin Curd and Jan A. Cover, 102–118. New York: W.W. Norton and Company.
- . 1992/2000. "The Trouble with the Historical Philosophy of Science." In *The Road Since Structure: Philosophical Essays, 1970–1993*, ed. James Conant and John Haglund, 105–120. Chicago: The University of Chicago Press.
- Lakatos, Imre. 1970. "Falsification and the Methodology of Scientific Research Programmes." In *The Methodology of Scientific Research Programmes: Philosophical Papers*, ed. John Worrall and Gregory Currie, vol. 1, 8–101. Cambridge, Cambridge University Press.
- Laudan, Larry. 1990. "Normative Naturalism." *Philosophy of Science* 57(1): 44–59.
- Leitgeb, Hannes and André Carus. 2020. "Rudolf Carnap." In *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/entries/carnap/>. Accessed December 2022.
- Limbeck-Lilienau, Christoph. 2018. "The First Vienna Circle: Myth or Reality?" *Hungarian Philosophical Review* 62(4): 50–65.
- Lipton, Peter. 2007. "Alien Abduction: Inference to the Best Explanation and the Management of Testimony." *Episteme* 4(3): 238–251.
- Loomis, Eric, and Cory Juhl. 2006. "Explication." In *The Philosophy of Science: An Encyclopedia*, ed. Sahotra Starker and Jessica Pfeifer, 287–294. New York: Routledge.
- Longino, Helen. 1987. "Can There Be A Feminist Science?" *Hypatia* 2(3): 51–64.
- Lutz, Sebastian. 2020. "Armchair Philosophy Naturalized." *Synthese* 197: 1099–1125.
- Maffie, James. 1990a. "Recent Work on Naturalized Epistemology." *American Philosophical Quarterly* 27(4): 281–293.
- . 1990b. "Naturalism and the Normativity of Epistemology." *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 59(3): 333–349.
- Magee, Bryan. 1997. *Confessions of a Philosopher*. London: Phoenix Books.
- Maher, Patrick. 2007. "Explication Defended." *Studia Logica* 86: 331–341.
- . 2010. "Explication of Inductive Probability." *Journal of Philosophical Logic* 39(6): 593–616.
- Mancosu, Paolo. 2005. "Harvard 1940–1941: Tarski, Carnap and Quine on a Finitistic Language of Mathematics for science." *History and Philosophy of Logic* 26: 327–357.
- . 2008. "Tarski, Neurath, and Kokoszyńska on the Semantic Conception of Truth." In *New Essays on Tarski and Philosophy*, ed. Douglas Patterson, 192–224. Oxford: Oxford University Press.
- McGuinness, Brian, ed. 1979. *Wittgenstein and the Vienna Circle: Conversations Recorded by Friedrich Waismann*. London: Basil Blackwell.
- . 1987. *Unified Science*. Dordrecht: D. Reidel Publishing Company.
- Menger, Karl. 1994. *Reminiscences of the Vienna Circle and the Mathematical Colloquium*, ed. Louise Goddard, Brian McGuinness and Abe Sklar. Dordrecht: Springer.
- Mills Boyd, Nora. 2018. "Evidence Enriched." *Philosophy of Science* 85(3): 403–421.
- Mormann, Thomas. 1996. "Encyclopedism As an Anti-Cartesian Account of Language and Science." In *Encyclopedia and Utopia: The Life and Work of Otto Neurath (1882–1945)*, ed. Elisabeth Nemeth and Friedrich Stadler, 87–96. Dordrecht: Kluwer Academic Publishers.
- . 1999. "Neurath's Opposition To Tarskian Semantics." In *Alfred Tarski and the Vienna Circle*, ed. Jan Woleński and Eckehart Köhler, 165–178. Dordrecht: Kluwer Academic Publications.
- Morris, Charles. 1960. "On the History of the International Encyclopedia of Unified Science." *Synthese* 12(4): 517–521.

- Müller, Karl. H. (1991). "Neurath's Theory of Pictorial-Statistical Representation." In *Rediscovering the Forgotten Vienna Circle: Austrian Studies on Otto Neurath and the Vienna Circle*, ed. Thomas Uebel, 223–251. Dordrecht: Kluwer Academic. Translated by Thomas Uebel.
- Murphey, Murray G. 2012. *The Development of Quine's Philosophy*. New York: Springer.
- Nelson, Jack. 1996. "The Last Dogma of Empiricism?" In *Feminist Interpretations of Quine*, ed. Lynn Hankinson Neldon and Jack Nelson, 307–347. Philadelphia: Temple University Press.
- Nelson, Lynn Hankinson. (1990). *Who Knows: From Quine to a Feminist Empiricism*. Philadelphia: Temple University Press.
- Nelson, Lynn Hankinson and Jack Nelson, eds. 2003. *Feminist Interpretations of Quine*. Pennsylvania: The Pennsylvania State University Press.
- Nemeth, Elisabeth. 1982a/1991. "The Unity of Planned Economy and the Unity of Science." In *Rediscovering the Forgotten Vienna Circle: Austrian Studies on Otto Neurath and the Vienna Circle*, ed. Thomas Uebel, 275–84. Dordrecht: Kluwer Academic. Translated by Thomas Uebel.
- . 1982b/1991. "Otto Neurath's Utopias—The Will to Hope." In *Rediscovering the Forgotten Vienna Circle: Austrian Studies on Otto Neurath and the Vienna Circle*, ed. Thomas Uebel, 285–92. Dordrecht: Kluwer Academic. Translated by Thomas Uebel.
- . 2007. "Logical Empiricism and the History and Sociology of Science." In *The Cambridge Companion to Logical Empiricism*, ed. Alan Richardson and Thomas Uebel, 278–302. Cambridge: Cambridge University Press.
- . 2019. "Visualizing Relations in Society and Economics: Otto Neurath's Isotype-Method Against the Background of his Economic Thought." In *Neurath Rediscovered: New Sources and Perspectives*, ed. Jordi Cat and Adam T. Tuboly, 117–140. Cham: Springer Nature.
- Nemeth, Elisabeth and Friedrich Stadler, eds. 1996. *Encyclopedia and Utopia: The Life and Work of Otto Neurath (1882–1945)*. Dordrecht: Kluwer Academic Publishers.
- Neurath, Bohr. 1912/1973. "The Problem of the Pleasure Maxim." In *Empiricism and Sociology*, ed. Marie Neurath and Robert S. Cohen, 113–122. Dordrecht: D. Reidel.
- . 1913/1983. "The Lost Wanderers of Descartes and the Auxiliary Motive." In *Philosophical Papers 1913–1946*, ed. R. S. Cohen and Marie Neurath, 1–12. Dordrecht: D. Reidel.
- . 1915/1973. "On the Foundations of the History of Optics." In *Empiricism and Sociology*, ed. Marie Neurath and Robert S. Cohen, 101–112. Dordrecht: D. Reidel.
- . 1916/1983. "On the Classification of Systems of Hypotheses." In *Philosophical Papers 1913–1946*, ed. Robert S. Cohen and Marie Neurath, 13–31. Dordrecht: D. Reidel.
- . 1921/1973. "Anti-Spengler." In *Empiricism and Sociology*, ed. Marie Neurath and Robert S. Cohen, 158–213. Dordrecht: D. Reidel.
- . 1928/1973. "Personal Life and Class Struggle." In *Empiricism and Sociology*, ed. Marie Neurath and Robert S. Cohen, 249–98. Dordrecht: D. Reidel.
- . 1930/1983. "Ways of the Scientific World-Conception." In *Philosophical Papers 1913–1946*, ed. Robert S. Cohen and Marie Neurath, 32–47. Dordrecht: D. Reidel.
- . 1931a/1983. "Physicalism: The Philosophy of the Viennese Circle." In *Philosophical Papers 1913–1946*, ed. Robert S. Cohen and Marie Neurath, 48–51. Dordrecht: D. Reidel.
- . 1931b/1983. "Physicalism." In *Philosophical Papers 1913–1946*, ed. Robert S. Cohen and Marie Neurath, 52–57. Dordrecht: D. Reidel.
- . 1931c/1983. "Sociology in the Framework of Physicalism." In *Philosophical Papers 1913–1946*, ed. Robert S. Cohen and Marie Neurath, 58–90. Dordrecht: D. Reidel.
- . 1931d/1973. "Empirical Sociology. The Scientific Content Of History And Political Economy." In *Empiricism and Sociology*, ed. Marie Neurath and Robert S. Cohen, 319–421. Dordrecht: D. Reidel.
- . 1932a/1983. "Protocol Statements." In *Philosophical Papers 1913–1946*, ed. Robert S. Cohen and Marie Neurath, 91–99. Dordrecht: D. Reidel.
- . 1932b/1987. "Unified Science and Psychology." In *Unified Science*, ed. Robert S. Cohen, 1–23. Dordrecht: D. Reidel.
- . 1934/1983. "Radical Physicalism and the 'Real World'." In *Philosophical Papers 1913–1946*, ed. Robert S. Cohen and Marie Neurath, 100–114. Dordrecht: D. Reidel.

- 1935a/1983. "The Unity of Science as a Task." In *Philosophical Papers 1913–1946*, ed. Robert S. Cohen and Marie Neurath, 115–120. Dordrecht: D. Reidel.
- 1935b/1983. "Pseudorationalism of Falsification." In *Philosophical Papers 1913–1946*, ed. Robert S. Cohen and Marie Neurath, 121–131. Dordrecht: D. Reidel.
- 1936a/1983. "Individual Sciences, Unified Science, Pseudorationalism." In *Philosophical Papers 1913–1946*, ed. Robert S. Cohen and Marie Neurath, 132–138. Dordrecht: D. Reidel.
- 1936b/1983. "An International Encyclopedia of Unified Science." In *Philosophical Papers 1913–1946*, ed. Robert S. Cohen and Marie Neurath, 139–144. Dordrecht: D. Reidel.
- 1936c/1983. "Encyclopedia as a 'Model'." In *Philosophical Papers 1913–1946*, ed. Robert S. Cohen and Marie Neurath, 145–158. Dordrecht: D. Reidel.
- 1936d/1983. "Physicalism and the Investigation of Knowledge." In *Philosophical Papers 1913–1946*, ed. Robert S. Cohen and Marie Neurath, 159–171. Dordrecht: D. Reidel.
- 1937a/1983. "The New Encyclopedia of Scientific Empiricism." In *Philosophical Papers 1913–1946*, ed. Robert S. Cohen and Marie Neurath, 189–199. Dordrecht: D. Reidel.
- 1937b/1983. "The Departmentalization of Unified Science." In *Philosophical Papers 1913–1946*, ed. Robert S. Cohen and Marie Neurath, 200–205. Dordrecht: D. Reidel.
- 1937c/1983. "Unified Science and Its Encyclopedia." In *Philosophical Papers 1913–1946*, ed. Robert S. Cohen and Marie Neurath, 173–182. Dordrecht: D. Reidel.
- 1937d/1987. "The New Encyclopedia." In *Unified Science*, ed. Robert S. Cohen, 132–141. Dordrecht: D. Reidel.
- 1937e/1983. "Comments on the Papers by Black, Kokoszynska, Williams." In *Philosophical Papers 1913–1946*, ed. Robert S. Cohen and Marie Neurath, 206–208. Dordrecht: D. Reidel.
- 1938/1971. "Unified Science as Encyclopedic Integration." In *Encyclopedia and Unified Science, (International Encyclopedia of Unified Science)*, ed. Otto Neurath, Niels Bohr, John Dewey, Bertrand Russell, Rudolf Carnap, and Charles Morris, vol. 1.1, 1–27. Chicago: University of Chicago Press.
- 1939/1983. "The Social Sciences and Unified Science." In *Philosophical Papers 1913–1946*, ed. Robert S. Cohen and Marie Neurath, 209–212. Dordrecht: D. Reidel.
- 1941/1983. "Universal Jargon and Terminology." In *Philosophical Papers 1913–1946*, ed. Robert S. Cohen and Marie Neurath, 213–229. Dordrecht: D. Reidel.
- 1944/1971. "Foundations of the Social Sciences." *International Encyclopedia of Unified Science*, vol. 2, number 1. Chicago: University of Chicago Press.
- 1945/1996. "Visual Education: Humanisation vs. Popularisation." In *Encyclopedia and Utopia: The Life and Work of Otto Neurath (1882–1945)*, ed. Elisabeth Nemeth and Friedrich Stadler, 245–335. Dordrecht: Kluwer Academic Publishers. Unfinished MS, Repr.
- 1946a/1983. "The Orchestration of the Sciences by the Encyclopedism of Logical Empiricism." In *Philosophical Papers 1913–1946*, ed. Robert S. Cohen and Marie Neurath, 230–242. Dordrecht: D. Reidel.
- 1946b/1983. "Prediction and Induction." In *Philosophical Papers 1913–1946*, ed. Robert S. Cohen and Marie Neurath, 243–246. Dordrecht: D. Reidel.
- 1946c. "For the Discussion: Just Annotations, Not a Reply." *Philosophy and Phenomenological Research* 6(4): 526–528.
- 1946d. "After six years." *Synthese* 5: 77–82.
- 2011. "Unity of Science and Logical Empiricism: A Reply." In *Otto Neurath and the Unity of Science*, ed. John Symons, Olga Pombo, and Juan Manuel Torres, 15–30. New York: Springer Dordrecht. Translated by Thomas Bonk.
- Neurath, Bohr, John Dewey, Bertrand Russell, Rudolph Carnap and Charles W. Morris, eds. 1938/1971. *Encyclopedia and Unified Science (International Encyclopedia of Unified Science)*, vol. 1.1. Chicago: University of Chicago Press. Repr.
- Niaz, Mansoor. 2000. "The Oil Drop Experiment: A Rational Reconstruction of the Millikan-Ehrenhaft Controversy and its Implications for Chemistry Textbooks." *Journal of Research in Science Teaching* 37: 480–508.
- 2005. "An Appraisal of the Controversial Nature of the Oil Drop Experiment: Is Closure Possible?" *The British Journal for the Philosophy of Science* 56(4): 681–702.

- Niiniluoto, I. 1999. "Theories of Truth: Vienna, Berlin, and Warsaw." In *Alfred Tarski and the Vienna Circle*, ed. Jan Woleński and Eckehart Köhler, 17–26. Dordrecht: Kluwer Academic Publications.
- Oberdan, Thomas. 1992. "The Concept of Truth in Carnap's 'Logical Syntax of Language'." *Synthese* 93(1/2): 239–260.
- . 1993. *Protocols, Truth, Convention*. Amsterdam: Rodopi.
- . 1996. "Postscript to Protocols: Reflections on Empiricism." In *Origins of Logical Empiricism*, edited by Ronald N. Giere, Herbert Feigl, and Alan W. Richardson, 269–91. Minneapolis: University of Minnesota Press.
- Okruhlik, Kathleen. 2004. "Logical Empiricism, Feminism, and Neurath's Auxiliary Motive." *Hypatia* 19(1): 48–72.
- Olsson, Erik. 2015. "Gettier and the Method of Explication: A 60 Year Old Solution to a 50 Year Old Problem." *Philosophical Studies* 172: 57–72.
- . 2021. "Coherentist Theories of Epistemic Justification." *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/entries/justep-coherence/>. Accessed December 2022
- O'Neill, John. 2003. "Unified Science as Political Philosophy: Positivism, Pluralism and Liberalism." *Studies in History and Philosophy of Science* 34(3): 575–596.
- . 2007. "Pluralism and Economic Institutions." In *Neurath's Economics in Context*, ed. Elisabeth Nemeth, Stefan W. Schmitz and Thomas Uebel, 77–100. Dordrecht: Kluwer.
- O'Neill, John and Thomas Uebel. 2004. "Horkheimer and Neurath: Restarting a Disrupted Debate." *European Journal of Philosophy* 12: 75–105.
- Oppenheim, Paul and Hilary Putnam. 1958. "Unity of Science as a Working Hypothesis." *Minnesota Studies in the Philosophy of Science* 2: 3–36.
- Pavese, Carlotta. 2020. "Lewis Carroll's Regress and the Presuppositional Structure of Arguments." *Linguistics and Philosophy* 1: 1–38.
- Passmore, John. 1967. "Logical Positivism." In *Encyclopedia of Philosophy*, ed. Donald M. Borcherst, , vol. 5, 2nd ed., 524–530. New York: Thomson Gale.
- Pincok, Christopher. 2005. "A Reserved Reading of Carnap's Aufbau." *Pacific Philosophical Quarterly* 86(4): 518–543.
- . 2007. "Carnap, Russell, and the external world." In *The Cambridge Companion to Carnap*, ed. Micheal Friedman and Richard Creath, 106–128. Cambridge: Cambridge University Press.
- Pombo, Olga. 2011. "Neurath and the Encyclopaedic Project of Unity of Science." In *Otto Neurath and the Unity of Science*, ed. John Symons, Olga Pombo, and Juan Manuel Torres, 59–70. New York: Springer Dordrecht.
- Popkin, Richard H. 1992. "Schlick and Scepticism." In *The Third Force in Seventeenth-Century Thought*, 254–267. Leiden: Brill.
- Popper, K. 1935/2002. *The Logic of Scientific Discovery*. London: Routledge. Translated by the author. Repr.
- . 1973. "Memories of Otto Neurath." In *Empiricism and Sociology*, ed. Marie Neurath and Robert S. Cohen, 51–56. Dordrecht: D. Reidel.
- . 1974/2002. *Unended Quest: An Intellectual Autobiography*. London: Routledge. Repr.
- . 1979. *Objective Knowledge: An Evolutionary Approach*, Revised ed. Oxford: Clarendon Press.
- Psillos, Stathis. 1999. *Scientific Realism: How Science Tracks Truth*. London: Routledge.
- Putnam, Hilary. 1962. "The Analytic and the Synthetic." In *Mind, Language and Reality. Philosophical Papers*, vol. 2, 33–69. Cambridge: Cambridge University Press.
- Quine, Willard Van Orman. 1934/1990. "Lectures on Carnap." In *Dear Carnap, Dear Van: The Quine-Carnap Correspondence and Related Works*, ed. Richard Creath, 45–103. Berkeley: University of California Press.
- . 1936/1966. "Truth by Convention." In *The Ways of Paradox and Other Essays*, 70–99. New York: Random House.
- . 1943. "Notes on Existence and Necessity." *The Journal of Philosophy* 40(5): 113–127.
- . 1946a/2008. "On the Notion of an Analytic Statement." In *Confessions of a Confirmed Extensionalist and Other Essays*, 24–35. Cambridge: Harvard University Press.

- . 1946b/2008. “W.V. Quine’s 1946 Lectures on David Hume’s Philosophy.” In *Confessions of a Confirmed Extensionalist and Other Essays*, 50–136. Cambridge: Harvard University Press.
- . 1951a/1980. “Two Dogmas of Empiricism.” In *From a Logical Point of View*, 2nd Revised ed., 20–46. Cambridge: Harvard University Press.
- . 1951b. “On Carnap’s Views on Ontology.” *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 2(5): 65–72.
- . 1954/1966. “Carnap and Logical Truth.” In *The Ways of Paradox and Other Essays*, 100–125. New York: Random House.
- . 1960/2013. *Word and Object*, New ed. Cambridge: MIT Press.
- . 1963. “Carnap and Logical Truth.” In *The Philosophy of Rudolf Carnap*, ed. Paul Arthur Schilpp, 385–406. Illinois: Open Court.
- . 1968/1969. “Ontological Relativity.” In *Ontological Relativity and Other Essays*, Willard Von Orman Quine, 26–68. New York: Columbia University Press.
- . 1969a. *Ontological Relativity and Other Essays*. New York: Columbia University Press.
- . 1969b. “Epistemology Naturalized.” In *Ontological Relativity and Other Essays*, Willard Von Orman Quine, 69–90. New York: Columbia University Press.
- . 1970. “Homage to Rudolf Carnap” In *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association 1970*, XXII–XXV.
- . 1974. *The Roots of Reference*. La Salle: Open Court.
- . 1975. “The Nature of Natural Knowledge.” In *Mind and Language: Wolfson College Lectures 1974*, ed. Samuel Guttenplan, 67–81. Oxford: Clarendon Press.
- . 1981a. *Theories and Things*. Cambridge: Belknap Press.
- . 1981b. “Five Milestones of Empiricism.” In *Theories and Things*, Willard Von Orman Quine, 67–72. Cambridge: Belknap Press.
- . 1986. *Philosophy of Logic*, 2nd ed. Cambridge: Harvard University Press.
- . 1987. “Reply to Morton White.” In *The Philosophy of W. V. Quine*, ed. Lewis Edwin Hahn and Paul Arthur Schilpp, Expanded ed., 663–665. Illinois: Open Court.
- . 1990a. “Comment on Koppelberg.” In *Perspectives on Quine*, ed. Robert Barrett and Roger Gibson, 212. Cambridge: Basil Blackwell.
- . 1990b. “Comment on Hintikka.” In *Perspectives on Quine*, ed. Robert Barrett and Roger Gibson, 176. Cambridge: Basil Blackwell.
- . 1991. “Two Dogmas in Retrospect.” *Canadian Journal of Philosophy* 21(3): 265–274.
- . 1992. *Pursuit of Truth*, Revised ed. Cambridge: Harvard University Press.
- Reck, Eric. 2012. “Carnapian Explication: A Case Study and Critique.” In *Carnap’s Ideal of Explication and Naturalism*, ed. Pierre Wagner, 96–116. New York: Palgrave Macmillan.
- Reisch, George. 1991. “Did Kuhn Kill Logical Empiricism?” *Philosophy of Science* 58(2): 264–277.
- . 1997. “How postmodern was Neurath’s Idea of Unity of Science?” In *Studies in History and Philosophy of Science Part A* 28(3): 439–451.
- . 2003. “On the International Encyclopedia, the Neurath-Carnap Disputes, and the Second World War.” In *Logical Empiricism: Historical and Contemporary Perspectives*, ed. Paolo Parrini, Wesley C. Salmon, and Merrilee H. Salmon, 94–108. Pittsburgh: University of Pittsburgh Press.
- . 2005. *How the Cold War Transformed Philosophy of Science: To the Icy Slopes of Logic*. Cambridge: Cambridge University Press.
- . 2007. “From ‘the Life of the Present’ to the ‘Icy Slope of Logic’: Logical Empiricism, the Unity of Science Movement, and the Cold War.” In *The Cambridge Companion to Logical Empiricism*, ed. Alan Richardson and Thomas Uebel, 58–87. Cambridge: Cambridge University Press.
- Reisch, George and Adam Tamas Tuboly. 2021. “Philipp Frank: A Crusader for Scientific Philosophy.” In *The Humanistic Background of Science*, ed. Philipp Frank, 1–67. New York: SUNY Press.
- Richardson, Alan. 1990. “How Not to Russell Carnap’s Aufbau.” In *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association, 1990*, 3–14.

- . 1996. "From Epistemology to the Logic of Science: Carnap's Philosophy of Empirical Knowledge in the 1930s." In *Origins of Logical Empiricism*, ed. Ronald N. Giere and Alan W. Richardson, 309–332. University of Minnesota Press: Minneapolis.
- . 1997a. "'Two Dogmas about Logical Empiricism: Carnap and Quine on Logic, Epistemology, and Empiricism.'" *Philosophical Topics* 25(2): 145–168.
- . 1997b. *Carnap's Construction of the World*. Cambridge: Cambridge University Press.
- . 2003. "Logical Empiricism, American Pragmatism, and the Fate of Scientific Philosophy in North America." In *Logical Empiricism in North America*, ed. Gary L. Hardcastle and Alan W. Richardson, 1–24. Minneapolis: University of Minnesota Press.
- . 2004. "Tolerating Semantics: Carnap's Philosophical Point of View." In *Carnap Brought Home: The View from Jena*, ed. Steve Awodey and Carsten Klein, 63–78. Chicago: Open Court.
- . 2007. "Carnapian pragmatism." In *The Cambridge Companion to Carnap*, ed. Micheal Friedman and Richard Creath, 295–315. Cambridge: Cambridge University Press.
- Ricketts, Thomas. 1994. "Carnap's Principle of Tolerance, Empiricism, and Conventionalism." In *Reading Putnam*, ed. Peter Clark and Bob Hale, 176–200. Cambridge: Blackwell.
- . 1996. "Carnap: From Logical Syntax to Semantics." In *Origins of Logical Empiricism*, ed. Ronald N. Giere and Alan W. Richardson, 231–250. University of Minnesota Press: Minneapolis.
- . 2004. "Frege, Carnap, and Quine: Continuities and Discontinuities." In *Carnap Brought Home: The View from Jena*, ed. Steve Awodey and Carsten Klein, 181–202. Chicago: Open Court.
- . 2007. "Tolerance and logicism: logical syntax and the philosophy of mathematics." In *The Cambridge Companion to Carnap*, ed. Micheal Friedman and Richard Creath, 200–225. Cambridge: Cambridge University Press.
- Romizi, Donata. 2012. "The Vienna Circle's 'Scientific World-Conception': Philosophy of Science in the Political Arena." *HOPOS: The Journal of the International Society for the History of Philosophy of Science* 2(2): 205–242.
- Rorty, Richard. 1997. "Introduction." In *Empiricism and the Philosophy of Mind*, ed. Wilfrid Sellars, 1–12. Cambridge: Harvard University Press.
- . 1998. "Against Unity." *The Wilson Quarterly* 22(1): 28–38.
- Rosenberg, Alex. 1996. "A Field Guide to Recent Species of Naturalism." *The British Journal for the Philosophy of Science* 47(1): 1–29.
- Roth, Paul. 2003. "Feminism and Naturalism: If Asked for Theories, Just Say 'No'." In *Feminist Interpretations of Quine*, ed. Lynn Hankinson Neldon and Jack Nelson, 271–305. Philadelphia: Temple University Press.
- Rudner, Richard. 1953. "The Scientist Qua Scientist Makes Value Judgments." *Philosophy of Science* 20(1): 1–6.
- Russell, Bertrand. 1937. "On Verification: The Presidential Address." *Proceedings of the Aristotelian Society* 38: 1–20.
- . 1940/1995. *An Inquiry Into Meaning and Truth*. New York: Routledge. Repr.
- . 1948/2009. *Human Knowledge: Its Scope and Limits*. London: Routledge.
- Russell, Gillian. 2014. "Quine on the Analytic/Synthetic Distinction." In *A Companion to W.V.O. Quine*, ed. Gilbert Harman and Ernie Lepore, 181–202. Chichester: John Wiley and Sons.
- Russell, L. J. 1934. "Communication and Verification II." *Proceedings of the Aristotelian Society, Supplementary Volumes* 13: 174–193.
- Rutte, H. 1979/1991. "Neurath Contra Schlick. On the Discussion of Truth in the Vienna Circle." In *Rediscovering the Forgotten Vienna Circle: Austrian Studies on Otto Neurath and the Vienna Circle*, ed. Thomas Uebel, 169–174. Dordrecht: Kluwer Academic. Translated by Thomas Uebel.
- . 1982a/1991. "The Philosopher Otto Neurath." In *Rediscovering the Forgotten Vienna Circle: Austrian Studies on Otto Neurath and the Vienna Circle*, ed. Thomas Uebel, 81–94. Dordrecht: Kluwer Academic. Translated by Thomas Uebel.

- 1982b/1991. “On Neurath’s Empiricism and His Critique of Empiricism.” In *Rediscovering the Forgotten Vienna Circle: Austrian Studies on Otto Neurath and the Vienna Circle*, ed. Thomas Uebel, 175–190. Dordrecht: Kluwer Academic. Translated by Thomas Uebel.
- Ruytinx, Jacques. 1983. “Neurath, Protocol Sentences And The Spirit Of Logical Positivism.” *Revue Internationale De Philosophie* 37(144/145): 37–46.
- Saatsi, Juha. 2019. “What is Theoretical Progress of Science?” *Synthese* 196: 611–631.
- Sartre, Jean-Paul. 1946/2007. *Existentialism is a Humanism*. London: Methuen. Translated by P. Mairet. Repr.
- Scheffler, Israel. 1982. *Science and Subjectivity*, 2nd ed. Indianapolis: Hackett.
- Schlick, M. 1928/1979. “Preface to: Friedrich Waismann, Logik, Sprache, Philosophie.” In *Philosophical Papers Volume II (1925–1936)*, ed. Henk L. Mulder and Barbara F.B. Van De Velde-Schlick, 130–138. Dordrecht: D. Reidel.
- 1929/1979. “Philosophy and Natural Science.” In *Philosophical Papers Volume II (1925–1936)*, ed. Henk L. Mulder and Barbara F.B. Van De Velde-Schlick, 139–153. Dordrecht: D. Reidel.
- 1930a/1979. “The Turning-Point in Philosophy.” In *Philosophical Papers Volume II (1925–1936)*, ed. Henk L. Mulder and Barbara F.B. Van De Velde-Schlick, 154–160. Dordrecht: D. Reidel.
- 1930b/1979. “Is There a Factual a Priori?” In *Philosophical Papers Volume II (1925–1936)*, ed. Henk L. Mulder and Barbara F.B. Van De Velde-Schlick, 161–170. Dordrecht: D. Reidel.
- 1930c/1979. “The Future of Philosophy.” In *Philosophical Papers Volume II (1925–1936)*, ed. Henk L. Mulder and Barbara F.B. Van De Velde-Schlick, 171–175. Dordrecht: D. Reidel.
- 1932a/1979. “Positivism and Realism.” In *Philosophical Papers Volume II (1925–1936)*, ed. Henk L. Mulder and Barbara F.B. Van De Velde-Schlick, 259–284. Dordrecht: D. Reidel.
- 1932b/1979. “Form and Content. An introduction to Philosophical Thinking.” In *Philosophical Papers Volume II (1925–1936)*, ed. Henk L. Mulder and Barbara F.B. Van De Velde-Schlick, 285–369. Dordrecht: D. Reidel.
- 1934/1979. “On The Foundation of Knowledge.” In *Philosophical Papers Volume II (1925–1936)*, ed. Henk L. Mulder and Barbara F.B. Van De Velde-Schlick, 370–387. Dordrecht: D. Reidel.
- 1935a/1979. “Facts and Propositions.” In *Philosophical Papers Volume II (1925–1936)*, ed. Henk L. Mulder and Barbara F.B. Van De Velde-Schlick, 400–404. Dordrecht: D. Reidel.
- 1935b/1979. “Introduction and On ‘Affirmations’”. In *Philosophical Papers Volume II (1925–1936)*, ed. Henk L. Mulder and Barbara F.B. Van De Velde-Schlick, 405–413. Dordrecht: D. Reidel.
- 1936/1979. “The Vienna School and Traditional Philosophy.” In *Philosophical Papers Volume II (1925–1936)*, ed. Henk L. Mulder and Barbara F.B. Van De Velde-Schlick, 499–513. Dordrecht: D. Reidel.
- Schwartz, Stephen P. 2012. *A Brief History of Analytic Philosophy: From Russell to Rawls*. Oxford: Wiley-Blackwell.
- Shepherd, Joshua and James Justus. 2015. “X-Phi and Carnapian Explication.” *Erkenntnis* 80: 381–402.
- Stadler, F. 1982/1991. “Otto Neurath–Moritz Schlick. On the Philosophical and Political Antagonism in the Vienna Circle.” In *Rediscovering the Forgotten Vienna Circle: Austrian Studies on Otto Neurath and the Vienna Circle*, ed. Thomas Uebel, 159–168. Dordrecht: Kluwer Academic. Translated by Thomas Uebel.
- 1989/1991. “Otto Neurath–Encyclopedist, School Reformer and People’s Educator.” In *Rediscovering the Forgotten Vienna Circle: Austrian Studies on Otto Neurath and the Vienna Circle*, ed. Thomas Uebel, 255–264. Dordrecht: Kluwer Academic. Translated by Thomas Uebel.
- . 2007. “The Vienna Circle: Context, Profile, and Development.” In *The Cambridge Companion to Logical Empiricism*, ed. Alan Richardson and Thomas Uebel, 13–40. Cambridge: Cambridge University Press.

- . 2015. *The Vienna Circle: Studies in the Origin, Development, and Influence of Logical Empiricism*. New York: Springer.
- Stein, Howard. 1992. "Was Carnap Entirely Wrong, after All?" *Synthese* 93(1/2): 275–295.
- Stich, Stephen. 1993. "Naturalizing Epistemology: Quine, Simon and the Prospects for Pragmatism." In *Collected Papers Vol. 2: Knowledge, Rationality, and Morality, 1978–2010*, 99–112. Oxford: Oxford University Press.
- Stoljar, Daniel and Nic Damnjanovic. 2010. "The Deflationary Theory of Truth." In *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/archives/fall2014/entries/truth-deflationary/>. Accessed December 2022.
- Stöltzner, Michael. 1996. "The Auxiliary Motive in the Forest and in Optics." In *Encyclopedia and Utopia: The Life and Work of Otto Neurath (1882–1945)*, ed. Elisabeth Nemeth and Friedrich Stadler, 113–126. Dordrecht: Kluwer Academic Publishers.
- Strawson, Peter F. 1963. "Carnap's Views on Constructed Systems versus Natural Languages in Analytic Philosophy." In *The Philosophy of Rudolf Carnap*, ed. Paul Arthur Schilpp, 503–518. Illinois: Open Court.
- . 1992. *Analysis and Metaphysics: An Introduction to Philosophy*. Oxford: Oxford University Press.
- Stroud, Barry. 1981. "The Significance of Naturalized Epistemology." *Midwest Studies in Epistemology* 6(1): 455–472.
- Stuchlik, Joshua. 2011. "Felicology: Neurath's Naturalization of Ethics." *HOPOS: The Journal of the International Society for the History of Philosophy of Science* 1(2): 183–208.
- Tarski, A. 1936/1956. "The Establishment of Scientific Semantics." In *Logic, Semantics, Metamathematics: Papers from 1923 to 1938*, 401–408. Oxford: Clarendon Press. Translated by J.H. Woodger.
- Tega, Walter. 1996. "Atlases, Cities, Mosaics – Neurath and the Encyclopédie." In *Encyclopedia and Utopia: The Life and Work of Otto Neurath (1882–1945)*, ed. Elisabeth Nemeth and Friedrich Stadler, 63–78. Dordrecht: Kluwer Academic Publishers.
- Tuboly, Adam T. 2017. "Philipp Frank's decline and the crisis of logical empiricism." *Studies in East European Thought* 69(3): 257–276.
- . 2019. "United by Action: Neurath in England." In *Neurath Rediscovered: New Sources and Perspectives*, ed. Jordi Cat and Adam T. Tuboly, 89–113. Cham: Springer Nature.
- . 2021. "Introduction: From Spying to Canonizing – Ayer and His Language, Truth and Logic." In *The Historical and Philosophical Significance of Ayer's Language, Truth and Logic*, ed. Adam T. Tuboly, 1–37. Cham: Palgrave Macmillan.
- Tuboly, Adam T. and Jordi Cat, eds. 2019. "Appendix 2: The 1940–1945 Neurath-Carnap Correspondence." In *Neurath Reconsidered: New Sources and Perspectives*, Jordi Cat and Adam T. Tuboly, 512–683. Cham: Springer Nature.
- Uebel, Thomas, ed. 1991a. *Rediscovering the Forgotten Vienna Circle: Austrian Studies on Otto Neurath and the Vienna Circle*. Dordrecht: Kluwer Academic Publishers.
- Uebel, Thomas. 1991b. "Neurath's Programme for Naturalistic Epistemology." *Studies in History and Philosophy of Science Part A* 22(4): 623–646.
- . 1992. *Overcoming Logical Positivism from Within: The Emergence of Neurath's Naturalism in the Vienna Circles Protocol Sentence Debate*. Amsterdam: Rodopi.
- . 1995. "Physicalism in Wittgenstein and the Vienna Circle." In *Physics, Philosophy, and the Scientific Community*, ed. Kostas Gavroglu, John Stachel, and Marx W. Wartofsky, 327–356. Dordrecht: Kluwer.
- . 1996. "Normativity and Convention: On the Constructivist Element of Neurath's Naturalism." In *Encyclopedia and Utopia: The Life and Work of Otto Neurath (1882–1945)*, ed. Elisabeth Nemeth and Friedrich Stadler, 97–112. Dordrecht: Kluwer Academic Publishers.
- . 2000. "Logical Empiricism and the Sociology of Knowledge: The Case of Neurath and Frank." *Philosophy of Science* 67: S138–S150.
- . 2001. "Carnap and Neurath in Exile: Can their Disputes Be Resolved?" *International Studies in the Philosophy of Science* 15(2): 211–220.

- . 2003a. “On the Austrian Roots of Logical Empiricism: The Case of the First Vienna Circle.” In *Logical Empiricism. Historical and Contemporary Perspectives*, ed. Paolo Parrini, Wesley C. Salmon, and Merrilee H. Salmon, 67–93. Pittsburgh: University of Pittsburgh Press.
- . 2003b. “Philipp Frank’s History of the Vienna Circle: A Programmatic Retrospective.” In *Logical Empiricism in North America*, ed. Gary L. Hardcastle and Alan W. Richardson, 149–169. Minneapolis: University of Minnesota Press.
- . 2004a. “Education, Enlightenment and Positivism: The Vienna Circle’s Scientific World-Conception Revisited.” *Science and Education* 13: 41–66.
- . 2004b. “Naturalism and Scepticism.” In *Essays on Wittgenstein and Austrian Philosophy in Honour of J. C. Nyiri*, ed. Tamás Demeter, 63–86. Amsterdam: Rodopi.
- . 2005. “Learning Logical Tolerance: Hans Hahn on the Foundations of Mathematics.” *History and Philosophy of Logic* 26(3): 175–209.
- . 2007a. “Carnap and the Vienna Circle: Rational Reconstructionism Refined.” In *The Cambridge Companion to Carnap*, ed. Micheal Friedman and Richard Creath, 153–175. Cambridge: Cambridge University Press.
- . 2007b. “Philosophy of Social Science in Early Logical Empiricism: The Case of Radical Physicalism.” In *The Cambridge Companion to Logical Empiricism*, ed. Alan Richardson and Thomas Uebel, 250–277. Cambridge: Cambridge University Press.
- . 2007c. “Interpreting Neurath’s Protocols. Reply to Nottelmann.” *Journal for General Philosophy of Science* 38(2): 383–391.
- . 2007d. *Empiricism at The Crossroads*. New York: Open Court.
- . 2007e. “Otto Neurath as an Austrian Economist: Behind the Scenes of the Early Socialist Calculation Debate.” In *Neurath’s Economics in Context*, ed. Elisabeth Nemeth, Stefan W. Schmitz and Thomas Uebel, 37–60. Dordrecht: Kluwer.
- . 2008. “Writing a Revolution: On the Production and Early Reception of the Vienna Circle’s Manifesto.” *Perspectives on Science* 16(1): 70–102.
- . 2009a. “Neurath’s Protocol Statements Revisited: Sketch of a Theory of Scientific Testimony.” *Studies in History and Philosophy of Science Part A* 40(1): 4–13.
- . 2009b. “Carnap’s Logical Syntax in the Context of the Vienna Circle.” In *Carnap’s Logical Syntax of Language*, ed. Pierre Wagner, 53–78. Basingstoke: Palgrave Macmillan.
- . 2010. “The Nature and Status of Scientific Metatheory: The Debate Between Otto Neurath and Åke Petzäll.” In *The Vienna Circle in the Nordic Countries*, ed. Juha Manninen and Friedrich Stadler, 181–201. Dordrecht: Springer.
- . 2011. “Beyond the Formalist Criterion of Cognitive Significance: Philipp Frank’s Later Antimetaphysics”. *HOPOS: The Journal of the International Society for the History of Philosophy of Science* 1(1): 47–72.
- . 2012. “The Bipartite Conception of Metatheory and the Dialectical Conception of Explication.” In *Carnap’s Ideal of Explication and Naturalism*, ed. Pierre Wagner, 117–130. New York: Palgrave Macmillan.
- . 2013a. “Pragmatics in Carnap and Morris and the Bipartite Metatheory Conception.” *Erkenntnis* 78(3): 523–546.
- . 2013b. “‘Logical Positivism’ – ‘Logical Empiricism’: What’s in a Name?” *Perspectives on Science* 21(1): 58–99.
- . 2015. “Three Challenges to the Complementarity of the Logic and the Pragmatics of Science.” *Studies in History and Philosophy of Science Part A* 53: 23–32.
- . 2018. “Carnap’s Transformation of Epistemology and the Development of His Metaphilosophy.” *The Monist* 101(4): 367–87.
- . 2019. “Verificationism and (Some of) Its Discontents.” *Journal for the History of Analytical Philosophy* 7(4): 1–31.
- . 2020a. “Schlick and Wittgenstein: The Theory of Affirmations Revisited.” *Journal of the History of Philosophy* 58(1): 141–166.
- . 2020b. “Intersubjective Accountability: Politics and Philosophy in the Left Vienna Circle.” *Perspectives on Science* 28(1): 35–62.
- Van Fraassen, Bas. 2002. *The Empirical Stance*. London: Yale University Press.

- Van Heijenoort, Jean. 1967. "Logic as Calculus and Logic as Language." *Synthese* 17(3): 324–330.
- Verhaegh, Sander. 2018. *Working from Within: The Nature and Development of Quine's Naturalism*. New York: Oxford University Press.
- von Juhos, Béla. 1935. "Empiricism and Physicalism." *Analysis* 2(6): 8–92.
- Vrahimis, Andreas. 2020. "Scientism, Social Praxis and overcoming Metaphysics: A debate between Logical Empiricism and the Frankfurt School." In *HOPOS: The Journal of the International Society for the History of Philosophy of Science* 10(2): 562–597.
- Vrahimis, Andreas. 2021. "Language, Truth and Logic and the Anglophone Reception of the Vienna Circle." In *The Historical and Philosophical Significance of Ayer's Language, Truth and Logic*, ed. Adam T. Tuboly, 41–68. Cham: Palgrave Macmillan.
- Waismann, Friedrich. 1930/1979. "Theses." In *Wittgenstein and the Vienna Circle: Conversations Recorded by Friedrich Waismann*, ed. Brian McGuinness, 243–245. London: Basil Blackwell.
- Wartofsky, Marx W. 1982. "Positivism and Politics: The Vienna Circle as Social Movement." *Grazer Philosophische Studien* 16(1): 79–101.
- Weinberg, Julius R. 1936. *An Examination of Logical Positivism*. London: Routledge and Kegan Paul.
- Wolenski, Jan. 1996. "Against Truth as Coherence." *Logic and Logical Philosophy* 4: 41–51.
- Yap, Audrey. 2010. "Feminism and Carnap's Principle of Tolerance." *Hypatia* 25(2): 437–454.
- Young, James O. 2018. "The Coherence Theory of Truth." In *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta, Fall 2018 ed. <https://plato.stanford.edu/archives/fall2018/entries/truth-coherence/>. Accessed 23 July 2020.
- Zemplén, Gábor Á. 2019. "Neurath's Theory of Theory Classification: History, Optics and Epistemology." In *Neurath Reconsidered: New Sources and Perspectives*, ed. Jordi Cat and Adam T. Tuboly, 217–236. Cham: Springer Nature.
- Zilsel, Edgar. 1942. "The Sociological Roots of Science." In *The Social Origins of Modern Science*, 7–21. Dordrecht: Springer Science+Business Media.
- Zolo, Danilo. 1989. *Reflexive Epistemology: The Philosophical Legacy of Otto Neurath*. Dordrecht: Kluwer Academic Publishers. Translated by David McKie.
- . 1990. "Reflexive Epistemology and Social Complexity: The Legacy of Otto Neurath." *Philosophy of the Social Sciences* 20(2): 149–169.