Instrumentalism

(to appear in *The Philosophy of Science: An Encyclopedia*, ed. Jessica Pfeifer and Sahotra Sarkar. New York: Routledge, Inc.)

P. Kyle Stanford
Department of Logic and Philosophy of Science
University of California, Irvine
5100 Social Sciences Plaza
Irvine, CA 92697-5100
U.S.A.
stanford@uci.edu

Though John Dewey coined the term 'instrumentalism' to describe an extremely broad pragmatist attitude towards ideas or concepts in general, the distinctive application of that label within the philosophy of science is to positions that regard scientific theories not as literal and/or accurate descriptions of the natural world, but instead as mere tools or 'instruments' for making empirical predictions and achieving other practical ends. This general instrumentalist thesis has, however, historically been associated with a wide variety of motivations, arguments, and further commitments, most centrally concerning the semantic and/or epistemic status of theoretical discourse (see below). Unifying all these positions is the insistence that one can and should make full pragmatic use of scientific theories either without believing the claims they seem to make about nature (or some parts thereof) or without regarding them as actually making such claims in the first place. This entry will leave aside the question of whether the term 'instrumentalism' is properly restricted to only some subset of such views, seeking instead to illustrate the historical and conceptual relations they bear to one another and to related positions in the philosophy of science.

Loci Classici

Broadly instrumentalist sentiments concerning scientific theories have a remarkably long intellectual pedigree: indeed, Popper's famous critique (1963, Ch. 3) of the position as intellectually sterile counts Andreas Osiander (author of the unsigned Preface to Copernicus's On the Revolutions of the Celestial Spheres), Cardinal Bellarmino, and Bishop Berkeley as notable early defenders of the view (for criticism of Popper's historical claims, see Fine 2001), even while resisting Duhem's claim to find its historical antecedents in classical Greek thinkers. Furthermore (as Popper and others note) the more recently influential instrumentalism of Ernst Mach is rooted in a critique of Newtonian mechanics (and its concepts of absolute space, time and motion) strikingly similar to Berkeley's own. Mach also resembles Berkeley in embracing a radical phenomenalism, insisting that what is represented " behind the appearances exists only in our understanding, and has for us only the value of a memoria technica or formula" (1911, 49). He argues that laws of nature (e.g. Snell's Law) and theoretical hypotheses (e.g. the atomic hypothesis) are simply conceptual devices for the systematic classification, summary, organization, and coordinated expression and prediction of innumerable particular appearances (1893, 582f). Thus, Mach insists that theoretical concepts like 'atoms' are merely "provisional helps" and are ultimately to be dispensed with not because they seek unsuccessfully to describe a reality beyond appearances but rather because they successfully but only indirectly describe coordinated and systematized collections of experiences themselves.

The instrumentalist impetus familiar from more recent philosophy of science, however, is rooted more fundamentally in developments within physics at the turn of the century, and in the related logical, epistemic, and historical concerns about the status of scientific theories articulated by thinkers like Pierre Duhem and Henri Poincaré (see Worrall 1982). The progress of physical science had by this time begun to suggest that there might be quite genuine cases of differences between actual competing scientific theories that could not possibly be adjudicated by any straightforward appeal to empirical tests or observations. To use a famous example of Poincaré's (though not a case of actual competing theories), any set of measurements of the angles in a triangle marked out by appropriately oriented perfectly rigid rods can be accommodated by the assignment of any number of different combinations of underlying spatial geometries and compensating 'congruence relations' for the rods in question; if the sum of the angles differs from 180 degrees, for instance, one may either interpret the underlying geometry as Euclidean and conclude that the distance marked out by each rod varies with its position and/or orientation, or assume that the distance marked out by each rod remains constant and conclude that the underlying geometry of the relevant space is non-Euclidean. Poincaré's response to this problem of theoretical underdetermination was conventionalism; that is, he regarded such theoretical matters as the assignment of a particular physical geometry to space as matters of choice or convention to be decided on grounds of greatest convenience (see CONVENTIONALISM; POINCARE, HENRI). And this in turn implied, he suggested, the distinctively instrumentalist conclusion that the quite useful ascription of a particular geometry to space by a theory should not be construed as literally attributing anything (truly or falsely) to nature itself: "[T]he question: Is Euclidean geometry true?…has no meaning. We might as well ask if the metric system is true, and if the old weights and measures are false…One geometry cannot be more true than another; it can only be more convenient." (Poincaré [1905] 1952, 50)

Duhem also worried about underdetermination, albeit of a very different sort (see Stanford 2001), and added a further concern about the role played by idealizations in physical theories. In addition, both Duhem and Poincaré were deeply impressed by the long history of repeated and radical discontinuities in the dominant theoretical conceptions of particular domains of nature. But both argued that this history of scientific revolution and wholesale replacement is characteristic only of our efforts to "surmise realities hidden under data observable by the senses" (Duhem [1914] 1954, 274) which in fact "merely nam[e]…the images we substituted for the real objects which Nature will hide for ever from our eyes" (Poincaré [1905] 1952, 161). Thus, while both thinkers retained full confidence in the "experimental laws" or generalizations about observable phenomena uncovered by scientific investigations, each denied that such investigations were able to penetrate (or that "mathematical theories" described) the actual constitution of nature, and Duhem went so far as to consign the explanatory ambitions of theories to the realm of metaphysics rather than science.

Thus, both Duhem and Poincaré seemed to recognize scientific theories or theorists as often aspiring to describe an underlying, inaccessible reality and/or explain observable events by appeal to it and simply rejected these ambitions as ultimately either unscientific or unsatisfiable in some way. But both thinkers ranged at different times and in different works through a wide variety of importantly divergent attitudes (and not all the same ones between them) towards the cognitive, semantic, and epistemic status of theories: these attitudes included the view that extant scientific theories were not in fact making claims about inaccessible realities behind observable phenomena, that the scientific enterprise need not do so, and that it should not. Moreover, there is reasonable controversy over classifying either thinker as ultimately an instrumentalist in any of these straightforward senses: in his last work Poincaré whole-heartedly embraced the reality of atoms, while Duhem consistently held that scientific theories are able to establish "natural classifications" of the phenomena (for balanced discussion, see Psillos 1999, Ch. 2).

Even this brief excursion through instrumentalist themes in Mach, Duhem and Poincaré offers some sense of the variety of distinctive further commitments that have been conjoined with the general claim that scientific theories should be understood simply as useful instruments rather than accurate descriptions of inaccessible domains of nature. Among such further commitments are the suggestions that (i) theoretical discourse is simply a device for organizing or systematizing beliefs about observational experience, and its meaning is therefore exhausted by or reducible to any implications it has concerning observable states of affairs (reductive instrumentalism); (ii) theoretical discourse has no meaning, semantic content or assertoric force at all beyond the license it provides to infer some observable states from others (syntactic instrumentalism); (iii) even if such discourse is both meaningful and irreducible, it can nonetheless be eliminated from science altogether (eliminative instrumentalism); and (iv) even if the literal claims of theoretical science about the natural world are neither reducible, nor meaningless, nor even eliminable, such claims are nonetheless not to be believed (epistemic instrumentalism).

The Language of Science: Reductive, Syntactic, and Eliminative Instrumentalism

It is quite striking that even some of Duhem's and Poincaré's explicit reservations about scientific theories have a semantic or linguistic character: Duhem claims that "hypotheses [are] not judgments about the nature of things, only premises intended to provide consequences conforming to experimental laws" ([1914] 1954, 39) and that theoretical propositions "are neither true nor false…only convenient or inconvenient" ([1914] 1954, 334), while Poincaré adds that the "object of mathematical theories is not to reveal to us the real nature of things", but "only…to co-ordinate the physical laws with which experiment makes us acquainted..." ([1905] 1952, 211). Perhaps less surprising is the fact that such a generally linguistic or semantic strategy of analysis was appealing to logical positivist thinkers.

The early logical positivists' efforts to effect a reduction of all scientific language to a privileged phenomenological or observational basis (a project pursued most notably by the early Carnap, but also influentially by Bridgman) quite naturally grounded an instrumentalism about scientific theories of the sort described above as reductive (see BRIDGMAN, PERCY; CARNAP, RUDOLPH). The positivists' (and later, the logical empiricists') embrace of instrumentalist sentiments, however, was also motivated by the general suspicion that competing realist claims about the status or content of scientific theories themselves constituted meaningless verbiage of the "metaphysical" sort they sought to banish from science altogether. Furthermore, even after the reductive project came to be widely regarded as a failure and such thinkers had given up the notion that the semantic content of apparently theoretical discourse was 'really' exhausted by its implications concerning collections of observable events or subjective experiences, the distinctively syntactic variety of instrumentalism offered a fallback position. More specifically, some positivists suggested that theoretical claims are properly regarded as devoid of <u>any</u> semantic content whatsoever beyond the license they provide to draw inferences from one observable state of affairs to another. On this account, theoretical claims do not carry any straightforward ontological commitments regarding unobservable entities, even if they cannot be fully reduced to claims about immediately accessible experiences or states of affairs. And in the spirit of Duhem and Poincaré, this view held theoretical claims to be <u>non-assertoric</u>; that is, appearances to the contrary, they are not claims about what the world is like and do not possess truth values at all.

Of course, this somewhat counterintuitive view of the semantics of theoretical claims might be evaded by embracing the arguably more natural view (equally in the spirit of Duhem and Poincaré) that such theoretical discourse is simply <u>eliminable</u> from science altogether. This eliminative form of instrumentalism also gained considerable currency among positivist thinkers, especially following the formulation and proof of an influential theorem by William Craig. Craig's Theorem showed that for any recursively axiomatized first-order theory T, given any effectively specified sub-vocabulary O of T (mutually exclusive of and exhaustive with the remainder of the vocabulary of T), one can effectively construct another theory T' whose theorems are exactly those of T that contain no nonlogical expressions besides those in O. As Hempel was the first to realize, this theorem implies that if the nonlogical vocabulary of any given scientific theory is partitioned into theoretical and observational components, the theory can be replaced with a 'functionally equivalent' Craig-transform that preserves all the deductive relationships between observation sentences established by T itself, since (by Craig's

Theorem) "any chain of laws and interpretive statements establishing [definite connections among observable phenomena] should then be replaceable by a law which directly links observational antecedents to observational consequents" ([1958] 1965, 186). This implied in turn, Hempel noted, that theoretical terms could be eliminated from theories altogether without any loss in the purely observable consequences (deductively) obtainable from them, creating the following "Theoretician's Dilemma":

> If the terms and principles of a theory serve their purpose [of deductively systematizing the theory's observational consequences] they are unnecessary, as just pointed out; and if they do not serve their purpose they are surely unnecessary. But given any theory, its terms and principles either serve their purpose or they do not. Hence, the terms and principles of any theory are unnecessary. ([1958] 1965, 186)

The apparent feasibility of this eliminative instrumentalist program was further advanced by a related (and earlier, though largely unrecognized at the time) innovation of Frank Ramsey's: he proposed replacing any finitely axiomatized theory with a sentence that existentially generalizes on all the theoretical predicates of that theory (e.g. replacing $\forall x\mathrm{T}x$ with $\exists\phi\forall x\phi x$). This so-called "Ramsey sentence", he argued, has the same observational consequences as the original theory and therefore captures all the "factual content" of the original (Ramsey [1929] 1978).

The significance of Craig's theorem was, however, immediately controversial. Nagel, for instance, famously argued (1961, 136-7) that it is of quite limited relevance to the actual eliminability of theoretical discourse from science because i) there is no guarantee that the axioms of T' delivered by Craig's method will not be "so cumbersome that no effective logical use can be made of them"; ii) in fact, the axioms of T' will be infinite in number, no matter how simple the axioms of T itself, and correspond one-to-one with all the true statements expressible in the language of T', rendering them "quite valueless for the purposes of scientific inquiry"; and iii) Craig's method can actually be applied only if one knows, in advance of any deductions made from them, all the true statements in the restricted observational language. In addition, Glymour (1980, Ch. II) offers elegant technical objections to Ramsey's proposal, most importantly that as a theory of truth it fails to respect even the most elementary forms of demonstrative inference: the Ramsey sentence of $A \wedge B$ may be necessarily false, for example, while the Ramsey sentences of A and of B are individually true.

More recently, however, it is the profound differences between actual scientific theories and the sorts of artificial formal systems to which tools such as Craig's Theorem and Ramsey's technique apply which have led these formal results to be regarded as increasingly irrelevant to the genuine prospects for instrumentalism. More specifically, philosophers of science have become increasingly convinced that i) there is no strict, principled, or systematic division of the vocabulary of a theory into observational and theoretical parts; ii) the parts of a theory bear important logical, epistemic, and cognitive relations to one another that go far beyond what is captured by mere deductive systematization; and iii) scientific theories may not be best regarded as axiomatic formal systems in any case. Thus, at least part of the solution to the theoretician's dilemma, as Hempel himself recognized, is to reject the claim that the only function of theoretical terms is to deductively systematize a theory's observational consequences.

Credibility and Belief: Epistemic Instrumentalism

Even as the philosophical fortunes of these distinctive semantic and eliminativist theses have declined, interest has remained strong in the broader instrumentalist conception of theories as tools for pursuing practical ends rather than accurate descriptions of nature itself. The most influential recent approaches have pursued this conception by exchanging the positivists' reductive, syntactic, and eliminativist commitments for epistemic alternatives. That is, more recently influential forms of instrumentalism grant both the assertoric force and ineliminability of theoretical claims, but insist that such theories should simply be used for prediction of experimental outcomes and other practical goals without believing the claims they do in fact make about nature itself (or some parts thereof). A further recent trend has been to make the case(s) for instrumentalism piecemeal: arguing that quite specific features of a given scientific theory (e.g., quantum mechanics or evolutionary biology) either require or recommend an instrumentalist stance towards that particular theory.

One prominent example of general instrumentalism of this epistemic variety is Bas Van Fraassen's Constructive Empiricism (1980). Like Duhem and Poincaré, Van Fraassen appeals to the underdetermination of theories by evidence to challenge the conclusion that empirically successful scientific theories describe what inaccessible domains of nature are really like, and he insists that even a reflective endorsement of the actual inferential and other practices of science itself requires only a cognitive attitude of acceptance towards theories, rather than belief. He argues that it is epistemically supererogatory to believe any more of scientific theories than that they are empirically adequate, that is, that what they say about observable phenomena is true, and he insists that epistemic prudence recommends agnosticism regarding even the most successful theories' claims about unobservables. Thus, Constructive Empiricism regards scientific theories as reliable tools for anticipating how observables will behave while resisting the conclusion that such theories describe what unobservable domains of nature are really like, but on epistemic rather than semantic grounds.

Of course, Constructive Empiricism still relies fundamentally on an extremely controversial distinction between observables and unobservables, so it is important to note that the distinctively epistemic form of instrumentalism need not rely upon any such distinction: as Fine argues, the guiding commitment of instrumentalism is simply to the reliability of a causal story, which "treats all entities (observable or not) perfectly on par":

> Of course if the cause happens to be observable, then the reliability of the story leads me to expect to observe it (other things being equal). If I make the observation, I then have independent grounds for thinking the cause to be real. If I do not make the observation or if the cause is not observable, then my commitment is just to the reliability of the causal story, and not to the reality of the cause (1991, 86).

Perhaps the most fully developed form of epistemic instrumentalism that eschews any important distinction between observables and unobservables is the historically-oriented variety inspired by thinkers like Thomas Kuhn and pursued more recently by Larry Laudan. Like Duhem and Poincaré, these thinkers draw centrally on the history of repeated fundamental changes over time in the descriptions of nature offered by dominant scientific theories in support of a skeptical attitude towards the claims of the dominant scientific theories of the present day. Kuhn not only appeals to this history to undermine the notion that contemporary science is in possession of any final theoretical truth about a stable natural world, but also famously claims that the very "notion of a match between the ontology of a theory and its 'real' counterpart in nature now seems to me illusive in principle"; nonetheless he insists that scientific theories have improved over time "as instruments for puzzle-solving" ([1962] 1996, 206). Laudan argues (1981a) that the long historical record of successful but ultimately rejected scientific theories undermines any justification for inferring even the approximate truth of contemporary scientific characterizations of nature (observable or not) from their dramatic empirical successes. Nonetheless, he insists (1977, 1981b) not only that such theories can and should be used to tackle and solve a wide variety of empirical and conceptual problems, but also that there is a clear sense in which cumulative progress in this regard has been achieved over time, by attaining with the theoretical instruments of science an ever larger and more various set of effective solutions to such problems.

As these influential formulations of the view illustrate, epistemic instrumentalism seems committed to some distinction between believing a theory to be true and accepting or using it <u>without</u> believing what it says. Perhaps unsurprisingly, the cogency of this distinction has itself been the target of recent influential criticisms of epistemic instrumentalism, on the grounds that these cognitive attitudes simply cannot be distinguished in the way that one or more forms of instrumentalism requires. Horwich (1991) points out, for example, that some accounts of belief itself simply identify it as the mental state responsible for use, while Blackburn (1984) argues that there is no room for a distinction between merely "accepting" a statement with a truth-condition and simply believing it to be true (see also Fine 1986, esp. Sec. 4). By contrast, Sober (2002) defends the distinction, pointing out not only that idealized models known to be false are often accepted or used as the basis for accurate predictions across a range of phenomena, but also that recent work in model selection theory shows why models (statements containing adjustable parameters) known to be false will routinely serve as the basis for more accurate predictions of new data than competitors known to have higher likelihood conferred on them by the available data or even by competitors known to be true. Thus, he argues, not only is there a genuine difference between seeking instrumental or predictive reliability and seeking truth, but this distinction is respected within scientific practice itself, which typically seeks to maximize predictive accuracy when choosing models (with adjustable parameters) but seeks to identify the truth when choosing among fitted models (once parameters have been adjusted).

In a related vein, Nagel (1961, 139) famously argues that there is a "merely verbal" difference between the instrumentalist contention that a theory offers satisfactory techniques of inference and the realist contention that it is true. More recently, Stein (1989) has argued that the dispute between realism and instrumentalism is not well-joined: once realism has been sophisticated (as he suggests it must be) to give up its pretensions to metaphysically transcendent theorizing, to eschew aspirations to noumenal truth and reference, and to abandon the idea that a property of a theory might somehow explain its success in a way that does not simply point out the use that has been made of the theory, and once instrumentalism has been sophisticated (as he suggests it must be) to recognize the scope of a theory's role as an instrument to include not just calculating experimental outcomes, but also adequately representing phenomena in detail across the entire domain of nature and providing resources for further inquiry, there remains no appreciable difference (or no difference that makes a difference) between the two positions. That is, Stein argues for a convergence between the appropriately restricted ambitions a sophisticated realism holds out for theories and the appropriately expanded ambitions a sophisticated instrumentalism holds out for them. Furthermore, he argues that in the work of the deepest scientists (his examples are Maxwell, Newton, and Einstein) the two attitudes are present together in such a way that the alleged contradiction between them simply vanishes. Thus, even as instrumentalism persists as a viable and influential position in the contemporary philosophy of science, its comparative merits and even the coherence of its formulation remain the subject of deservedly intense controversy.

References

Blackburn, Simon. Spreading the Word. Oxford: Clarendon Press, 1984.

Duhem, Pierre. The Aim and Structure of Physical Theory. Translated from the second edition by Philip P. Wiener. Princeton: Princeton University Press, [1914] 1954.

Fine, Arthur. "Unnatural Attitudes: Realist and Instrumentalist Attachments to Science." Mind 95 (1986): 149-179.

_____. "Piecemeal Realism." Philosophical Studies 61 (1991): 79-96.

_____. "The Scientific Image Twenty Years Later." Philosophical Studies 106 (2001): 107-122.

Glymour, Clark. Theory and Evidence. Princeton: Princeton University Press, 1980.

Hempel, Carl. "The Theoretician's Dilemma: A Study in the Logic of Theory Construction." As reprinted in Aspects of Scientific Explanation and Other Essays in the Philosophy of Science. New York: Free Press, [1958] 1965.

Horwich, Paul. "On the Nature and Norms of Theoretical Commitment." Philosophy of Science 58 (1991): 1-14.

Kuhn, Thomas S. The Structure of Scientific Revolutions, 3d. Ed. Chicago, University of Chicago Press, [1962] 1996.

Laudan, Larry. Progress and it's Problems. Berkeley: University of California Press, 1977.

_____. "A Confutation of Scientific Realism." Philosophy of Science 48 (1981a): 19-49.

_____. "A Problem-Solving Approach to Scientific Progress." In Scientific Revolutions, edited by Ian Hacking, 144-155. New York: Oxford University Press, 1981b.

Mach, Ernst. The Science of Mechanics, 6th Edition. Translated by T. J. McCormack. La Salle, IL: Open Court, [1893] 1960.

_____. History and Root of the Principle of the Conservation of Energy. Translated by P. E. B. Jourdain. Chicago: Open Court, 1911.

Nagel, Ernst. The Structure of Science. New York: Harcourt, Brace and World, 1961.

Poincaré, Henri. Science and Hypothesis. Reprint of first English translation; originally published as La Science et L'Hypothèse (Paris, 1902). New York: Dover, [1905] 1952.

Popper, Karl. R. Conjectures and Refutations. London: Routledge and Kegan Paul, 1963.

Psillos, Stathis. Scientific Realism: How Science Tracks Truth. New York: Routledge, 1999.

Ramsey, F. P. "Theories." Reprinted in D. H. Mellor (ed.) Foundations: Essays in Philosophy, Logic, Mathematics and Economics. London: RKP, [1929] 1978.

Sober, Elliott. "Instrumentalism, Parsimony, and the Akaike Framework." Philosophy of Science 69 (Supplement) (2002): S112-S123.

Stanford, P. Kyle. "Refusing the Devil's Bargain: What Kind of Underdetermination Should We Take Seriously?" Philosophy of Science 68 (Supplement) (2001): S1-S12.

Stein, Howard. "Yes, but…Some Skeptical Remarks on Realism and Anti-Realism."

Dialectica 43 (1989): 47-65.

Van Fraassen, Bas. The Scientific Image. Oxford: Clarendon Press, 1980.

Worrall, John. "Scientific Realism and Scientific Change." Philosophical Quarterly 32 (1982): 201-131.

Acknowledgements

My thanks to Arthur Fine, Bas Van Fraassen, Bill Demopoulos, Elliott Sober, Larry Laudan, David Malament, Aldo Antonelli, Jeff Barrett, Stathis Psillos, Philip Kitcher, and the editors for helpful discussion and suggestions. This material is based upon work supported by the National Science Foundation under Grant No. SES-0094001. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation (NSF).