1 Playing with fire

Although an innovative astronomer and an important contributor to the development of planetary science, the late Carl Sagan is probably best remembered among the general public for two of his other activities: his popularization of contemporary natural science (especially astrophysics) and his highly public and unapologetic condemnation of “pseudoscience” concerning crystals, ESP, and alien abductions. The two activities fit together quite well, as they are united by a commitment to spreading a particular sensibility out beyond professional specialists and into the wider community. In a collection of essays entitled *The Demon-Haunted World*, Sagan borrows a metaphor from Thomas Ady’s 17th-century tract condemning witch hunts to describe his public and popular work as an effort to shine an illuminating light into the dark corners of the contemporary world: to light a candle in the hopes of banishing the shadows. The candle he sought to light and to wield against the darkness was what he called *science*:

> In science we may start with experimental results, data, observations, measurements, “facts.” We invent, if we can, a rich array of possible explanations and systematically confront each explanation with the facts. In the course of their training, scientists are equipped with a baloney detection kit. The kit is brought out as a matter of course whenever new ideas are offered for consideration. If the new idea survives examination by the tools in our kit, we grant it warm, although tentative, acceptance. If you’re so inclined, if you don’t want to buy baloney even when it’s reassuring to do so, there are precautions that can be taken.
>
> (Sagan 1997, 209–210)

Sagan’s account of the mechanics of science is probably fairly familiar to us, as it tracks quite closely with the notion of “falsification” famously propounded by Karl Popper (1992): science, in Popper’s formulation, proceeds and progresses through successive efforts to *dis*prove conjectures, rather than through efforts to verify or justify them. But Sagan’s metaphor—science as a candle in the darkness—should be scarcely less familiar, drawing as it does on a longstanding tradition in the philosophy of knowledge that equates knowing with seeing, and reason—often exemplified by science—with a source of light. Famously, John

Locke drew on this metaphor in his *An Essay Concerning Human Understanding*, admonishing his readers to use their natural faculties of reason to the best of their ability: "It will be no excuse to an idle and untoward servant, who would not attend his business by candle light, to plead that he had not broad sunshine. The Candle that is set up in us shines bright enough for all our purposes" (Locke 1959a, 30). Further, Locke deployed the notion of reason as a defense against popular deception in a manner quite reminiscent of Sagan's stance:

> Reason is natural revelation, whereby the eternal Father of light and fountain of all knowledge, communicates to mankind that portion of truth which he has laid within the reach of their natural faculties: revelation is natural reason enlarged by a new set of discoveries communicated by God immediately; which reason vouches the truth of, by the testimony and proofs it gives that they come from God. So that he that takes away reason to make way for revelation, puts out the light of both, and does much what the same as if he would persuade a man to put out his eyes, the better to receive the remote light of an invisible star by a telescope.
>
> (Locke 1959b, 431)

Setting aside the language of divinity for a moment, we can see a clear continuity between Locke and Sagan. Both point to a natural faculty that can be developed and deployed against error, and both symbolically equate that faculty with "light"—and oppose it to the "darkness" of misconception and superstition. Similarly, both privilege science as a superior way of gaining and evaluating knowledge—Sagan uses the term "science," while Locke, preferring the term "reason," explicitly associates himself and his argument with great scientists of the day such as Newton and Boyle. Whatever else it is good for, science appears in their conception as our best defense against error.

Of course, such arguments are not only advanced by philosophers and astronomers. Closer to home, as it were, David Laitin (2003, 169) advances a very similar image of science—including social science—as containing "ample procedures for figuring out if our best judgments are misplaced" and hence serving as "the surest hope for valid inference." Laitin pairs this declaration with a denunciation of Bent Flyvbjerg's *Making Social Science Matter* (2001) for allegedly violating the strictures of science and opening the door to a kind of anything-goes relativism—the ultimate nightmare about what the abandonment of the ground of "science" might mean in practice.[1] And in their popular and oft-cited methods handbook, Gary King, Robert Keohane, and Sidney Verba flatly declare: "research designed to help us understand social reality can only succeed if it follows the logic of scientific inference" (King, Keohane, and Verba 1994, 229). The juxtaposition of science and (potential) error, therefore, seems just as prominent in our field as it is in other domains.

Arguments such as these pose extremely fundamental questions about the character of our scholarly enterprise. Scholars of politics who advance such claims are quite clearly drawing on the cultural prestige associated with the notion of "science" in the contemporary age (Litfin 1994) as part of an effort to shape the practices of their colleagues involved in the effort to produce knowledge about the social world. To invoke "science" is to call to mind a panoply of notions connected with truth, progress, reason, and the like—and, perhaps more importantly, to implicitly reference a record of demonstrated empirical success. Appeals such as this function this way particularly in internal debates among scholars of the social world, as tossing an appeal to "science" into such debates is like playing a very valuable trump-card that implicitly, if not explicitly, calls the entire status of the scholarly field into question. Within the field of International Relations (IR)[2] in particular, the "science question" has long vexed scholars, coming to a head in the field's second "great debate" between self-identified traditionalists and scientists (Knorr and Rosenau 1969) but never really getting resolved or losing its scholarly resonance (see the discussion in Kratochwil 2006). Especially under such circumstances, it is impossible to invoke the notion of "science"—let alone to propose turning to either the practice or the philosophy of science in an effort to clarify or improve our own scholarship!—in any kind of purely typological manner. Playing the science card raises the stakes.

The science question in IR

It is important to note at the outset that the role played by "science" in our field is at least conditionally, if not completely, independent of any detailed philosophical or conceptual sense afforded to the term. In debates about the proper conduct of IR scholarship, we typically operate with caricatures and generalities rather than precise specifications, speaking loosely of "*the* scientific method" or "*the* philosophy of science" as though either of those two things actually existed. Although there have been some notable exceptions in recent years, most references to and invocations of "science" seem to operate with an image of knowledge-production that is a curious amalgamation of Sagan's skeptical "baloney detection kit," an embrace of mathematical formalism, and a desire for law-like generalizations that hold true across cases (given appropriate scope conditions, of course). This is a curious amalgam because the first defines a skeptical *attitude*, the second defines a formalist *method*, and the third defines an epistemic *goal*—and none of these are perfectly characteristic of any actually existing scientific practice. In debates about knowledge-production in our field, what is most often in play is not a specific account of science, but a vague and general sensibility.

Of course, this is in no way just a comment on the present state of the field. Throughout the history of IR, the term "science" has been flung around in extremely cavalier ways, standing-in generally as the positive pole of a contrast that an author wishes to draw between her or his approach to generating and evaluating claims about world politics and some reviled alternative. For example:

> This book has two purposes. The first is to detect and understand the forces that determine political relations among nations, and to comprehend the ways in which those forces act upon one another and upon international political

> relations and institutions. In most other branches of the social sciences this purpose would be taken for granted, because the natural aim of all scientific undertakings is to discover the forces underlying social phenomena and the mode of their operation.
>
> (Morgenthau 1985, 18)

Thus Hans Morgenthau claimed early in his textbook *Politics Among Nations*, characterizing his approach as a "scientific undertaking" with little more than a vague gesture in the direction of "forces underlying social phenomena." There is no more specific discussion of the character or value of science in the book, although Morgenthau generally takes it for granted that only a scientific study can provide the basis for a responsible pursuit of a peaceful world; that, indeed, is the second "purpose" of his book (ibid., 20). The general notion or idea of "science," and the cultural prestige associated with it, suffices to legitimate Morgenthau's enterprise.

Morgenthau was very aware of this cultural prestige, having railed at length against the over-scientizing of the contemporary age in his 1946 masterpiece *Scientific Man vs. Power Politics*:

> Politics is an art and not a science, and what is required for its mastery is not the rationality of the engineer but the wisdom and the moral strength of the statesman . . . The age has tried to make politics a science. By doing so, it has demonstrated its intellectual confusion, moral blindness, and political decay.
>
> (Morgenthau 1946, 10)

The problem, Morgenthau argued, is that we put too *much* stock in science, and thus overlook the distinctiveness of the political and social world. In his typically Weberian fashion, Morgenthau argued that we make a category mistake when we expect science to solve our political problems; instead, we should respect the limits of human knowing, and keep science in its place. "For the liberal, science is a prophecy confirmed by reason; for the conservative, it is the revelation of the past confirmed by experience" (Morgenthau 1946, 32). Casting himself on the "conservative" side of the ledger, Morgenthau engaged in a very interesting double intellectual operation: on one hand, criticizing the over-reliance on science, but on the other hand, claiming some of its cultural prestige for his own project of knowledge-production. The result, whether by accident or by design, is the simultaneous preservation of the notion that we ought to have "scientific" knowledge of world politics, along with a good deal of ambiguity about precisely what that means in practice.

In pursuing this line of argument, Morgenthau was simply following the precedent laid down by E.H. Carr in *his* announcement of a scientific study of world politics. Carr talked about science, but never precisely defined the term except to contrast science with both unchecked idealism and unchecked realism (Carr 2001, 87). The science Carr announced would avoid both of those political-partisan stances, instead aiming for a more comprehensive view. But the scientific study of world politics, Carr acknowledged, would not be a simple transplantation of procedures from the natural sciences:

> The laboratory worker engaged in investigating the causes of cancer may have been originally inspired by the purpose of eradicating the disease. But this purpose is, in the strictest sense, irrelevant to the investigation and separable from it. His conclusion can be nothing more than a true report on facts. It cannot help to make the facts other than they are; for the facts exist independently of what anyone thinks about them. In the political sciences, which are concerned with human behavior, there are no such facts. The investigator is inspired by the desire to cure some ill of the body politic. Among the causes of the trouble, he diagnoses the fact that human beings normally react to certain conditions in a certain way. But this is not a fact comparable with the fact that human bodies react in a certain way to certain drugs. It is a fact which may be changed by the desire to change it . . . The purpose is not, as in the physical sciences, irrelevant to the investigation and separable from it: it is itself one of the facts.
>
> (Carr 2001, 4–5)

This does not tell us much about what it *means* for something to be a science. Indeed, Carr's claim is quite difficult to elucidate, because it is unclear just what is "scientific" about *both* a report on facts that are independent of human recognition *and* a report on facts that can be changed by the desire to change them—and Carr gave his readers little explicit guidance on this issue. Neither did Morgenthau, who similarly claimed that "social conditions" are more closely interwoven with scientific inquiry in the social sciences (Morgenthau 1946, 162). Both of these seminal IR scholars were quite confident that the study of world politics can and should be a "scientific" one, but it was not a central concern of either author to spell out precisely what it means for a study to be scientific. Instead, both were content simply to invoke the notion of "science" in the course of justifying their approaches.

Matters became more specific with the next of the field's "great debates"—a controversy "over the merits of the traditional and scientific approaches to the study of international politics," in which the main protagonists were Hedley Bull, arguing for tradition, and a diverse cast of characters arguing for science (Knorr and Rosenau 1969, iii). Bull characterized the opposition between these two approaches as mostly a matter of style and technique, with the traditional approach emphasizing "judgment" derived from an intimate experience with the history and philosophy of politics, and the scientific approach aspiring "to a theory of international relations whose propositions are based either upon logical or mathematical proof, or upon strict, empirical procedures of verification" (Bull 1969, 20–21). That this was largely a tactical difference became clear with Bull's declaration that:

> The theory of international relations should undoubtedly attempt to be scientific in the sense of being a coherent, precise, and orderly body of knowledge, and in the sense of being consistent with the philosophical foundations of modern science. Insofar as the scientific approach is a protest against slipshod thinking and dogmatism, or against a residual providentialism, there is everything to be said for it.
>
> (ibid., 36)

In this broad sense, Bull's definition of science was strikingly similar to that of Carr or Morgenthau. What he objected to were quantitative and formal techniques, and the drive towards generalization—precisely the features privileged and defended by self-identified "scientists" such as J. David Singer and Marion Levy. Levy was quite clear that "a generalized system of theory . . . hopefully with deductive interdependencies among the members of the set" (Levy 1969, 92) is the ultimate goal of any science, and he agreed with Singer that "we will never build much of a theory, no matter how high and wide we stack our *beliefs*" (ibid., 71)—the conduct of science means moving beyond beliefs and evaluating those beliefs in the light of systematic empirical evidence. In this debate, scientists took traditionalists to task for simply resting, content with their intuitions; traditionalists took scientists to task for their remoteness from the subject-matter.

But all sides of the debate agreed that the point of studying world politics is to produce empirically grounded and justified claims. This made the controversy a disagreement about the relative contribution of general propositions and hypothetical models, on one hand, and detailed historical reconstructions, on the other, to the understanding of world politics. Read in this way, the debate featured much less of an unbridgeable divide than might have at first appeared: everyone wanted to be "scientific" in the broad sense, and to produce coherent and orderly knowledge, but they disagreed as to which techniques were actually "scientific" in the relevant sense. However, it is significant that this was *not* Bull's rhetorical strategy; instead of defining and defending a broad account of science against the more elaborate and specific account advanced by his (largely American) opponents, Bull in effect *conceded* the notion of "science" to his opponents and took his stand elsewhere. The fact that Bull's broad definition of science is buried within the sixth of his seven critiques of formalist quantification and the quest for general propositions indicates something of how far it was away from the main thrust of his argumentative strategy.

Thus, the actual result of the "second great debate" in IR was to link "science" with quantification, formal models, and general propositions, replacing Carr and Morgenthau's vague notion of science with something more precise while retaining the cultural prestige of the notion. Singer, Levy, and other self-identified "scientists" made numerous references to the successes of physics and economics, holding out hope that IR could enjoy similar successes by becoming equally "scientific." The editors of the volume containing many of the important essays constituting the controversy even pioneered a strategy of reconciling the two approaches under a common banner, a strategy that further reinforced the equating of "science" with the formulation of general propositions:

> [W]hy could not the traditionalists take on the burden of casting their conclusions in the form of hypotheses testable in other situations? This would not undermine their inquiries, but it would maximize their possible contribution to the work of their more scientific colleagues. Likewise, why could not the scientists append summaries to their studies that straightforwardly identify their major propositions and findings? Such additions would not jeopardize their procedures, but they would make the products of their research more accessible to those who prefer nonscientific modes of inquiry.
>
> (Knorr and Rosenau 1969, 18)

Notice that, in this passage, the main "burden" falls on the traditionalists, who have to adopt a form of presentation that makes their claims ready for evaluation by the techniques preferred by self-identified "scientists." The only thing that the "scientists" have to do, apparently, is to produce a plain-English account of their study—a communicative, rather than a methodological, modification. Testable hypotheses and general claims are thus portrayed as almost unquestionable goals of IR scholarship, hardly even needing the label "science" to distinguish them from alternatives. But the label continues to serve a useful function in reaffirming the status of those fundamental assumptions—as when, a quarter-century later, King, Keohane, and Verba declared that "the social science we espouse seeks to make descriptive and causal inferences about the world" (King, Keohane, and Verba 1994, 7) and passed quite seamlessly from that claim to a series of discussions about strategies for testing hypothetical generalizations.

In fact, "science," in IR, has come to mean more or less precisely what Bull's opponents asserted that it meant, and the historical controversy between the traditionalists and the scientists has been recoded or reconceptualized as a dispute about styles of presentation or argumentation. "'Science' versus 'tradition'" has morphed into "'quantitative' versus 'qualitative'," a characterization that effectively strips any fundamental philosophical or conceptual issues out of the dispute (Yanow and Schwartz-Shea 2006, xv–xix). Knorr and Rosenau noted this at the time of the initial debate:

> Why, then, could not the traditionalists employ rather than deplore the quantitative findings of the scientists, refining them as seems suitable to their own way of thinking? And why could not the scientist use rather than abuse the qualitative insights of the traditionalists, subjecting them to the rigors of their procedures in the same way they do their own ideas?
>
> (Knorr and Rosenau 1969, 18)

While it remains a bit unclear how traditionalists uninterested in general propositions might "employ" quantitative findings, the idea that a "scientist" could take a traditionalist's conclusion or insight and subject it to procedures of hypothesis testing (especially if the traditionalist had followed their advice to state the insight in the *form* of a testable hypothesis, thus relieving the "scientist" of any conceptual labor of translation) is both a well-defined intellectual operation and a clear example of the priority accorded to "science" understood as the quest for generalized theoretical knowledge. That this priority of general propositions over insight based on intimate familiarity with particular situations persisted can be seen in King, Keohane, and Verba's more recent suggestion that "nonstatistical research will produce more reliable results if researchers pay attention to the rules of scientific inference—rules that are sometimes more clearly stated in the style of quantitative research" (King, Keohane, and Verba 1994, 6). This applies above all to "qualitative" studies, where researchers can only guarantee their "scientific" status by seeking to distinguish systematic from nonsystematic components of a situation even in their descriptions of that situation (ibid., 56). Every scholarly practice, then, is to be subordinated to the specific notion of "science" established as dominant in the discipline during the debate with Hedley Bull.

Of course, this outcome was somewhat foreshadowed by Bull's own confused position about science (Kratochwil 2006, 9). Because Bull failed to articulate a clear *alternative* to systematic generalization across historical cases, for example, he opened his position up to the rejoinder that there was no compelling reason *not* to subject the results of a detailed empirical-historical account to broader evaluation. Especially since this technique seemed to have proven so helpful in other fields of inquiry, the argument in favor of the "scientists" appeared almost unassailable. In practice, the most prominent dissenters focused more on pointing out the shortcomings of the "scientific" position than on elucidating a concrete alternative, calling for greater reflexivity among scholars (Lapid 1989) or affecting a whole-scale turn towards political and normative theory (Connolly 1989). Critics of generalized theoretical systems, such as Richard Ashley (1983; 1984), followed in Bull's footsteps by leaving the notion of "science" itself untouched in the field and permitting the self-proclaimed "scientists" to continue their monopoly on defining the term.

This strategy was evident even in the most successful effort to garner some "thinking space" (George and Campbell 1990) in the field for empirical scholarship not particularly interested in the formulation and evaluation of theoretical generalizations. Martin Hollis and Steve Smith's *Explaining and Understanding International Relations* was one of the first books to elucidate cogently a form of empirical knowledge-production that was not simply a deficient or low-tech version of the hypothesis testing/generalization approach. Hollis and Smith began with the delineation of two "intellectual traditions" animating the production of empirical knowledge in the social sciences: one derived from the natural sciences and the other derived from nineteenth-century hermeneutics. "Explaining" designates the first approach; "understanding," the other. Hollis and Smith then quickly proceeded to draw a series of other distinctions that map onto this same basic division: "outsider" versus "insider" accounts, causes versus meanings, and preferences versus rules (Hollis and Smith 1990, 1–7). The authors argued that these two bundles—causal outsider accounts using preferences to explain what actors do in world politics, and meaningful insider accounts using social rules to understand what actors do in world politics—were virtually incommensurable, leaving us with a situation in which there are always two separate stories to tell about any given empirical situation. The authors were also meticulous in avoiding any kind of comparative analysis of the two approaches, concluding the book with a dialogue between themselves that highlights the strengths and shortcomings of each approach in terms of the other (ibid., 203–214).

The clear implication of the Hollis and Smith depiction of empirical inquiry in IR was that "scientists" did not have a monopoly on knowledge-construction; there was an established, vibrant tradition operating with very different assumptions about how knowledge ought to be produced, and it was in some sense equal in value to its "scientific" alternative. The argument established a diversity of modes of inquiry, but at a fairly significant cost. "Explanation," rooted in "the attempt to apply the methods of natural science to the world of international relations" (ibid., 45), received causation and preferences, while "understanding" was left with the explication of social rules and the delineation of the motives of actors[3]—a stance that, incidentally, left many understanding-accounts vulnerable to critiques that they were actor-reductionist or perhaps even idealist.[4] More to the point, the Hollis and Smith strategy allowed the self-proclaimed "scientists" to continue to claim both the centuries-old tradition of the natural sciences *and* the cultural prestige associated with that tradition. Partisans or practitioners of "understanding" had no such proud parentage to claim, but instead had to be content with a bevy of German philosophers and British anthropologists.

From this potted history of some key debates in the field of IR, I would like to draw two conclusions. First, "science" has been a notion in play in IR debates since the very beginning of the scholarly study of world politics. Indeed, we could easily go back *before* the establishment of the study of world politics as a distinct scholarly endeavor and find "science" playing an important role in debates about the status of international law (Schmidt 1998, 104–106) and in the efforts of scholars of politics to distinguish themselves and their work from purely partisan political activity in the very early part of the twentieth century (Adcock 2003, 501–506)—to say nothing of the continuing role played by "science" in the shaping of the discipline of Political Science, within which so much of IR scholarship is located (Gunnell 1993). For the moment, it is sufficient to note that the shapers of the field of IR have been concerned about the scientific status of their scholarship for a very long time. Because of this long-standing history, "science" remains a notion to conjure with in the field of IR; it is a veritable "rhetorical commonplace" (Jackson 2006, 27–32), which is available for deployment within all kinds of controversies. And a powerful resource it is, too: charging that a piece of work is not "scientific" carries immensely negative connotations, both because of the field-specific history I have sketched here and because of the broader cultural prestige enjoyed by "science" (Moses and Knutsen 2007, 155–156).

This leads to my second conclusion: the *function* of the commonplace "science" within IR is primarily a *disciplining* function. When "science" makes an appearance, it is a pretty good bet that the text in which the term is invoked is more or less explicitly trying to reshape how inquiry is conducted, and doing so by drawing on the rhetorical power of "science" in order to privilege some modes of inquiry at the expense of others. If "science" is a good and valuable thing, then non-"science" cannot be as worthwhile an endeavor. Simply rejecting "science," or elaborating an alternative such as "understanding," leaves the whole discursive arrangement intact, and does not really offer a reasonable or effective rejoinder to the charge that the non-"scientific" work that one is doing is not somehow of lesser value. There is no effective way around this unless the whole field abandons any claims to or aspirations of being scientific. Absent this unlikely possibility, the question of science remains almost unavoidable for IR scholarship.

The demarcation problem

Philosophers of science sometimes refer to the "science question" as the *demarcation problem*: the quest for a set of criteria that can adequately demarcate science from non-science. "Adequately" here generally means something more profound than the disciplining deployment I have been discussing; philosophers working on the demarcation problem are looking for defensible logical or conceptual criteria, powerful enough that their application to a given scholarly controversy will yield a philosophically valuable determination of the scientific status of a given claim or position or approach, and help to explain the success of that science. Such philosophical work does, of course, draw on the cultural prestige of the commonplace "science," but seeks to give content to that label such that the claim to be "scientific" might rest on firm foundations rather than on a vague appreciation for modern technological marvels such as the computer or the airplane.

Inasmuch as philosophical elaborations of demarcation criteria are based on detailed study of successful (and sometimes unsuccessful) sciences, a philosophical solution to the demarcation problem would provide an answer to the question of how IR ought to proceed as a scientific field. In fact, until very recently, the most prominent use of philosophy of science in IR has been precisely along these lines and has featured efforts to spell out concrete steps that need to be undertaken in order to make IR more, or more properly, scientific. The basic structure of the argument is quite simple: according to some philosopher, successful science *S* engages in scientific practices $sp_1 \ldots sp_n$; we want IR to be a science too; ergo, we ought to engage in $sp_1 \ldots sp_n$ in IR. Elaborating such sets of practices by referring to something that is rather uncontroversially a science, such as evolutionary biology (Bernstein et al. 2000) or paleontology (Van Belle 2006), implicitly invokes a set of demarcation criteria that both define the science in question as a science, and encompass the subject matter of IR in such a way that practices the author identifies in one domain can be easily transported into the other domain. The uncontroversial identification of the "scientific" domain *as a science* spares the person making the argument from having to spell out explicitly just what it is that defines something as a science: we know it when we see it, after all, and if something works in physics or in paleontology it ought to work in IR, right?

The problem, of course, is that without a clear explication of the criteria that make a given practice of knowledge-production scientific, we have no good way to answer that question. Maybe there is something specific about, say, the empirical domain of physics that enables it to be uniquely scientific in a way that simply will not work if applied to the study of human beings and their social relations. Or maybe different approaches to knowledge-production have their own internal standards and practices, such that trying to apply techniques and procedures from one domain to another is nonsensical at best and harmful at worst. It is impossible to make a decision about matters such as this without a much clearer and more precise elaboration of what a science *is*, which is where philosophers of science might enter the picture. If philosophers agreed on a set of criteria that served to demarcate science from non-science, then we would have a defensible basis on which to examine claims about particular ways in which knowledge-production practices in IR ought to be disciplined.

Unfortunately, philosophers have come to no global consensus about what defines a field of inquiry as a "science" or a practice of knowledge-production as "scientific." Even worse, different attempts to determine such criteria proceed in wildly divergent directions and elucidate incompatible or contradictory positions on the importance of logical consistency, empirical observability, and predictive accuracy (among other criteria) to a compelling definition of science. Under these circumstances, a turn to the philosophy of science is unlikely to be able to put an end to the science question in IR by resolving the issue once and for all.

The roots of the traditional demarcation problem in the philosophy of science go back to the early twentieth-century "logical positivists" of the Vienna Circle. Confronted with Marx, Freud, Einstein, and a whole slew of theories about racial and national "destinies," the logical positivists sought to elucidate a foolproof way to distinguish between a scientific and a non-scientific statement. Besides being an interesting intellectual puzzle, the scientific status of a claim was also a pressing political and social problem: it mattered a great deal whether a denunciation of the received wisdom about sexuality, time, space, or governmental authority should be considered "scientific" and thus worthy of respect, or unscientific and hence intellectually valueless (Moses and Knutsen 2007, 38–39; Lakatos 2000, 22–24). The logical positivists' major criterion for distinguishing a scientific from a non-scientific claim was *verifiability*, which maintained that a claim could only be scientific if all of its terms could be checked or confirmed through an examination of the empirical world. The verifiability criterion would rule out claims involving "'entelechy' in biology, 'historical destiny of a race' or 'self-unfolding of absolute reason' in history," because they were not verifiable—but were instead "mere metaphors without cognitive content" (Hempel 1965b, 237).

However, the verifiability criterion also raised problems for notions such as "force" or "cause," which had long been staples of natural-scientific work. Indeed, a sensibility in many ways quite akin to that of the Vienna Circle led Ludwig Wittgenstein to banish causality from the scientific lexicon altogether: "There is no compulsion making one thing happen because another has happened. The only necessity that exists is *logical* necessity" (Wittgenstein 1961, §6.37). In general, logical positivists preferred to speak of a nomological explanation of an event, "showing that its occurrence could have been inferred . . . by applying certain laws of universal or of statistical form to specified antecedent circumstances" (Hempel 1965c, 302). Causality was thus redefined to mean a law-like relationship between phenomena. But this only displaced the problem, because *law-like claims are not verifiable*. All that exists, empirically, are specific objects and entities inhabiting particular situations, and if we were to confine ourselves strictly to what we can verify we could not say with certainty that, for instance, "books fall to the floor when dropped." All that we could say would be that *this* book fell to the floor when dropped, and *that* book fell to the floor when dropped, and so on ... and we would never reach a law-like statement about books and floors *in general*, no matter how many books we dropped. Rewriting the law-like statement so that it was only probabilistic would not solve the problem, inasmuch as a gap would still remain between "books have been observed to fall quite often to the floor when dropped" and "books quite often fall to the floor when dropped."

Of course, this was a known issue. David Hume had made a similar point over a century earlier:

> All inferences from experience suppose, as their foundation, that the future will resemble the past, and that similar powers will be conjoined with similar sensible qualities. If there be any suspicion, that the course of nature may change, and that the past may be no rule for the future, all experience becomes useless . . . In vain do you pretend to have learned the nature of bodies from your past experience. Their secret nature, and consequently, all their effects and influence, may change, without any change in their sensible qualities. This happens sometimes, and with regard to some objects: Why may it not happen always, and with regard to all objects? What logic, what process of argument secures you against this supposition?
>
> (Hume 1977, 24)

Logical positivists worried extensively about this problem and designed increasingly sophisticated ways to try to get around it,[5] but they all floundered on the same basic conceptual gap between particular observations and law-like claims. And this, in turn, would mean that no law-like claim was scientific, because no means could be found for verifying it.

Karl Popper's solution to these logical problems involved an inversion of the basic stance of the logical positivists: since law-like claims could never be verified, and since scientific claims were phrased in law-like—often universal—terms, perhaps it made sense to stop asking whether a claim could be proven *true* and instead ask whether a claim could be proven *false* (Popper 1992, 92). If a law-like claim were treated as a hypothetical conjecture instead of being regarded as the logical endpoint of a process of empirical observation and inductive reasoning, the conceptual gap between general laws and particular observations could be subsumed under the procedure of *falsification*: instead of vainly trying to assemble enough particulars to ground a law, a researcher could instead toss a law-like conjecture out into the world and then use particular observations to try to disprove it (Popper 1979, 29–30). This, in turn, suggested a different demarcation criterion for scientific claims: instead of being verifiable, they should be falsifiable. Indeed, Popper even added the requirement that the conditions under which a claim would be disproven should be stated in advance of conducting any empirical research; if one could not state such criteria, then one did not have a scientific claim.

The Popperian criterion of falsifiability enjoys a great deal of support, especially among practicing scientists—charges that some claim or piece of research is "unfalsifiable" are often used in a transparently disciplining manner, to exclude that claim or piece of research from serious consideration (Taylor 1996, 30–31). The idea that claims must be testable through the collection of empirical evidence has, to some extent, become commonsensical in many discussions of science, taken for granted to the point that an explicit defense of the idea is not considered to be necessary. For example, in debates about evolution and "creation science," one regularly sees each side accusing the other of holding onto their core assumption in defiance of the available evidence, and thus not adhering to the principle of falsifiability (Beil 2008); but nowhere in those debates will one find a *defense* of falsifiability as a criterion demarcating science from non-science. Instead, debate using the Popperian criterion revolves around the two behavioral implications of the falsifiability principle: researchers should be actively trying to falsify their conjectural claims, and only tentatively and provisionally accepting claims that survive a more or less rigorous series of tests; and researchers should abandon claims that have been falsified, because knowledge only expands if discredited propositions are discarded. Hence the focus of evaluation shifts from claims themselves (as long as they are falsifiable) to the behavior of the communities of researchers working with them, and science ceases to be a purely logical endeavor—it is, rather, a *practical* one.

One problem with falsifiability, however, is that it does not appear to work very well even when applied to established sciences such as physics. That was the chief empirical argument of Thomas Kuhn, who spent a lot of time observing the actual history and practice of science when writing his classic book *The Structure of Scientific Revolutions* (Kuhn 1970b). He discovered that practicing physicists do not, in fact, spend a lot of time attempting to falsify foundational claims about the world. In fact, they seem to take a lot of claims for granted in the conduct of their everyday research work, and when confronted with results that would appear to call into question those foundational claims, they were more likely to creatively reinterpret the results (for instance, by postulating an exogenous intervening factor) than simply to abandon their claims. Kuhn argued that acceptance of these foundational claims was, in fact, the precondition of scientific work:

> When engaged with a normal research problem, the scientist must *premise* current theory as the rules of his game. His object is to solve a puzzle, preferably one at which others have failed, and current theory is required to define that puzzle and to guarantee that, given sufficient brilliance, it can be solved.
> (Kuhn 1970a, 4–5)

"Normal science," as Kuhn defined it, was characterized by puzzle-solving, not by ongoing efforts to falsify any and all conjectures and claims. Actual scientists did not, in practice, adhere to the behavioral implications of falsifiability; hence there was either something wrong with the principle of falsifiability, or with the practice of science itself. Kuhn preferred the former; Popper, in a rather striking contrast to his own principle of falsifiability, stuck to his claim in defiance of the empirical evidence about scientific practice, claiming that Kuhn's normal scientist "has been badly taught" and "is a victim of indoctrination" rather than possessing a properly critical intellect (Popper 1970, 53).

In a way, the disagreement between Kuhn and Popper about what constitutes science illustrates another difficulty involved in attempting to implement the principle of falsifiability in the first place. Take a (Popperian) statement such as "science is characterized by the making of bold conjectures and the attempt to falsify them," and confront it with evidence that practicing scientists do not, in fact, behave in this way; what is the result? Perhaps the statement is rejected because of the discrepant evidence, but perhaps the statement's author questions the accuracy of the potentially falsifying empirical claim, or the definitions involved in the collection of that data, or the meaning of the phrase "science is," or any one of dozens of other things that might be done to call into question the precise relationship between the statement and the evidence. The point is that falsifying a statement is a very complex endeavor, and some philosophers (notably Quine) have argued that one can in principle *always* preserve a theoretical statement by adjusting various background assumptions: the meanings of key terms, the scope of the claim, or the theory built into the way that the empirical data was collected and organized in the first place (Chernoff 2005, 183–184). All of these considerations mean that it is almost impossible to determine when and whether a claim has been falsified, making falsifiability a deeply problematic way to demarcate science from non-science (Hay 2002, 83–84).

It is important to note that the disagreement between Kuhn and Popper about falsifiability as a demarcation criterion is not merely an empirical dispute. Instead, falsifiability versus normal science rests on profoundly divergent views about how knowledgeable actors—scientists, to be sure, but also people in general—relate to one another and to the world that they are studying. For all of his criticisms of logical positivism, Popper retains one of their key presumptions throughout his work: the presumption that it is always possible to translate claims from one conceptual vocabulary into another one. To the extent that there are "frameworks" of assumptions standing behind our statements, Popper suggests, if we want to be intellectually honest and critical we have to break through those frameworks, lest we allow "ourselves to be caught in a mental prison" (Popper 1996, 53). Falsifiability, like verifiability, depends on the idea that a statement and the pieces of empirical evidence used to evaluate it must all be expressible in ways that would make them clear to any competent observer. Both falsifiability and verifiability would fall apart if they were relativized to a specific conceptual vocabulary, because that would make any statement's scientific status dependent on the language used to express it—and render the principle in question not a very useful demarcation criterion.

However, in many ways, this is precisely what Kuhn's argument *does*. Kuhn embeds scientific statements in the "paradigmatic" framework within which they occur and are evaluated, making it virtually impossible for anyone not working in a given paradigm to determine whether any particular statement is or is not falsifiable or verifiable—or whether the statement presents a viable puzzle to be solved. In this way Kuhn disrupts the very idea of "science" as a single unified field of endeavor, replacing that image with one of islands of incommensurable research. Needless to say, a science made up of incommensurable islands need not have, and most likely does not have, any common standards or criteria for the production of knowledge; nor does it have a single measurement of progress (Kuhn 2000, 85–86).[6] The unity of science—the assumption of perfect translatability that underpinned both logical positivism and Popperian falsifiability—is disrupted by Kuhn's suggestion that science is instead marked by radical discontinuity. Needless to say, the Popperian demarcation criterion drops out of contention too.

In an effort to get around these problems, Imre Lakatos famously proposed that analysts shift away from the evaluation of the scientific status of individual statements, and instead examine a series of statements—a "research programme"—in order to ascertain whether it is progressing or degenerating over time. Lakatos accepted much of Kuhn's account of science, including the idea that one cannot simply subject hypothetical statements to empirical testing in order to ascertain whether the statement is close to the truth. Although Lakatos rejected Kuhn's strong claims about the incommensurability of rival scientific theories (Lakatos 1978a, 112), he retained the idea that direct comparison of rival claims—either with one another or with the empirical world—is impossible. This necessitated the formulation of a second-order conceptual language, revolving around the rational reconstruction of scientific controversies *after the fact*, which would permit the comparison of research programmes in terms of their "progressive" or "degenerative" character (Lakatos 1978b). Were scientific theories directly testable, this conceptual architecture would not be needed, as one could more or less straightforwardly seek to falsify them by adducing the appropriate evidence (Jackson and Nexon 2009). Hence Lakatos' efforts should be seen as an effort to retain certain elements of the traditional definition of science while acknowledging the weakness (or, less charitably, the failure) of the Popperian account on methodological and empirical grounds.

All of this philosophical controversy about the definition of "science"—and I have only scratched the surface here, referencing mainly authors whose names have been commonly invoked in existing demarcation debates within the field of IR—makes it deeply problematic to claim, as IR scholars often do, that there are *any* criteria for the definition of "science" that are "standard in philosophy of science" (Vasquez 1995, 230). Instead, we are confronted with a situation in which a variety of standards and criteria present themselves, and absent a widespread consensus about these issues in the philosophy of science the door is opened for IR scholars to, in effect, reach into an alien field of study and pull out something that fits their immediate aims, while retaining the cultural prestige of "science" as a rhetorical warrant for their disciplinary maneuver. Far from solving the science question, this kind of intellectual instrumentalism simply muddies the conceptual waters even further.

Even worse, in staging these opportunistic raids into foreign scholarly territory, IR scholars routinely ignore the fact that demarcation debates among philosophers of science are generally concerned with shoring up or preserving notions such as "progress" and "truth" in the face of what might at first seem like discrepant evidence about how actual scientists do their empirical work. Philosophers engaged in demarcating science from non-science are thus, and necessarily, engaged in something of a *normative* enterprise (Laudan 1996, 217–218; Lakatos 1978a, 118–121). IR scholars also ignore the fact that philosophers of science engaging in these discussions are working in a *transcendental* mode, and are faced with obviously successful knowledge-producing endeavors, the success of which they are trying to account for in terms of their "scientific" character. No such obvious successes exist in IR, which changes the terms of the debate quite radically (Chernoff 2005, 54–55). Indeed, IR scholars routinely ignore Lakatos' firm division between the "*methodological* appraisal of a programme" and "firm *heuristic* advice about what to do" (Lakatos 1978a, 117)—a division that renders deeply problematic any effort to learn what science is from the study of other sciences, with intent to apply those lessons elsewhere. Finally, IR scholars ignore the fact that many contemporary philosophers of science would agree with Larry Laudan's observation that "the problem of demarcation . . . is spurious" because even a cursory examination of how various scientific endeavors proceed indicates that they are "not all cut from the same epistemic cloth" (Laudan 1996, 221). By simply taking what we like from the philosophical literature, we miss the context of, and the controversy surrounding, discussions about demarcation among philosophers.

All of this means that it is futile to look to the philosophy of science expecting a simple and clear answer to the question of how we ought to produce knowledge about world politics, because no such consensus answer is even remotely in evidence. Philosophers of science simply do not speak with one voice when it comes to demarcating and analyzing scientific practice.

Science, broadly understood

Faced with the impossibility of putting an end to the science question within IR by turning to the philosophy of science, what should we do? Since we cannot resolve the question of what science is by appealing to a consensus in philosophy, one option is to become philosophers of science ourselves, and to spend our time and our scholarly efforts trying to resolve thorny and abstract issues about the status of theory and evidence and the limits of epistemic certainty. But this is an unappealing option for a scholarly field defined, if loosely, by its empirical focus (world politics), and it would be roughly akin to advising physicists to become philosophers of physics in order to resolve the question of what physics was and whether it was a science. This also mis-states the relationship between philosophical debates and scientific practice; practicing scientists have a pretty good working definition of what it means for something to be "scientific," but this "is less a matter of strategy than of ongoing evaluative practice," conducted in the course of everyday knowledge-producing activities (Taylor 1996, 133). We do not expect physicists to give philosophical answers to questions about the scientific status of their scholarship; we expect them to produce knowledge of the physical world. Similarly, we should not expect IR scholars to engage in "philosophy of IR" to the detriment of generating knowledge about world politics; the latter, not the former, is our main vocational task.

If we should not all become philosophers of science, perhaps we should simply continue what we have been doing: deploying philosophical snippets in the course of our "ongoing evaluative practice" of one another's scholarship about world politics. After all, we are not philosophers of science, so why should it matter whether we are taking philosophical claims out of context? This option is equally unappealing, but for different reasons. For one thing, the rhetorical power of an appeal to "science" within IR, as within other scholarly fields that have inherited a "science question" from their forebears (Steinmetz 2005a), depends on a claim—perhaps implicit—that the criteria identified as "scientific" *are in fact* the kinds of knowledge-production practices that, if adopted, will establish IR as a science. In principle, at least, this is a claim that can be evaluated, and more importantly, it is a claim that can be true or false. Whether it is true or whether it is false has enormous implications for whether we ought to engage in the specified course of action. While the lack of consensus among philosophers of science should put to rest the idea that any given knowledge-production practices are *uniquely* scientific, it is still entirely possible to ground claims to scientific status in firmer philosophical arguments, and thus to move beyond the merely tactical use of a term such as "science."

Besides this logical reason, there is also an ethical reason why we should stop taking philosophical claims about "science" out of context and using them to shore up our positions within disciplinary debates: when we invoke "science," we are in a very practical sense playing with fire. The cultural prestige of "science" is such that tapping that commonplace in a debate is really akin to bringing out the big guns, raising the temperature of the controversy to the point where one wonders how far we are from an accusation of "relativism" and an accompanying violation of Godwin's Law.[7] Under such circumstances, it is even more important to ask whether the appeal to "science" is philosophically appropriate.

A third option would be simply to de-escalate our controversies about research practices and refrain from invoking "science" in such discussions at all. Larry Laudan suggests that philosophers of science ought to do just this, shifting their attention to "the question of reliable knowledge" and giving up any attempt to define the boundaries of scientific practice (Laudan 1996, 222). But Laudan's proposal, I would argue, is only feasible within a scholarly field not as dominated by the science question as IR has historically been. Whether the philosophy of science is itself a science remains a much less pressing question than the question of whether the study of world politics is or can be a science. In addition, the cultural prestige of "science" makes the notion a very appealing rhetorical weapon;

a simple promise not to use it is probably not credible, and as long as "science" retains its broader appeal, it will likely be too tempting for one party of a debate to reach for the commonplace in the course of discussion. Simply removing the claim to "science" from IR discussions is, therefore, probably quite a futile endeavor.

Hence, the best response to the fact that the science question cannot be simply resolved by a turn to philosophy is to *replace* the narrow definition(s) of "science" circulating in the field with a definition that simply cannot be deployed by partisans of any single approach to the study of world politics as part of an effort to render their opponents' claims unworthy of serious consideration. What we should be avoiding, as a field, are derisive caricatures of one another's work as "storytelling," "mindless number-crunching," or "philosophical mumbo-jumbo," and the accompanying characterization of those approaches as "unscientific" and hence not worthy of intellectual engagement. Similarly, we ought to be avoiding caricatures of self-proclaimed "scientific" work as being out of touch with the actual world, incapable of appreciating the complexity of social life, or necessarily wedded to the preservation of the status quo. Instead, a principle of charity (Blackburn 1994, 62) is called for: treat other arguments about world politics as serious attempts to generate knowledge. But as long as "science" remains in circulation in the field in the vague form in which it presently exists, such charitable readings are unlikely to survive, as it is too tempting simply to wield "science" as an excuse for not engaging claims at odds with one's own.

In order to craft a sufficiently broad definition of science, it is important not to replicate the errors and weaknesses associated with the disciplining deployments I have been criticizing. As such, it is unlikely that an acceptable definition of science can be produced by looking for fundamental "rules of inference on which" the "validity" of "scientific research . . . depends" (King, Keohane, and Verba 1994, 9). The reason is simple: different kinds of empirical research in IR adhere to different "rules of inference," and some reject inference itself in favor of (for example) thick description or structural overdetermination or discourse analysis. Hence, making some set of "rules of inference" the criterion for scientific status simply replicates the same disciplining move under the guise of advancing a putatively neutral set of methods and techniques. Arguably, *any* attempt to specify universal rules and procedures is doomed to collapse into a disciplining move, since there are no rules so universally agreed upon that their adoption would be uncontroversial. The commonality of "science" in IR, then, cannot be sought in rules or procedures for handling evidence or evaluating claims.

Perhaps the common element animating a field-wide definition of science can be found not in the supposed methods of science, but in the *goals* of science. Colin Wight suggests that "what distinguishes scientific knowledge is not the method of knowledge acquisition, nor the immutable nature of the knowledge produced, but the aim of the knowledge itself," which he takes to be the "explanatory content" of scientific knowledge (Wight 2006, 61). Defining science in this way seems promising, as long as the precise definition of "explanatory" is allowed to vary so as to encompass a variety of approaches to explaining phenomena in world politics. Unfortunately, Wight promptly goes further in specifying a sense of "explanatory" that excludes more than a few ways of studying world politics:

> What marks scientific knowledge out from other forms of knowledge is that it attempts to go beyond appearances and provide explanations at a deeper level of understanding. This implies that the scientist believes that there is a world beyond the appearances that helps explain those appearances.
>
> (ibid., 18)

Thus Wight offers a unity of *ontology*—the belief in a mind-independent reality to which our concrete researches should be directed (Wendt 1999, 52–53)—as the crucial element in science. But this locking down of a precise meaning of "explanatory" drives us right back into the disciplining move of accepting one philosophically controversial account of science and shaping our empirical work in IR in accord with it—and dismissing other kinds of work as not sufficiently "scientific."[8] Absent a universal consensus about the validity of presuming the existence of a "world beyond . . . appearances," this is not a solution to our problem.

Indeed, perhaps the only solution that does not presume a non-existent philosophical consensus about the definition of "science" would be an account of science that, in effect, equated science with empirical inquiry designed to produce knowledge. Such an account would not give a lot of specific guidance as to how empirical research should be conducted, but it would serve to differentiate the production of knowledge about world politics from other things that one might do with respect to world politics—other things that might be valuable in their own way, but which would not be reducible or equivalent to knowledge-production. Such an account would also allow the criteria for *good* knowledge about world politics to vary between approaches; designating all empirical inquiry designed to produce knowledge as science in no way says that all knowledge-claims are equally good ones. It simply shifts the question—along the lines of both Laudan's and Lakatos' criticisms of the demarcation problem—from "Is this piece of work scientific?" to "Is this piece of work a good piece of work?" Naturally, *answering* that question in any particular situation will require us to elaborate and specify standards for good work, but by getting the rhetorical trump-card "science" out of the mix, a broad definition allows us to focus on the knowledge-production techniques in our own field instead of focusing on what we think other fields are doing.

This may be the most important contribution of a broad and pluralistic definition of science: to cure IR of its perennial envy of other fields of scholarly inquiry by highlighting the important conceptual work on the matter of science that has already been done *within the social sciences themselves*. Almost four decades ago, Albert O. Hirschman called for precisely this kind of self-assertion by practitioners of the study of politics, arguing (as an economist!) that political scientists need not accept the colonization of their field by economists:

> [R]eciprocity has been lacking in recent interdisciplinary work as economists have claimed that concepts developed for the purpose of analyzing phenomena of scarcity and resource allocation can be successfully used for explaining political phenomena as diverse as power, democracy, and nationalism. They have thus succeeded in occupying large portions of the neighboring discipline while political scientists—whose inferiority complex vis-à-vis the tool-rich economist is equaled only by that of the economist vis-à-vis the physicist—have shown themselves quite eager to be colonized and have often actively joined the invaders. Perhaps it takes an economist to reawaken feelings of identity and pride among our oppressed colleagues and to give them a sense of confidence that their concepts too have not only *grandeur*, but *rayonnement* as well?
>
> (Hirschman 1970a, 19–20)

What Hirschman claims about *substantive* concepts, I mean to suggest, is equally true of *methodological* concepts: those of us engaged in the scholarly study of social and political life have our own proud tradition of reflection on the science question, and the broad definition I want to propose comes directly from the seminal reflections of Max Weber on this topic. Adoption of this broadly Weberian account of science, I suggest, can quite neatly resolve the problems I have been discussing.

For Weber, what defines "science" is not its manner or its method, but its goal—a goal that, in the first instance, differentiates it from partisan politics:

> The taking of practical-political positions and the scientific analysis of political structures and party positions are two very different things. If you are speaking about democracy in a popular meeting, you do not need to make a mystery of your personal position; instead, clearly taking a recognizable position is your damned duty and responsibility. The words you use are not tools of scientific analysis, but political advertisements against the positions of others. They are not ploughshares for the loosening of the soil of contemplative thought, but swords for use against your opponents: weapons.
>
> (Weber 1917, 14–15)

The distinction that Weber is drawing here is a *logical* distinction between two different ways of using words and concepts. In the realm of practical politics, the key goal is the achieving of results; the clarity or defensibility of those words and concepts is of decidedly secondary importance. But in the realm of scientific analysis, the order is inverted: what matters most of all is the systematic application of a set of theories and concepts so as to produce a "thoughtful ordering of empirical actuality" (Weber 1999a, 160). Weber elaborates:

> The social science that we want to concern ourselves with is a *science of actuality*. We want to understand *in its particularity* the encompassing actuality of the life in which we are placed—on one hand, the coherence and cultural *significance* of individual occurrences in their contemporary

> configuration, and on the other hand, the reasons for those occurrences being historically so and not otherwise.
>
> (Weber 1999a, 170–171)

For Weber, then, there is no fundamental opposition between "explaining" and "understanding," as both are equally scientific. Instead of reading Weber as a partisan for one or another specific *kind* of social science, as Hollis and Smith (1990, 72–82) do, we should understand Weber's project as the attempt to define a basic and broad notion of "social science" *within which* we might then discuss or debate (for example) the extent to which we ought to take an actor's description of her or his action as a point of departure for our analysis. Thus Weber's encompassing definition of science, which we might think of as "systematic empirical analysis that aims to produce knowledge rather than to produce innerworldly effects," provides a big enough tent to put out the fires associated with accusations of being "unscientific."

Another way to put this is that Weber's definition is that science, including social science, should be concerned with empirical *facts* rather than with evaluative *judgments*. Weber distinguishes between an idealized analytical concept of "Christianity" that might be used to generate factual knowledge about some particular sect or arrangement, and an evaluative definition of "Christianity" that might provide a basis on which to judge whether some particular doctrine or arrangement was or was not actually *Christian*:

> Here it is *no longer* a matter of a purely theoretical process of *referring* to values empirically, but instead of value-*judgments* which have been taken over into the "concept" of Christianity. *Because* the ideal-type claims empirical *validity*, it towers into the region of the evaluative *interpretation* of Christianity. The ground of empirical science is forsaken; before us stands a profession of faith, and *not* an ideal-typical *conceptual* construct.
>
> (Weber 1999a, 199)

In IR terms, we might think of this as an admonition that we ought not to confuse a concept such as "sovereignty" or "human rights" that we might use in generating empirical facts about world politics with a normative standard that we might use to judge or evaluate world politics. For Hedley Bull, the distinction between "order" and "justice" illustrated this nicely: Bull treated order primarily as "an actual or possible condition or state of affairs in world politics," and thus as an instrument for generating factual knowledge of social relations, while arguing that justice "belongs to the class of moral ideas, ideas which treat human actions as right in themselves" (Bull 1977, 77–78). Justice, for Bull, is therefore a concept useful for a normative evaluation of those same social relations: an evaluative commentary on the facts, rather than the production of factual knowledge. These are logically distinct endeavors.[9]

However, it does not follow from the dictum that science ought to be focused on the production of factual knowledge that the practice of academic analysis is somehow devoid of values. Indeed, Weber argues:

> There is simply no "objective" scientific analysis of cultural life—or, put perhaps somewhat more narrowly but certainly not essentially differently for our purposes—of a "social phenomenon" *independent* of special and "one-sided" points of view, according to which—explicitly or tacitly, consciously or unconsciously—they are selected, analyzed, and representationally organized as an object of research.
>
> (Weber 1999a, 170)

The inescapability of value-commitments does not mean that "*research* can only have *results* which are 'subjective' in the sense that they are *valid* for one person and not for others" (ibid., 183–184). Indeed, as I have been arguing, the distinctiveness of science for Weber is not that it embodies no value-commitments, but that it does something distinctive with those commitments. Value-commitments place a specific duty on the practicing (social) scientist:

> A systematically correct scientific demonstration in the social sciences, if it wants to achieve its goal, must be recognized as correct even by a Chinese (or, more accurately, it must constantly *strive* to attain this goal, although it may not be completely reachable due to a dearth of documentation). Further, if the *logical* analysis of the content of an ideal and of its ultimate axioms, and the demonstration of the consequences that arise from pursuing it logically and practically, wants to be valid and successful, it must be valid for someone who lacks the "sense" of our ethical imperative and who would (and often will) refuse our ideal and the concrete *valuations* that flow from it. None of these refusals come anywhere near the scientific value of the *analysis*.
>
> (ibid., 155–156)

The basic point here is that even someone who rejects our values should be able to acknowledge the validity of our empirical results within the context of our perspective. The fact that we have a perspective—that our results were produced by the application of concepts and procedures derived from a specific set of values—is philosophically and epistemologically important, but it has little or no bearing on the question of whether a piece of work is "scientific" or not. Instead, the decisive issue is *internal validity*: whether, given our assumptions, our conclusions follow rigorously from the evidence and logical argumentation that we provide.

None of this is to say that normative evaluation of world politics is not a good and worthwhile activity, or to say that the distinction between science and politics denigrates the actual practice of politics. Nor is the implication here that the scholarly field of IR ought to be exclusively "scientific," even in the broad Weberian sense I have proposed here. It is, rather, to distinguish logically between a number of ends to which we might apply our scholarly efforts. We could engage in the generation of political arguments and commentaries; we could engage in the normative evaluation of actually existing political and social arrangements; or we could engage in the systematic production of factual knowledge about those political and social arrangements. Calling only the third of these “science” preserves the integrity of all three ends: in order for the claim to scientific status to have any *value* in the political or normative realms, it is logically necessary for science to be distinct from those endeavors. Otherwise, calling a claim “scientific” is perhaps nothing but shorthand for saying that one agrees or disagrees with it, perhaps on political or normative grounds. Whether a scientific claim ought to trump a political one, or whether normative claims ought to build on scientific ones, are open questions, but they cannot even be *asked* if one does not start from the position that science constitutes a distinct endeavor. Not necessarily a better or worse endeavor, but a *distinct* one.

2 Philosophical wagers

The broad, Weberian definition of science I have sketched in the previous chapter is designed to accomplish two tasks. First, it effectively makes science equivalent to systematic inquiry designed to produce factual knowledge. Second, it differentiates science from politics and from normative evaluation. As such, this broad definition of science makes it virtually impossible for the charge of being "unscientific" to be used as a way to discredit a piece of scholarship that intends to contribute to our factual knowledge of the world. The only kinds of works against which such a charge could be legitimately deployed—works of normative analysis and works of political advocacy or commentary, and probably works of art—would, almost certainly, not be particularly interested in classifying themselves as "scientific." Even critical-theoretical scholarship in the Frankfurt School (Linklater 2007) or neo-Gramscian (Cox 1996b) traditions, which routinely emphasizes the evaluative aspects of scholarship, relies on factual claims about the empirical world in order to give its critical interventions sufficient force (Geuss 1981, 109). The critical-theoretical argument about scholarship and values is, in the language I have introduced here, an argument that the scientific parts of scholarship ought to be supplemented by normative or even partisan-political parts. As long as Weber's admonition about making it clear "where the analytical researcher becomes silent and the advocating person begins to speak" (Weber 1999a, 167) is adhered to, this poses no special problems for a broad definition of science.

That said, the Weberian definition of science does not tell us very much about precisely what we ought to be doing when we conduct research on world politics. This is also by design, since linking any *specific* approach to worldly knowledge-production with the label "science" simply re-opens the unproductive disciplining debates so prominent in the field of IR over its history. The only way that such a strategy would be justified would be if there were broad philosophical consensus on the definition of science, but this is simply not the case. Hence, deploying claims derived from, or authors working on, the philosophy of science for the purpose of defining science—and therefore disciplining all empirical research in the field of IR—appears to be an enterprise fraught with peril. If philosophers of science as a group do not agree on what science is, what intellectual warrant do we have to pluck out one or another position on science from within their discussions and place it as a standard in front of our particular campaign to alter the field?

However, the fact that we should not be looking to philosophy of science as a way to resolve definitively the science question does not mean that IR scholars have no use for the philosophy of science. If we stop expecting that philosophy of science contains some kind of master strategy that will, if implemented in IR, make us truly "scientific," perhaps we can start to appreciate the *actual* value of philosophical reflections on knowledge-production: systematically clarifying the implications, especially the methodological implications, of taking a particular stand on how to produce knowledge. A broad definition of science, by design, does not provide us with any standards for good research, or indeed any specific advice for how to go about doing research, beyond the two basic admonitions to focus on factual knowledge of the world, and to separate this activity logically and conceptually from the promulgation of normative judgments and partisan-political stances. But methodological advice and standards are indispensable components of any actually existing line of scientific research; practicing researchers necessarily operate with a wide variety of techniques designed to facilitate and improve their research, and to criticize constructively the research produced by others. Philosophy of science, as a reflection on scientific research practice, can help us to make explicit some of the tacit principles with which researchers in particular traditions are already operating. In other words, philosophy of science can help us to *clarify* IR research practices, with an eye towards making them more coherent and potentially more productive.

This makes the utility of the philosophy of science for IR primarily a *methodological* utility. By "methodology" in this context I mean something quite different than "methods:" methods are techniques for gathering and analyzing bits of data, whereas methodology is "a concern with the logical structure and procedure of scientific enquiry" (Sartori 1970, 1,033). Philosophy of science is not going to teach anyone how to run a multivariate regression testing hypotheses about democracy and economic growth, or how to craft an ethnographic account of the activities of the Ministry of Foreign Affairs, but it can help us think through the decision to utilize those methods, and make sure that we are using research methods in ways that complement one another or generally hang together. We do not spend much time in the field wrestling with such methodological questions; instead, we engage in discussion about methods, debating such technical issues as the relative merits of different techniques of case-selection and case-comparison (George and Bennett 2005; McKeown 1999; Mahoney and Goertz 2004) or how to identify the appropriate documents for use in a discourse analysis (Hansen 2006, 51–54; Bially Mattern 2004, 63–68). These are important questions of method, but they are not questions of methodology, inasmuch as these discussions presume a whole variety of things about the definition of knowledge and the overall goal of empirical research. Indeed, absent at least a broad agreement on strategic questions about the character and status of knowledge, it is unlikely that the tactical debates about how best to achieve those strategic goals could even take place.

That we do not do a lot of this kind of reflection in IR, or in the sciences generally, is quite understandable when one remembers that our primary professional job is the production of knowledge about the world, and our primary specialized training is in specific techniques of data-collection and data-analysis. Philosophy of science is not even a required course in many, if not most, Ph.D. programs in IR (Schwartz-Shea 2003), further contributing to our challenges in engaging in these kinds of conceptual discussions. However, it is tremendously important that we not lose sight of methodological issues as we craft and evaluate pieces of empirical research, both because methods without methodology can be quite myopic in lacking a big picture within which specific techniques might make sense, and because in the absence of explicit methodological reflection there is a not inconsiderable chance that scholars working in various lines of research will continue to consider their way of conducting research to be uniquely "scientific" rather than *a* way of doing scientific research. Methodological reflection, assisted by readings in the philosophy of science, is the cure for both of these ills.

Ontology, philosophical and scientific

By linking philosophy of science to methodology, and foregrounding methodological reflection in thinking about how to do empirical research, I am deliberately breaking with a tradition of denigrating methodology that is common among philosophers and scientists alike. In that tradition, methodological questions come late in the game, after more fundamental issues have been sorted out; hence the proper place of philosophy of science would be prior to methodology. Three section-headings from Audie Klotz and Cecelia Lynch's recent book on research techniques (Klotz and Lynch 2007), and the sequence in which they occur in the book's first chapter, tell the story:

> Ontology: how do researchers conceptualize what they study?
> Epistemology: how do researchers know what they know?
> Methodology: how do researchers select their tools?

The sequence here, which is echoed in numerous contemporary guides to research, runs from ontology (concerning *being*, and what exists in the world) to epistemology (concerning *knowing*, and how observers formulate and evaluate statements about the world) and only then to methodology—here as elsewhere in the literature defined as the selection of specific research tools. Colin Wight clarifies this sequence, contrasting an "inclusive" definition of methodology such as the one I have advanced with a "less expansive notion" (such as that presumed by Klotz and Lynch) that equates methodology with "the differing methods of gaining knowledge *relative to the object of inquiry*" (Wight 2006, 258; emphasis added). I have highlighted the crucial clause in Wight's claim, since by linking methodology to "the object of inquiry" he also privileges ontology and epistemology over methodology. Indeed, Wight explicitly claims that "methodologies are always, or at least should be, ontologically specific . . . the methods used to study atomic particles, for example, would be wholly inappropriate when applied to the study of social processes" (ibid., 259).[1] Therefore we ought to begin with the world and compose our research strategies accordingly—a position that involves putting ontology first, and maintaining that "it is the nature of objects that determines their cognitive possibilities for us" (Bhaskar 1998, 25).

Wight further argues that a privileging of methodology in the abstract might lead to efforts to define a single "scientific method," and thus act as "a potential barrier to methodological innovation and pluralism" (Wight 2006, 258). His fear seems partially justified when we consider the fact that contemporary efforts to define a universal, categorical scientific approach—especially within the social sciences—stake their claim *precisely* on the distinction between claims about the world and claims about the design and goals of empirical research, as when King, Keohane, and Verba (1994, 20, 29–30) distance themselves from "parsimony" (a claim about the composition of the world) in favor of "leverage" (a principle of hypothesis-construction). Hence, we appear to have a choice between starting with the world and conforming our methodology to that world, or starting with methodology and thus losing the world as we try to articulate universal standards for scientific research—universal standards that I have been claiming *do not exist* in any intellectually defensible way.

On Wight's account, the role of philosophy of science would be to clarify our *ontological* assumptions, not our methodological practices. Philosophy of science has been used to do this in the field of IR in recent years, starting with Wendt's seminal paper on the agent-structure problem (Wendt 1987), which drew on critical realist philosophy of science to suggest that unobservable structures, both in the natural and the social worlds, were as real as the objects of sensory experience. Notions of "punctuated equilibrium" (Spruyt 1994) and "complexity" (Hoffmann and Riley 2002) have made their way into the study of world politics through a similar route: from natural science, through philosophical reflection, and finally into IR. The implicit logic driving such importations seems to be that if natural scientists, or those philosophers who reflect on the natural sciences, have a way of apprehending the world that works well for them, then maybe it will work equally well for us—even if certain technical aspects of empirical research need to be altered so as to take account of the ontological differences between mute natural objects and self-aware human beings (Bhaskar 1998, 159). In any event, ontology comes first.

However, I do not think that putting ontology first is the panacea that many seem to think it is. For one thing, if one puts ontology first then one is, at least provisionally, committed to a particular (if revisable) account of what the world is made up of: co-constituted agents and structures, states interacting under conditions of anarchy, global class relations, or what have you. This is a rather large leap to make on anyone's authority, let alone that of a philosopher of science. Along these lines, it is unclear what if any *warrant* we could provide for most ontological claims if ontology in this sense were to always "come first." If someone makes an ontological claim about something existing in the world, then we are faced with an intriguing *epistemological* problem of how possibly to know whether that claim is true, and the equally intriguing problem of selecting the proper *methods* to use in evaluating the claim (Chernoff 2009b, 391). But if epistemology and method are supposed to be fitted to ontology, then we are stuck with techniques and standards designed to respond to the specificity of the object under investigation. This problem is roughly akin to using state-centric measurements of cross-border transactions to determine whether globalization is eroding state borders, because the very object under investigation—"state borders"—is presupposed by the procedures of data-collection, meaning that the answer will always, and necessarily, assert the persistence of the state.

There is also a more fundamental problem with "putting ontology first," which is that ontology in contemporary philosophical usage can refer to two different, but related, components of a way of apprehending the world. On one hand, ontology can refer to a catalog of objects, processes, and factors that a given line of scientific research expects to exist or has evidence for the existence of: ontology as bestiary, so to speak, concerned with what exists, or with the general principles on which such existence might be determined. On the other hand, ontology can refer to the conceptual and philosophical basis on which claims about the world are formulated in the first place: ontology as our "hook-up" to the world, so to speak, concerned with how we as researchers are able to produce knowledge in the first place (Shotter 1993b, 73–79). Patomäki and Wight helpfully distinguish between these two uses of the term "ontology" by designating the former "scientific ontology" and the latter "philosophical ontology" (Patomäki and Wight 2000, 215); they also note that philosophical ontology is logically, and necessarily, prior to the construction of any scientific ontology, since we cannot make defensible claims about what exists until the basis on which we are doing so has been clarified.

So when we talk about putting ontology first, which kind of ontology do we mean? Since philosophical ontology takes logical and conceptual priority, one would think that philosophical ontology ought to come first. However, most advocates of putting ontology first seem more concerned with elaborating their particular scientific ontology, and putting *that* first: before epistemology, methodology, or concrete research methods. For Wight, this scientific ontology involves agents and structures as irreducible objects of "interdependent nature," meaning that they never occur separately but nonetheless remain essentially distinct from one another (Wight 2006, 296). For Wendt, this scientific ontology involves states as the actually existing persons of international society interacting so as to produce and sustain a variety of "cultures of anarchy" (Wendt 1999, 246–250). In that way, the call to put ontology first seems to mean approximately the same thing as having a clear definition of the entities and factors with which one is concerned: states (Nettl 1968), firms (Williamson 1998), transnational social movements (Keck and Sikkink 1998), and so forth. What it means to produce knowledge and how we produce knowledge could then be customized to the particular features of the entities and factors under investigation.

This pull away from philosophical ontology towards scientific ontology is so strong as to affect even works overtly concerned with ways of producing knowledge rather than with the objects of knowledge. A most prominent example of this is Hollis and Smith's widely read book *Explaining and Understanding International Relations* (Hollis and Smith 1990), which begins with some claims about philosophical ontology proper but then mixes in claims about objects and entities—elements of scientific ontology—in seeking to elaborate what it might mean to study world politics from different theoretical and conceptual standpoints. Hollis and Smith begin by contrasting "explaining" and "understanding" as separate "traditions" yielding different kinds of accounts of world politics, with an "explaining" story working from an outsider's perspective "in the manner of a natural scientist seeking to explain the workings of nature and treating the human realm as part of nature," while an "understanding" story works from the inside, "told so as to make us understand what the events mean, in a sense distinct from any meaning found in unearthing the laws of nature" (ibid., 1).[2] At the outset, then, we are in the realm of philosophical ontology, since what is at stake in the contrast between "explaining" and "understanding" is not the character of the world, but rather how we observers are hooked up to it. That this is the case can be easily glimpsed by asking whether it would make sense to generate both kinds of stories about any given situation, social or natural; to do this we need not look far to find both insider and outsider accounts of both the natural and social worlds.[3] Insider "understanding" and outsider "explaining" accounts can, in principle, be used to generate knowledge of *any* kind of object; as philosophical ontologies, they logically precede any possible scientific ontology or catalog of entities and factors.

Hollis and Smith, however, quickly slip into enumerating characteristics of objects, linking those enumerations to the two traditions with which they are concerned. Insider and outsider accounts, we quickly learn, conceptualize individual human beings quite differently:

> *X* is an actor conceived in the spirit of the scientific ["explaining"] tradition, *Y* the counterpart in the spirit of the interpretative ["understanding"] tradition . . . Being part of the natural world and a proper object of scientific study, *X* is predictable on the basis of *X*'s preferences and information, which are in turn the result of *X*'s nature and nurture . . . The fabric of *Y*'s social world is woven from rules and meanings, which define relationships among the inhabitants and give interpretations their purpose . . . *Y* is expected to pick an intelligent course through a variety of social engagements, to which actors bring something of themselves in exercising their social capacities.
>
> (ibid., 4–6)

We are no longer in the realm of philosophical ontology, and "explaining" and "understanding" now name substantive conceptions of things in the world rather than ways in which the researcher is connected to the world. The shift here is subtle, but important: in the space of a few pages we have gone from different ways of encountering the world (from the outside or from the inside) to different conceptions of objects in the world (*homo economicus* and *homo sociologicus*, so to speak).[4] Indeed, it would not be too much of a stretch to say that Hollis and Smith's argument that "explaining" and "understanding" accounts cannot be reconciled rests on the fact that, substantively speaking, the world envisioned by "explaining" and the world envisioned by "understanding" are *not the same world*, as the explaining-world is a world of structural constraints where people's social capacities have to be explained in terms of broader social forces, while the understanding-world is a world of historical endowments that offer possibilities that can only be actualized by playing out a set of social interactions (ibid., 209–212). But that is a disagreement that takes place almost exclusively on the terrain of scientific ontology, and involves "worldviews" rather than ways of being connected to the world.

The virtual disappearance of philosophical ontology from IR debates—and its ready replacement by sets of substantive considerations—carries with it a set of costs for IR scholarship. Chief among these is that every substantive disagreement is transformed into an empirical dispute, but without any clear guidelines for how such disputes are supposed to be resolved. That such empirical disputes are difficult to resolve is evidenced by a quick glance at the ongoing debates surrounding the question of whether "balancing" or "bandwagoning" behavior among states predominates at the level of the international system (Kaufman, Little, and Wohlforth 2007), or whether "ideas" or "material factors" were the most important cause of the end of the Cold War (Brooks and Wohlforth 2001; English 2002; Brooks and Wohlforth 2002). Further, what comes up in these debates on a regular basis are questions of methodology and research design: what kind of knowledge of the world we can and should produce, and how to go about producing such knowledge. However, in the absence of any sustained attention to philosophical ontology, *such questions are almost certainly irresolvable*, as any scholar can at almost any time retreat behind the safety of their particular view of the world—their scientific ontology—and the sets of research techniques designed to work in and with that world. Thus, realists read world politics as characterized by a struggle for power among independent political units, neoliberal institutionalists read world politics as characterized by a competitive set of mixed-motive games under conditions of interdependence, and when confronted by evidence emanating from the other camp, partisans of each worldview simply reassert their central postulates and go on reading the world in their own way.[5]

Of course, one way to resolve this fragmentation would be to impose a set of common standards—one might even call them "scientific" standards—on the field as a whole, and then subject every worldview to the same procedures of systematic evaluation. Besides the fact that the lack of consensus among philosophers of science makes any such imposition arbitrary in the extreme, there is a further problem in that the very idea of empirically adjudicating between scientific ontologies *presumes a certain philosophical ontology*—a philosophical ontology that implicitly animates both calls to put ontology before epistemology (Wendt 1999, 52) and calls to dispense with "meta-theory" in favor of a focus on substantive claims (for example, Friedman and Starr 1997). In both cases, scholars are enjoined to stop worrying about their "hook-up" to the world and simply focus on the world itself and the entities and factors in it, whether those are sovereign territorial states or patterns of global class domination or whatever. The philosophical ontology underlying all of these claims, the grounds on which a claim advocating a focus on the world rather than on our hook-up to the world can be sensibly articulated, is the apparently innocuous notion of "independently existing reality" (Patomäki and Wight 2000, 217)—the notion that there is a world "out there," beyond all of our knowledge-making practices, to which our claims refer and with which those claims can be compared in order to assess their veracity. This *mind–world dualism* is the philosophical ontology that makes meaningful the proposition that we can empirically evaluate scientific ontologies, because if there is a world existing "out there" in a mind-independent way, we can in principle compare any given scientific ontology to that world and see if it matches in some sense.[6]

In fact, mind–world dualism also underpins the very distinction with which I began this discussion of types of ontology: the separation between ontological concerns on the one hand, and epistemological and methodological concerns on the other. In order to coherently argue that knowledge-production is separate from and subordinate to the way that the world is, it is necessary to argue that the world exists independently of our knowledge of it, and that the world places limits on how we may produce knowledge of it. Epistemology as a separate philosophical focus only emerged after the early-Enlightenment redefinition of the situation of human beings as individual minds facing an external world, and from Descartes onward largely concerned itself with trying to bridge the gap between the mind and the world in a robust and defensible manner, asking whether we could trust sensory impressions, whether ideas were innate or arose from observation, and whether and in what sense generalizations could be considered valid (Taylor 1995, 3–5). I will unpack some of these controversies in subsequent chapters; for the moment, my point is simply that *all of these issues presume mind–world dualism*. In the absence of a firm separation between the mind and the world, there would be no mind–world gap to bridge and, indeed, no "epistemology" as such. If "mind" and "world" are not two separate and distinct things, then it literally makes no sense to speak of the world as independently existing, since mind would be always and already intertwined with the world; nor would it make sense to subordinate epistemological and methodological concerns to the specific features of the world, since those features cannot be sensibly referred to outside of the context of the practices of knowledge-production that we employ when investigating them.

So perhaps the most significant implication of the disappearance of an explicit consideration of philosophical ontology within IR debates, and the consequent rush to elaborate scientific ontologies and to design research techniques and approaches, is that mind–world dualism goes largely unnoticed and largely uncriticized. This would not present any particular problems or challenges, except for the fact that mind–world dualism is far from uncontroversial in philosophical circles, where it has been contested under a banner that should be very familiar to contemporary IR scholars: social construction. This is more than a mere coincidence of labels, as IR constructivists have been leveling challenges at mind–world dualism for at least two decades (Onuf 1989; Kratochwil 1989), but have often been charged by critics with failing to elucidate empirically testable propositions about world politics. In other words, constructivists are charged with failing to subject their scientific ontologies of rules and norms and transactional social practices to the kinds of evaluation procedures that are only meaningful *within* a philosophical ontology of mind–world dualism—procedures involving efforts to compare expected outcomes with observed outcomes, and so to test (for example) the relative causal weight of social identities versus structurally induced preferences (Fischer 1992; Schweller and Wohlforth 2000). We persistently fail to notice the logical absurdity of the situation—obviously it makes no sense to evaluate a claim *opposing* mind–world dualism by *presuming* mind–world dualism—in part because we do not think enough in IR about philosophical ontology and its implications for research practice.[7]

Philosophy of science can help us to think more clearly about these issues, not by providing us with solutions but by elaborating the logical consequences of adopting particular positions on issues such as mind–world dualism. In order to realize that potential, we have to affix philosophy of science not merely to scientific ontology, and not merely to epistemology or the choice of methods, but first and foremost to methodology broadly understood: methodology as philosophical ontology, setting the context within which particular practices of knowledge-production might make sense. Wight (2006, 258) is entirely correct that this account minimizes the "difference between methodology and philosophy of science," but I do not think that the dire consequences that he foresees for "innovation and pluralism" necessarily follow because I am not proposing new restrictive methodological or philosophical standards for "science." Indeed, the important thing about the philosophy of science for IR scholars and scholarship is precisely that there are a *variety* of claims about our hook-up to the world, and thus a variety of philosophical ontologies, each of which holds different implications for how we should go about producing factual knowledge about world politics. Hence, we should be pluralist about the answers to these philosophical issues; in *this* sense, we should indeed "put ontology first" (Shotter 1993a, 77–78). As long as we recognize the diversity of philosophical ontologies, there is no danger that a connection between philosophy of science and methodology broadly understood will lead to anything like a new orthodoxy.[8]

Core wagers: a practical typology

How should we organize that diversity so as to bring out the most salient points of agreement and disagreement? In order to produce a mapping of philosophical ontologies that will be of use to IR scholars, we are faced with the challenge of specifying a set of distinctions between approaches to the philosophy of empirical inquiry that might enable something similar to an informed discussion between aficionados of various perspectives. But philosophy of science as a field does not have a widely accepted organizational scheme dividing authors and positions into distinct schools of thought, and to the extent that particular authors self-identify with a tradition of inquiry, they generally do not do so in terms of philosophical ontology per se. Getting a grasp on the disputes among philosophers of science is a tricky business.

Indeed, surveys of work in the philosophy of science—and I am setting aside those putative "surveys" that have as their not-so-hidden aim the vindication of the author's own particular standpoint—adopt one of two strategies of presentation:

they either proceed historically, describing authors and debates more or less chronologically (for example, Godfrey-Smith 2003), or they proceed topically, organizing the discussion around issues such as justification or perception (for example, O'Brien 2006). Along the way, we sometimes hear of more or less coherent positions such as "realism," or supposedly coherent positions such as "positivism,"[9] but such positions encompass a wide variety of stances and claims that frequently overlap with one another in a way that makes it difficult to summarize the core commitments of each. Add to this the fact that certain positions are quite intimately connected to the work of a particular author—such as Duhem and conventionalism, or Quine and naturalism—and the task of enumerating a general overview starts to look quite daunting.

A clue about how to proceed might be found by redirecting our attention to the purpose of the exercise: to make the systematic reflections found in the philosophy of science accessible to IR scholars, and to do so in a way that foregrounds salient points of distinction. It is therefore not necessary to capture every debate in the philosophy of science; it is only necessary to produce a set of categories that helps to illuminate discussions within and issues pertinent to IR, and perhaps other social sciences. Such a set of distinctions—such a classification scheme—should, in John Dewey's terminology, be evaluated "functionally, not structurally and statically:" the central issue should be whether the classification permits and promotes the particular end to which it is directed (Dewey 1920, 150). In the present case, the end to be promoted is a robust contrast between perspectives, and this carries two consequences for the scheme: distinctions must be drawn sharply enough to clarify disagreements, but the resulting positions have to resemble one another sufficiently that scholars can meaningfully elaborate the consequences of adopting one or another of the positions. This certainly does not mean that positions and perspectives on the philosophy of science have to be made *commensurable* in a way that would permit some kind of direct empirical test between them; indeed, because of the nature of the philosophical issues under discussion, no such empirical testing is even *conceivable* (Smith 1989, 21). But it does mean that we have to construct positions that are susceptible to comparison and contrast, because they are at the very least trying to occupy the same conceptual terrain.

Dewey also gives some helpful advice for the construction of such a classification scheme:

> The teleological theory of classification does not therefore commit us to the notion that classes are purely verbal or purely mental. Organization is no more merely nominal or mental in any art, including the art of inquiry, than it is in a department store or railway system. The necessity of execution supplies objective criteria. Things have to be sorted out and arranged so that their grouping will promote successful action for ends.
>
> (Dewey 1920, 154)

Two important procedural suggestions emanate from this observation. First, and in line with calls to bring *practice* back in to the analysis of social action

(Neumann 2002), analysts neither should nor need to invent a classification scheme from scratch. Instead, analysts can and should take their bearings from extant classificatory practices, seeking only to bring some abstract order to the sorts of things that are already and empirically going on in the social domain under investigation. Applied to the present task, this means that we should take our bearings for a classification of positions in the philosophy of science with relevance to IR scholarship from the existing contrasts and distinctions that active IR scholars in fact draw in their work. But second, analysts need not be bound simply to reproduce or redescribe extant social practices; instead, and much like the skilled craftsperson in any other field of activity, scholarly analysts can and should abstract from particular practices in order to forge more useful tools for accomplishing specific purposes (Dewey 1920, 55). Hence the challenge is not simply to get various positions in the philosophy of social inquiry "right" (whatever *that* might mean operationally). Instead, the challenge is to abstract from existing controversies so as to focus them and ultimately make them more productive, and to do so in a pluralistic way that highlights a diversity of approaches to "science" rather than seeking imperialistically to foreclose discussion by promulgating a narrow and uniform definition.

With that by way of prelude, let me now offer a methodological principle and a provisional set of distinctions that, when combined, form what I believe is a useful typology for the discussion of the philosophy of science in IR. The methodological principle is that we should regard positions on the character and conduct of science as resting on provisional commitments—*wagers*—about matters of philosophical ontology that can really never be settled definitively.[10] "What is the nature of Being?" and "What is the purpose of human existence?", to give two of the best-known examples, are the sorts of ontological/theological/ethical questions to which particular scholars give answers that depend, in the final analysis, on a measure of *faith*, precisely because they cannot be revolved empirically or rationally. But commitments of this sort undergird every instance of scientific research, implicitly shaping what the goals of such research are thought to be and how the research goes about trying to accomplish those goals. Even the most flat-footed empiricist has implicitly decided that reality is made up of tangible, measurable stuff and that true knowledge consists in discovering how that stuff is related so that knowledgeable humans can conform their expectations to those relations. It is a measure of the conceptual and philosophical poverty of the field that we rarely if ever acknowledge, let alone *discuss*, such commitments. Instead, we focus on technical application, obscuring the world-constituting wagers that animate those technical procedures.

Wagers constitute worlds, in that they quite literally set the stage for the kinds of empirical and theoretical puzzles and challenges that a scholar takes to be meaningful and important. For example, if one does not believe that the purpose of social science is to contribute to human emancipation, then the deplorable living conditions of much of the world's population at the present time, or the impacts on daily life wrought by the increasing interconnectedness of global financial markets, look very different than they do to a scholar who believes—as, for example, James Bohman does—that "the social sciences play a special role in not only reconstructing . . . communicative capabilities, but also in developing reflexivity sufficient to allow speakers to make manifest the limitations of existing discursive practices" (Bohman 2002, 507).[11] At a minimum, a wager locates and specifies three things: the researcher, the world to be researched, and the character of the relationship between them. Bohman's critical-theoretical stance, for example, separates researcher from social actors to the extent that the researcher is empowered to introduce or induce, through the practice of social science, changes in existing practices that are intended to disclose the deficiencies of those practices as ways of approximating a broad and subtle notion of democracy. It also upholds the researcher's privileged—because social-scientific—grasp on the normative goal of democracy, even if the actual working-out of that ideal in practice depends on collaboration with social actors and even if that normative ideal is transcendentally related to the actual practices of social actors rather than being handed down from some ideal realm *á la* Immanuel Kant.

To put this a slightly different way, Bohman's position combines two analytically distinct wagers. The first involves the relationship between the researcher and the world, and speaks to the question of whether the objects of study have a more or less determinate essential character that is separate from the researcher's activity, or whether the process of research in some sense constitutes the object of study *en passant*, in the course of gathering and assembling data. Critical evaluation of a set of social practices seems to call for the first answer rather than the second one, as it is difficult to conceptualize the standpoint from which a social-scientific researcher could possibly critique existing practices without some detached ground from which to launch such critiques.[12] The second wager involves the kind of knowledge to which the social scientist is thought to have access, which in this case is super-empirical or transcendental (albeit, in Bohman's case, in the complex and Habermasian sense of that term) rather than confined to the empirical or experiential sphere. Together, these two wagers produce an image of knowledge-production and an account of scholarly social-scientific practice that make possible the kind of critical emancipatory activity that Bohman argues should characterize more of IR scholarship.

Not by accident, these two wagers seem to me to constitute two of the most important commitments of philosophical ontology made by IR scholars, and suitably abstracted they provide a useful way of clarifying debates about the philosophy of science in the field. As I have suggested above, the first wager—concerning the relationship or connection between the researcher and the researched world—presents an ideal-typical choice between *mind–world dualism* and its opposite, which I will call *mind–world monism*.[13] The former option maintains a separation between researcher and world such that research has to be directed toward properly crossing that gap, and valid knowledge must in the end be related to some sort of accurate correspondence between empirical and theoretical propositions on the one hand and the actual character of a mind-independent world[14] on the other. The latter, on the other hand, maintains that the researcher is a part of the world in such a way that speaking of "the world" as divorced from the activities of making sense of the world is literally nonsensical: "world" is endogenous to social practices of knowledge-production, including (but not limited to) scholarly practices, and hence scholarly knowledge-production is in no sense a simple description or recording of already-existing stable worldly objects. But mind–world monism is no more "idealist" (in the sense of privileging ideas about the world) than mind–world dualism is "realist" (in the sense of privileging the world); it is not the privileging of one or the other side of a mind–world dichotomy that makes a position monistic, but the rejection of the very distinction in the first place.[15]

The fact that the mind–world dualist position has often been characterized as "positivist" (Wendt 1999, 39–40) while the mind–world monist position is often characterized as "interpretivist" (Yanow and Schwartz-Shea 2006) is one of those examples of a less-than-useful classificatory scheme that does not really clarify the issues at stake in the philosophical distinction. Despite the best intentions of many who use this distinction, the result of contrasting "positivist" and "interpretivist" scholarship seems to be the kind of faux "synthesis" advocated by David Laitin (2003) in which participant-observation and other experience-near modes of data collection are assigned the role of gathering raw materials for the testing of covering-law hypotheses (see also King, Keohane, and Verba 1994, 36–41). "Positivist" versus "interpretivist," like "quantitative" versus "qualitative," collapses all-too-easily into a difference of *method*, rather than a difference of *methodology*, and the key wager about our hook-up to the world made in more anthropological modes of knowledge-production is obscured. The only way to avoid this is to clarify the terms of the distinction more clearly, something that my terminological shift is designed to do, both by avoiding the "interpretivism-as-raw-materials-gathering" misunderstanding presently operative in large parts of the field and by refocusing attention on the issues of philosophical ontology at the heart of the distinction properly understood.

The mind–world dualism/mind–world monism wager, however, is not the only core wager that we need in hand in order to order contemporary IR debates about social inquiry usefully. After all, both hypothesis-testers such as King, Keohane, and Verba (in common with the majority of American IR scholars and political scientists, protestations about mechanisms and "qualitative" strategies of inference to the contrary) and critical realists such as Wight and Bohman are mind–world dualists inasmuch as they posit an external world to which knowledge in some sense approximates. But there are clearly important differences between hypothesis-testers and critical realists, issues that critical realists indicate by critiquing the restriction of knowledge to those aspects of reality that can be more or less directly observed, experienced, and measured (Patomäki and Wight 2000, 218–219; Wight 2006, 25–26). The key issue here is whether knowledge is purely related to things that can be experienced and empirically observed, or whether it is possible to generate knowledge of in-principle unobservable objects.

Following language introduced by Roy Bhaskar (1975), I will refer to the position that maintains the possibility of knowing things about in-principle unobservables *transfactualism*, since it holds out the possibility of going beyond the facts to grasp the deeper processes and factors that generate those facts (Wight 2006, 18). The opposite position, *phenomenalism* (Harre 1985, 68–86),[16] maintains, to the contrary, that it is neither necessary nor possible for researchers to "transcend experience by some organ of unique character that carries [them] into the super-empirical" (Dewey 1920, 77)—that knowledge, to the contrary, is a matter of organizing past experiences so as to forge useful tools for the investigation of future, as-yet-unknown situations (Dewey 1910, 126–127). Between them, transfactualism and phenomenalism define the parameters of this second wager.

Putting these two wagers together generates the following 2 × 2 table of commitments in philosophical ontology and the methodologies that arise from those commitments:

Table 2.1

		Relationship between knowledge and observation	
		phenomenalism	*transfactualism*
Relationship between the knower and the known	mind–world dualism	neopositivism	critical realism
	mind–world monism	analyticism	reflexivity

Fleshing out the specifics of the four cells of this table—both explaining what each of these philosophical-ontological commitments entail in greater detail, and clarifying the methodological implications of each for IR scholarship—will be the task of the remainder of this book. In the rest of the present chapter I want to sketch out, in a preliminary way, some of the issues at stake. Before I do that, however, I want to make the *status* of the typology absolutely clear. This typology is not an exhaustive account of debates in the philosophy of science; it is not even articulated in terms that philosophers of science would necessarily use to describe their own positions. It is not an intervention into debates in the philosophy of science; indeed, by conceptually placing these four philosophical-ontological combinations on something of a level playing field, I am likely to be unintentionally annoying partisans of each camp. The typology is also focused on those positions within the philosophy of science that are concerned to clarify the implications that a particular combination of ontological commitments has on the actual practice of knowledge-production, and as such more or less completely ignores thoroughgoing skepticism of the sort that would call the very possibility of knowledge-production into question (for example, Williams 1995).

Finally, the typology is also *ideal-typical* in the precise sense that Max Weber used the term: instead of a representation or a depiction, it is a deliberate over-simplification of a complex empirical actuality for the purpose of highlighting certain themes or aspects that are never as clear in the actual world as they are in the ideal-typical depiction of it (Weber 1999a, 191). In this way, my procedure shares something with Imre Lakatos' approach to the characterization of debates and controversies:

> The history of science is always richer than its rational reconstruction. But rational construction or internal history is primary, external history only secondary, since the most important problems of external history are defined by internal history . . . Internal history is not just a selection of methodologically improved facts: it may be, on occasions, their radically improved version.
>
> (Lakatos 1978, 118–119)

"A selection of methodologically improved facts" strikes me as a very good summary of what my typology contains: not detailed nuances of specific positions taken by specific people (that would be "external history"), but a purposeful summary of the conceptual and philosophical *content* of those positions. However, unlike Lakatos, whose rational reconstructions are designed for use in retrospectively evaluating whether a given scientific research programme has been progressive or has degenerated, I am not primarily concerned with *evaluating* any of the four philosophical-ontological combinations in my typology. Instead, I hope to provoke two things: a clarification of the issues involved in, and the concrete research implications of, taking up any one of these positions; and a general sense of the importance of getting our philosophical ontology straight when making and evaluating factual claims about world politics. In other words, I want to *foreground* ontological concerns, not *re-ground* the field on some particular ontological basis.

Hence, the test of my typology is ultimately a practical one. First, how useful is thinking about mind–world dualism/monism and phenomenalism/transfactualism for clarifying the relevant philosophical issues? The four methodologies contained in the typology certainly have identifiable analogues within the philosophy of science. Neopositivism, arising from the conjunction of mind–world dualism and phenomenalism, points towards hypothesis testing and the attempt to falsify general claims against empirical evidence; none of that would make much sense without the presumptions of an externally existing world against which to test claims and the limitation of the objects of knowledge to those things we can observe and measure. Broadly speaking, this is the post-Popperian tradition in the philosophy of science. Similarly, critical realism, which departs from neopositivism (and from Popper) by pushing the limits of knowledge into the realm of the in-principle unobservable,[17] stands with the neopositivists in presuming that the world exists independently—otherwise, no sense could be given to the notion of objects and relations that were "real but unobservable," disclosed through abductive inference and other similar techniques. Analyticists[18] also depart from Popperian neopositivism, but not in the same way that critical realists do. Analyticists reject the notion that in-principle unobservable relations and objects are anything but instrumental devices used to make sense of the world that we *can* observe, whether with our unaided senses or with specialized detection equipment. Thus, for analyticists, knowledge is a useful ordering of experience, and it makes little sense to formulate and test hypotheses because the idea of an externally existing world against which to test them is nonsensical. And those committed to reflexivity reject both the notion of an externally existing world and the notion that knowledge is limited to experience; instead, they ground knowledge in the social situation of the researcher, arguing that what we know is inseparable from where we are situated when we produce knowledge. This is the province of social studies of science, and of certain types of feminist and post-colonial scholarship.

However, even if the typology clarifies philosophical and ontological issues, does it do so in a way that is useful for IR? In order to answer this second, and ultimately more important question, it is necessary to consider the alternative ways of dividing up the field so as to clarify debates and controversies. One of the most curious things about IR from a philosophical perspective is that we do not generally organize the field along conceptual or philosophical lines at all; rather, we divide into schools and research communities based on substantive topics and preferred causal factors. Thus "international security" and "international political economy" name subfields in IR, subfields that are not in any meaningful way characterized by common ways of analyzing particular topics. Similarly, we have a set of "isms" that often seem, in practice, to be little more than groups of scholars who maintain that military, economic, or ideational factors exercise the most influence over the course of world politics. Then we also have lines of research united by techniques and tools: rational-choice modeling, large-n statistical analysis, qualitative case studies. In the midst of all of this empirical chaos, we lack any good and defensible way to make choices, or to evaluate the choices that other scholars make, about how research is conducted.

We might think of this as a good thing for the diversity of the field as a whole, but we should not lose sight of the fact that global diversity is quite compatible with enforced local homogeneity, whether we are talking about cultures (Inayatullah and Blaney 2004, 124–125) or methodologies. Thus one possible result of field-wide diversity is not a freewheeling and problem-driven eclecticism (Sil 2000), but instead an archipelago of small groups of scholars doing their own thing in blithe disregard of the rest of the field. The first step towards avoiding that fate, I think, is to highlight the extent to which the various methodological commitments that scholars make are, ultimately, composed of the kind of philosophical-ontological wagers I have sketched here—wagers about which there is no simple final resolution. In addition, a philosophical-ontological typology of methodologies has the merit of placing commitments in a common conceptual space, so that when we disagree we are at least disagreeing about the same or similar things. Having a commonplace about which to disagree fosters conversation, not isolation.

Finally, one might ask whether this typology actually captures any debates that are actually going on within the field of IR, or whether this whole exercise represents yet another attempt to import a set of concerns derived from outside the field in an effort to press the field in some specific direction. It would be disingenuous of me to deny that I for one would greatly prefer a more philosophically self-aware IR, a field characterized by a broader consideration of the fundamental philosophical issues that are intimately intertwined with *any* effort to generate factual knowledge. It would be equally disingenuous of me to deny that the two axes of debate that I have ideal-typically isolated in my typology are also the issues that I think we ought to be having more debates about in the field; of course they are, and anyone who claims anything different about *any* conceptual typology or distinction is most likely not being entirely forthcoming. But I do not think that this kind of objection suffices to disqualify any substantive claim; instead, what matters is how well the claim does what it is supposed to do in practice, whether that is to reflect accurately an externally existing reality, or to order lived experience usefully, or what have you.

The typology I have sketched here—and have organized the remainder of this book around—does, I think, capture current controversies within IR, even though it remains true that most IR scholars are probably located in the upper left-hand quadrant and practice some form of neopositivism (which also helps to contribute to the continued absence of debate about the philosophical ontologies I am sketching, since most IR scholars already *share* a philosophical ontology, and what is understood need not be discussed). But IR certainly also features critical realists, analyticists, and scholars pressing for increased reflexivity. If my typology helps place such scholarship on more of an equal footing with neopositivism, it will have accomplished perhaps its most pressing task.