CONVENTIONALISM

Yemima Ben-Menahem

Conventionalism

YEMIMA BEN-MENAHEM

The Hebrew University of Jerusalem

CAMBRIDGE
UNIVERSITY PRESS

Preface

The cluster of problems surrounding the notion of convention and its counterpart, the notion of truth, have always been at the very heart of philosophical inquiry. This book examines a relatively recent round in this ongoing discussion, beginning with Poincaré and ending with Quine and the later Wittgenstein. It is only during this period that the notion of convention comes to be associated with an 'ism,' a distinct philosophical position. I will focus on the philosophy of science and mathematics, setting aside other realms of philosophy, such as ethics and political theory, in which questions about the role of convention also figure prominently. Although a wide spectrum of positions fall under the rubric "conventionalism," all explore the scope and limits of epistemic discretion. On the prevailing conception, conventionalism has been taken to extend the scope of discretion to the very stipulation of truth. The thrust of the present study is a critique of this reading.

The various chapters of this book are largely self-contained, but when brought to bear on one another, they provide not only a new understanding of conventionalism, but a reframing of central themes of twentieth-century philosophy.

My debts to teachers, colleagues, students, and others who have written on the aforementioned questions are, of course, numerous. I would like to mention, in particular, Yehuda Elkana, Hilary Putnam, and the late Frank Manuel, who introduced me to the history and philosophy of science; my late physics teacher Ruth Stern, who imparted to her students a feel for the beauty of physics; and my late friends Amos Funkenstein and Mara Beller, who passed away at the peak of their creative careers. I am grateful

Contents

1

Overview

The Varieties of Conventionalism

This book recounts the hitherto untold story of conventionalism. The profound impact conventionalism has had on seminal developments in both the science and the philosophy of the twentieth century is revealed through analysis of the writings of Poincaré, Duhem, Carnap, Wittgenstein, and Quine on the subject, and by examining the debate over conventionalism in the context of the theory of relativity and the foundations of mathematics. I trace the evolution of conventionalism from Poincaré's modest but precise initial conception through a number of extravagant extrapolations, all of which, I show, eventually collapsed under the weight of the problems they generated. My focus, however, is not history but analysis. The literature is replete with ambiguity as to what the meaning of 'convention' is, misunderstandings about the aims of conventionalism, and conflation of conventionalism with other philosophical positions, such as instrumentalism and relativism. The most serious confusion pertains to the notion of *truth by convention* typically associated with conventionalism. A central theme of this book is that conventionalism does *not* purport to base truth on convention, but rather, seeks to forestall the conflation of truth and convention.

Much of twentieth-century philosophy was characterized by engagement in determining the limits of meaning and countering the tendency to ascribe meaning to meaningless expressions. Conventionalism, correctly understood, is motivated by a desire to mitigate deceptive ascription of truth. To the conventionalist, the very idea of truth by convention is as incongruous as that of meaningful nonsense. Clearly, the exposure of nonsense is philosophically important only when we are deluded as to the meaning and meaningfulness of the expressions in question, not

when it is clear to all and sundry that they are nonsensical. Similarly, the exposure of convention is philosophically important only in contexts in which we tend to delude ourselves about the nature of the beliefs in question. Conventionalism thus seeks to expose conventions likely to be mistaken for truths, and calls our attention to the fact that we do have discretion even in contexts where we appear to have none. The axioms of geometry, the original focus of Poincaré's conventionalism, clearly illustrate this misleading character: traditionally, they are construed as necessary truths, but according to the conventionalist, they serve as definitions of the entities that satisfy them. Obvious conventions, for instance, that green means 'go,' red means 'stop' – or indeed, that the particular word 'stop' has this particular meaning – are of interest to the conventionalist solely to the extent that they can be employed as simpler analogues of the disguised conventions that are really at issue. I stress this point because David Lewis's *Convention* (Lewis 1969), probably the most thorough study of convention, does not actually address the problems that motivate conventionalism. Lewis might have disagreed with this assessment, for he perceived his book to be a direct response to Quine's critique of conventionalism. Lewis maintains that Quine challenged the platitude that language is ruled by convention, but failed to make his case. This failure, he argues, was to be expected, "for when a . . . philosopher challenges a platitude, it usually turns out that the platitude was essentially right" (1969, p. 1). However, it is not this platitude that is the subject of Quine's critique, but the highly controversial thesis that convention is the sole root of analyticity and necessity. Lewis explicitly rejects what Quine deems to be the conventionalist account of necessary truth. That language is ruled by convention "is not to say that necessary truths are created by convention: only that necessary truths, like geological truths, are conventionally stated in these words rather than in those" (1969, p. 1). But neither conventionalists nor their opponents challenge this thesis; the question they debate is whether there *are* any necessary truths. In replacing the notion of necessary truth with that of linguistic convention, the conventionalist takes truth to be first and foremost a matter of empirical fact. It goes without saying that there can be empirical facts about language; for example, it is a fact that in Hebrew, adjectives generally follow the nouns they modify. Yet this rule is not *itself* grounded in fact, and is thus a convention. The thesis Quine critiques is that necessary truths are analogous to such grammatical conventions. Further elucidation of the point of contention between Lewis and Quine and an appraisal of Lewis's defense of conventionalism will be taken up in

chapters 6 and 7; here it suffices to note that the focus of conventionalism is not convention per se, but rather, convention masquerading as truth.

In a way, then, I too defend a platitude – the platitude that truth is distinct from convention and cannot be generated by fiat. (I set aside cases such as predictions made true by voluntary actions; this is not the type of case adduced by conventionalists.) Part of my argument is interpretative; on my understanding, conventionalists such as Poincaré and Carnap do not sanction the postulation of truth. That these thinkers do not espouse the view commonly associated with conventionalism does not, of course, amount to a refutation of that view. But if the most profound versions of conventionalism do not argue for the creation of truth by convention, the notion of 'truth by convention' remains nothing more than a hollow idiom unsupported by argument, indeed, an oxymoron. Nevertheless, my defense of the platitude does not consist merely in showing that conventionalists, the received reading of their ideas notwithstanding, do not challenge it. It consists, further, in showing that methods and practices thought to sustain the postulation of truth, for instance, the method of implicit definition, in fact presuppose a background of nonconventional truths.

Conventionalism has elicited both radical readings, and readings that trivialize it. The former construe conventionalism as taking truth itself to be a matter of convention; the latter limit the role of convention to the choice of one particular word, sign, or formulation rather than another. Both types of readings fail to do justice to the conventionalist position, but it is the radical readings that seem to me to be further off the mark. Ultimately, conventionalism might end up doing no more than calling attention to our discretion to choose between different formulations of the same truth; in this sense, it would indeed be noncontroversial. In cases of interest to the conventionalist, however, it is far from trivial to demonstrate that we are in fact confronted with equivalent formulations rather than divergent and incompatible theories. Subsequent developments in physics, discussed in chapter 3, bring to the fore the nontrivial character of assessments of equivalence. As the example of geometry illustrates, the most profound (and controversial) element of Poincaré's argument is not the claim that the choice of a unit of measurement, say, meters rather than yards, is up to us, but the claim that, despite appearances to the contrary, the differences between alternative geometries are actually analogous to such trivial differences in units of measurement.

In saying that conventionalists seek to distinguish fact from convention, I do not impute to them the naive conception that there are 'bare'

facts. On the contrary, the recognition that facts are described via language, and the same facts can be variously described, is the common core of the different conventionalist arguments examined in this book. Indeed, the sameness of facts can only be established by establishing a *systematic correspondence* between types of description. The description-sensitivity of facts has also been stressed by nonconventionalist philosophers. It is embodied in the intentionality of explanation and the value-ladenness of typical descriptions of human action. This phenomenon, which has been much remarked upon and analyzed quite independently of the controversy over conventionalism, will not concern me in any detail in this book (I do address it in Ben-Menahem 2001a).

I must stress, however – and here I return to my theme – that description-sensitivity does not blur the notions of truth and objectivity or undermine their centrality to our attempts to comprehend the world. Facts under a description are facts, and the assertions we make about them can be true or false, justified or unjustified, probable or improbable, compatible or incompatible with specific assertions, and so on. In other words, description-sensitivity is not at odds with either realist conceptions of truth or the fact–convention distinction. (That there are hard cases, where the borderline is fuzzy, such as Quine's '(x) x is self-identical,' should not deter us from making the distinction in garden-variety cases.) At the same time, that the notions of truth and objectivity are meaningful and applicable does not make each and every application straightforward, effortless, or infallible; we are prone to error not only with regard to identifying and describing the facts, but also with regard to the logical relations between different descriptions. We might, for example, take two theories to be inconsistent with each other when in fact they are not. This is the type of mistake conventionalists are particularly alert to; precisely because they deem truth irreducible to convention, they are eager to clear up misunderstandings about what falls under the scope of each notion. While they are by no means alone in acknowledging the significance of modes of description, conventionalists have paid specific attention to two paradigm cases that underscore the question of how facts are to be described: the case of incompatible (or seemingly incompatible) theories that are nonetheless empirically equivalent, and the case of pseudostatements (theories, inquiries) for which the factual basis is specious. My favorite example of the latter is James's quote from Lessing, "Why is it that the rich have all the money?" (James 1955, p. 144), to which I return in chapters 6 and 7.

The birth of conventionalism in the writings of Henri Poincaré at the end of the nineteenth century was a major event in the history of philosophy, comparable in some respects to Kant's Copernican revolution. The problem of a priori and necessary truth, aptly referred to as "the largest sleeping giant of modern analytic epistemology" (Coffa 1986, p. 4), had taken another dramatic turn. For the first time, the roots of some such truths – the axioms of geometry – were being sought neither in objective reality, nor in the nature of thought as such, but in human decisions about the use of language. The traditional notion of necessity was giving way to a new, and liberating, image of conceptual freedom. On the new understanding, necessary truths were not, as is often claimed, construed as truths decided on by fiat. Rather, some so-called necessary truths were denied the status of truth altogether.

Since then, conventionalism has enriched both philosophy and science, serving as a springboard for some of the most significant contributions to twentieth-century philosophy. I do not claim that these contributions were always made by proponents of conventionalism; indeed, they were often made in the course of attempting to refute conventionalism or diminish its seductive force. While the chapters on Poincaré, Duhem, and Carnap are devoted to an analysis of the conventionalist arguments put forward by these writers, the chapters on Quine and Wittgenstein present central themes in their philosophies – the indeterminacy of translation and the rule-following paradox, respectively – as critical responses to conventionalism.

In general, conventionalists had a hard time coming up with a satisfactory, let alone agreed upon, formulation of their doctrine. This is particularly true of the more extravagant versions of conventionalism: the more ambitious conventionalism became in its endeavor to extend the scope of convention, the more vulnerable it was to counterarguments impugning its coherence or intelligibility. In a sense, therefore, the story of conventionalism is the story of a highly edifying philosophical failure. In terms of impact and inspiration, however, conventionalism has been a spectacular success. The prism of conventionalism affords insight not only into the history of philosophy in the twentieth century, but also into problems on the contemporary philosophical agenda. Let me mention three examples. First, as we will see in chapter 3, the debate over the conventionality of geometry, thought to have been decided against conventionalism by the general theory of relativity, is in fact as germane and open-ended today as when conventionalism was first conceived. Second,

the method of implicit definition, discussed in chapter 4, has been a major focus of contention between realists and conventionalists. Construed as a method sanctioning stipulation of the truth of a set of axioms, it has been viewed as epitomizing the conventionalist account of necessary truth, and fiercely criticized by realists from Frege and Russell to the present. I argue that despite its association with conventionalism in the writings of Poincaré, the method of implicit definition need not transgress realist intuitions about truth. The allegation that it does is based on a misconception as to what Poincaré and Hilbert had in mind when they referred to the axioms of geometry as definitions, and worse, a flawed grasp of the method of implicit definition itself. And lastly, we will see that fundamental issues in the theory of meaning have their roots in the debate over conventionalism. Specifically, both the Kuhn-Feyerabend thesis of incommensurability and the externalist rebuttal put forward by Putnam in "The Meaning of 'Meaning' " revisit issues debated earlier by Poincaré and his critics.

How is conventionalism to be defined? We are about to see that the term 'conventionalism' has come to have radically different meanings in different contexts. In the community of philosophers of science, conventionalism is associated with the underdetermination of theory, holism, and the Duhem-Quine thesis. Popper's polemic against what he calls "the conventionalist stratagem" (Popper 1959, pp. 80–1) is a response to Duhem's influential study, *The Aim and Structure of Physical Theory*. Other philosophers of science, among them Friedman, Laudan, and Sklar, also take the term 'conventionalism' to refer to the underdetermination of theory by observation; see Friedman (1983, 1999), Laudan (1977, 1990), Sklar (1974, 1985). By contrast, in the community of analytic philosophers, 'conventionalism' usually refers to an account of necessary truth: so-called necessary truths are conventional because they either express linguistic conventions, definitions and rules, or are directly based on such conventions. This is the view often construed as sanctioning the stipulation of truth via axioms serving as implicit definitions (e.g., Wright 1980) and attacked in Quine's "Truth by Convention" (1936) and "Carnap on Logical Truth" (1960). That Quine was a merciless critic of the conventionalist account of necessary truth, yet a passionate advocate of the underdetermination of science, does not, of course, establish that these are indeed independent positions. But upon closer inspection, we will find more direct evidence that the positions in question are not merely variants of an umbrella thesis, but different, and arguably incompatible, theses.

In the remainder of this chapter, I first set out a schematic description of the aforementioned understandings of conventionalism. The search for their common roots will lead back to the context in which conventionalism was first conceived – Poincaré's philosophical writings on the epistemic and metaphysical problems raised by non-Euclidean geometries. I will point out two distinct aspects of Poincaré's argument, each of which gave rise to a different reading of conventionalism. These readings, in turn, inspired extrapolations from Poincaré's original argument that extended the scope of underdetermination, on the one hand, and the method of implicit definition, on the other. The two understandings of conventionalism I have distinguished are directly linked to these extrapolations. After showing that both extrapolations raise problems that do not afflict Poincaré's original argument, I conclude by noting the impact of these problems on the development of the views of Carnap, Quine, and Wittgenstein.

The following is a schematic presentation of my account of the history of conventionalism.

Poincaré: the conventionality of geometry

a the axioms of geometry as conventions	b underdetermination of geometry by experience

Extrapolations

a_1 necessary truths in general as conventions	b_1 underdetermination of theory in general by experience

Two conventionalist theses

a_2 a conventionalist account of necessary truth	b_2 a conventionalist account of the scientific process

Problems

1 rule following	1 demonstrating underdetermination
2 Gödel's incompleteness theorems	2 the individuation of theories
3 truth by virtue of meaning	

I. TWO READINGS OF CONVENTIONALISM

a. Conventionalism as the Underdetermination of Theory

The underdetermination thesis owes one of its most detailed formulations to Duhem, but is also associated with Neurath's boat that must be rebuilt while at sea, Reichenbach's theory of equivalent descriptions, and

Quine's holistic model of science and language. The following schematic
and nonhistorical outline of this understanding of conventionalism uses
Quinean terminology; the original Duhemian formulation is examined
in chapter 2. In its simplest form, the problem of underdetermination is
an offshoot of the problem of induction. Ideally, we would want to deduce
general laws or theories from observational data (sentences describing
such data), but in reality, we must make do, at best, with deduction in the
reverse direction – the derivation of observational consequences from
hypothetical laws and theories. As it is conceivable that incompatible the-
ories yield the same predictions, we are unable to nail down a single law
or theory that stands in the desired logical and explanatory relation to
the data. Drawing on the analogy with the underdetermination of a set of
unknowns by a number of equations that does not suffice to determine
the values of these unknowns, this situation is referred to as the under-
determination of scientific theory. Of course, such underdetermination
is a function of a particular set of data; additional data may distinguish
between hitherto indistinguishable alternatives. Thus underdetermina-
tion may be transitory or enduring. There exist today several alternative
interpretations of quantum mechanics that seem empirically equivalent
thus far but may yet prove empirically distinguishable. The question arises
whether there is a stronger kind of underdetermination that can persist
in the face of *any* additional information or testing. Upholders of under-
determination answer this question in the affirmative: scientific theory is
underdetermined by the entire body of possible observations, for there
will always be empirically equivalent but mutually incompatible theories
implying the totality of these observations. Reichenbach was particularly
sensitive to the difference between equivalence relative to a restricted
body of evidence and genuine equivalence vis-à-vis the totality of possible
observations. Only the latter, he maintains, calls for conventional choice
between alternatives, but this choice, he stresses, has nothing to do with
truth and is merely a choice between various ways of formulating the
truth.

Thus conceived, the problem of underdetermination is linked to the
built-in asymmetry between confirmation and refutation. Refutational-
ism exploits this asymmetry to argue that underdetermination frustrates
verification, not refutation. The contribution of Duhem's holism here is
that once we acknowledge that typically, scientific hypotheses are tested
collectively, not individually, the alleged asymmetry all but vanishes. The
metaphor introduced by Quine in this context is that of the intercon-
nected web of belief, bordering on experience at its periphery, and

answering to the tribunal of experience as a whole. In case of failure, various options for revision are open to the scientist, from which she chooses in line with values such as simplicity and minimal mutilation. On this account, the scientific process involves the exercise of discretion. As scientific theories are not uniquely determined by logic and experience, they are, in essence, chosen on the basis of other considerations, conscious or unconscious. It is this discretion, with respect to either the values guiding the scientist's choice or the theoretical choices made in line with these values, that licenses the terms 'convention' and 'conventionalism' in this context. These value-based conventions are not arbitrary. The claim that the notion of a 'reasoned convention' is an oxymoron (Laudan 1990, p. 88) is at odds with the way the term 'convention' has been understood and used by proponents of underdetermination from Poincaré and Duhem to Neurath and Quine.

The strong thesis of underdetermination, namely, the thesis that the entire observational and experimental repertoire is compatible with empirically equivalent but incompatible theoretical alternatives, is impressed upon us by Quine's powerful metaphor; we seem able to practically visualize the various ways in which the inner parts of the web could be rearranged without severing their ties to the periphery. Yet we should note that at this point, strong underdetermination, while suggested by this compelling image, has not actually been demonstrated. Whether a more detailed examination of Duhem's and Quine's arguments yields such a demonstration is discussed in chapters 2 and 6; I answer in the negative in both cases. Whereas Poincaré succeeds in making a convincing case for the underdetermination of geometry by experience, the more general Duhem-Quine thesis of the underdetermination of science as a whole remains, I conclude, rather speculative.

Let me pause to compare the relation of empirical equivalence, germane to the thesis of underdetermination, with other possible relations between theories. The tightest relation is that of logical equivalence: each axiom (and hence each theorem) of one theory is logically equivalent to an axiom or theorem of the other, or to a combination thereof, and the consequence relation is preserved. Logically equivalent theories are in fact different formulations of the same theory. The relation that Poincaré posits between the various geometries, which we can call translation equivalence, differs from logical equivalence insofar as there is a sense in which different geometries are incompatible. Although we can translate the terms of one geometry into those of the others, these geometries are still incompatible under any interpretation that assigns *the*

same meanings to corresponding terms. In other words, whereas for logi-
cally equivalent theories, every model of one is ipso facto a model of the
other, for translation-equivalent theories (that are incompatible in this
sense) *no* model of one is a model of the other. The possibility of finding
within one theory a model for another, incompatible, theory mandates
that at least some terms – for example, 'straight line' and 'distance' in
Poincaré's dictionary – receive different interpretations in the two the-
ories. Hence the term 'translation' is used here in a nonstandard way:
while the ordinary notion of translation preserves both truth and mean-
ing, in the case of translation-equivalence, we preserve truth at the cost
of meaning-change. Davidson often emphasizes that preserving truth is
a constraint on (ordinary) translation. Poincaré's example shows that it
may be insufficient.

Empirically equivalent theories yield the same predictions or entail
the same class of observation sentences, but need not be either logically
equivalent or translation equivalent. In general, though, it is impossible
to substantiate the existence of empirical equivalence in any particular
case unless the stronger relation of translation equivalence is established.
Indeed, Poincaré's claim that no experiment can compel us to accept one
geometry rather than another was based on his argument that empiri-
cal equivalence is guaranteed by translation equivalence. This notion of
translation equivalence is akin to what Glymour (1971) calls theoretical
equivalence, but theoretical equivalence, and the translation it invokes, is
anchored in the principles of a particular theory. According to the prin-
ciple of relativity, for instance, systems in uniform motion relative to each
other are equivalent and cannot be distinguished by experiment. Here
too, the descriptions deemed equivalent by the theory in question can be
'translated' into one another. It is desirable that (from the perspective of
the theory we employ) empirically equivalent states will also be theoreti-
cally equivalent. In other words, it is desirable that empirical equivalence
be anchored in theoretical equivalence, but this desideratum, as we will
see in chapters 2 and 3, is not always met.

With Kuhn (1962) and Feyerabend (1962), a new relation, incom-
mensurability, came into vogue. Prima facie at least, the incommensu-
rability thesis and Poincaré's conventionalism have much in common.
Seemingly incompatible theories, such as two different geometries in
the case of Poincaré, or Newton's and Einstein's physical theories in the
case of Kuhn and Feyerabend, are declared to be free of any real con-
flict with each other. In both these examples, the paradoxical situation is

explained by meaning variance – the same terms have different meanings in the seemingly incompatible theories. In both cases, moreover, a theory is seen as implicitly defining its terms, so that any change in theory is, ipso facto, a change in the meanings of the implicitly defined terms, and consequently, in what the theory is about. But whereas Poincaré builds his argument around translatability, Kuhn and Feyerabend focus on *un*translatability. According to Kuhn (Feyerabend), different paradigms (theories) are incommensurable precisely because they cannot be translated into each other. Going beyond traditional relativism, which sees *truth* as internal or context dependent, incommensurability implies that from the perspective of one paradigm (theory), the alternative is not simply false, but makes no sense at all. Whereas Poincaré addresses situations in which we obtain, via translation, an equally meaningful, though seemingly incompatible theory, in the Kuhn-Feyerabend examples, we have no way of establishing any inter-theoretical relation. And while translation equivalence is a well-defined relation that can be rigorously demonstrated and does not hold between just any alternative theories, incommensurability, based as it is on a declaration of impossibility, is much more widely applicable but hardly ever demonstrable.

One of the most forceful critiques of the incommensurability-cum-untranslatability thesis is due to Davidson, who questions its intelligibility. The picture it paints of numerous alternatives of which we are aware, but cannot make sense, is itself senseless, according to Davidson. I agree. There is, however, one aspect of Davidson's argument I find disturbing: his formulation of the problem is insufficiently fine grained to distinguish the Kuhn-Feyerabend argument from Poincaré's. As Davidson uses the term "conceptual relativity," it refers to Kuhn's predicament of "different observers of the world who come to it with incommensurable systems of concepts" (1984, p. 187), but also, more generally, to any case in which there is essential recourse to more than a single language or mode of description. The latter characterization covers the translation equivalence of Euclidean and non-Euclidean geometry, which, clearly, is not a case of incommensurability. Here is Davidson's formulation:

We may now seem to have a formula for generating distinct conceptual schemes. We get a new out of an old scheme when the speakers of a language come to accept as true an important range of sentences they previously took to be false (and, of course, vice versa). We must not describe this change simply as a matter of their coming to view old falsehoods as truths, for a truth is a proposition, and what they come to accept, in accepting a sentence as true, is not the same thing