

Lionel Robbins and broad positivism: all the philosophy an economist needs

Don Ross

School of Economics, University of Cape Town

Department of Philosophy and Department of Finance, Economics and Quantitative
Methods, University of Alabama at Birmingham

don.ross@uct.ac.za
dross1@uab.edu

1. Introduction

In my opinion Lionel Robbins's *Essay on the Nature and Significance of Economic Science*¹ remains the gold standard among general descriptions of what I will call, for lack of a phrase both more precise and more accurate, the mainstream economic attitude. I am unsure how many contemporary economists would find it surprising that Robbins might merit such acknowledgement. In my large Economics Department² there is little discussion of methodology or history of thought. More to the point, I suspect that few among my thirty-five full-time Economics colleagues³ have ever read Robbins's *Essay* or any critical discussions of it, though probably most of them could paraphrase his definition of economics.

I do not point this out as any kind of implied criticism of economists. Their chief interest is in their professional activity, that is, in applying economic principles to predicting and explaining various social, commercial and behavioural phenomena. This is a different interest than interest in intellectual history, and it naturally becomes all-consuming; goodness knows there are plenty of puzzling phenomena to keep everybody busy, and furthermore predicting and explaining these phenomena has potential practical consequences if and only if one jumps on them quickly enough to avoid recommending policy responses appropriate to yesterday's but not today's problem settings. I have said what I did about my colleagues as a lead-in to a point about Robbins's *Essay*: though I would not advise them to read the 'official' literature in 'Methodology' unless they have idiosyncratic extra-professional interests, I think that every economist should read Robbins. One way I would summarize the relevant difference here is by saying that the *Essay* is not methodology except in a negative sense. That is, from a practical point of view there are merely a few things it advises economists *not* to do: they shouldn't try to infer sweeping generalizations from quantitative historical data, or set out to empirically test the basic postulates of economics such as the general positive correlation between

¹ Throughout this paper, all citations will be to the better-known 2nd, 1935, edition of the *Essay*, which introduced non-trivial revisions to the 1st edition.

² One of the three Departments of which I am a member is a Philosophy Department and the other two are Economics Departments. Only one of the latter is large enough to be a sensible basis for the example.

³ I'm now counting both Economics Departments.

relative prices and sizes of ratios of demand to supply. The *Essay* says little about what economists *should* do, i.e., about a methodology to which they ought to conform.

To say that Robbins's *Essay* is a gold standard for something-or-other in economics, but not methodology, is consistent with the fact that most economists have a low opinion of 'Methodology'. However, this is really a semantic point, symbolized by my capitalizing 'Methodology' and putting it in quotes. An example of a document that is surely methodological in the straightforward sense of the word and yet has plenty of respect around my department is Hayashi's *Econometrics* (2000). I say Hayashi's book is 'surely' methodology because it is, after all, nothing *but* a set of recipes telling an economist how to test models against data and how to gather and measure data in such a way that they can be used to test models of the kinds economists build. Econometrics isn't economics; it is precisely the core of (contemporary) economic *methodology*.

The 'Methodology' that most economists ignore is, in fact, philosophy. Economists generally doubt that philosophy is relevant to what they do. Their instinct in this respect is sound, though the simple expression of it just given is a bit too crude. *Minimal* self-consciousness about an activity – the extent of self-consciousness needed to be able to meaningfully say “we do economics, not psychology or sociology or demography” – implies a minimal degree of philosophical sophistication. More importantly, there is a certain level of philosophical sophistication required to defend the practice of economic analysis against the legions of critics who continuously assert that economics is in principle irrelevant or obsolete or pernicious or all three.⁴ I will argue that Robbins's *Essay* sets an outstanding example of where that minimum lies. This claim will be defended in a quite specific sense: Robbins shows that mainstream practice in economics is consistent with what I will call 'positivist' assumptions, in a broad sense which I will make precise. Anti-economists are often just philosophically sophisticated enough to think that economists practice positivism and that positivism has been refuted and/or is 'bad'. Economists, thinking their foes know more philosophy than they do, are often inclined either to try to show that they aren't positivists – as the other side defines positivism – or, more typically, just ignore the noise from the sidelines and get back to work. The latter response, however understandable, concedes the public square to the ignorant. I want to bring the good news that economists can avoid having to do this by investing only a small amount of time to distract them from their main work: Robbins's *Essay* is short. The something-or-other of which Robbins is still the gold standard is: a description and defense of the distinctive disciplinary structure of economics.

If Robbins's *Essay* tells the economist what she needs to know to be armed against hostile philosophy, why don't I just stop now, leaving off with the injunction to economists read Robbins? The answer is that the *Essay* has itself been extensively filtered through philosophical readings, and, in particular, has been read as if it were 'Methodology' – notwithstanding Robbins's explicit disavowals of this in the text (p. xv and, most clearly, p. 72). Historically influential exercises of this kind are books by Mark

⁴ It may be wondered how something could be both irrelevant and pernicious. This is a question that can only be asked of anti-economists, who routinely combine these charges. See Coleman (2002) for a baleful, but excellent, tour of instances.

Blaug (1980, chapter 3) and Bruce Caldwell (1982, chapter 6). I confess to having added a contribution to this confusion myself, in a way I will indicate later. My latest re-reading of the *Essay*, however, has convinced me that it offers reflections on the practice of economics that stop short of ‘Methodology’, but are consistent with some leading currents in contemporary philosophy of science, and furthermore are still essentially in accord with what economists do.

2. ‘Methodology’ and the philosophy of science

In general, Blaug and Caldwell agree in the way they fit Robbins’s *Essay* into the history of economic thought. According to them, he preserves core themes of the main tradition of English neoclassicism but makes these themes subservient to a version of a priorism about the fundamental assumptions of economic theory that he gleaned from the Austrians. In holding that basic economic postulates are known intuitively and are not open to empirical disconfirmation, Robbins is taken by both Blaug and Caldwell to be the natural foil for the ‘logical positivism’ of Terence Hutchison, who wrote his own book on methodology just a few years later (1938). Hutchison indeed intended Robbins as his foil. Whether Robbins really *is* an appropriate foil for a doctrine deserving the name ‘logical positivism’ is much more doubtful, as we will shortly see. The main difference between Blaug and Caldwell is that whereas the former is scathingly hostile to the Austrians, especially von Mises, on questions of methodology, Caldwell is relatively sympathetic, and provides good answers on von Mises’s behalf to Blaug’s unconcealed disdain (Blaug 1980, p. 92; Caldwell 1982, pp. 117-124).⁵ This difference seems largely to stem from Blaug’s wholehearted endorsement of Popperian falsificationism as the right methodological programme for economics, an enthusiasm Caldwell, who calls himself a “methodological pluralist”, does not share.

Thus Blaug and Caldwell offer a common interpretation of Robbins but divergent evaluations. I will shortly engage with the substance of the common interpretation. First, however, comments are in order about the larger narrative into which it is set. Much of the philosophy of science known to a typically trained contemporary economist comes directly or indirectly from Blaug’s account. It is thus unfortunate that the philosophy both Blaug and Caldwell describe as having influenced the history of economics is almost exclusively anglophile. Blaug describes the importance of Bentham and Mill for Jevons and Marshall, but says nothing about philosophical influences on Walras or Menger. With respect to the next generation of economists, Blaug and Caldwell tell us that the Austrians influenced Robbins, but nothing is related about who influenced the Austrians. Then both say that Hutchison absorbed the views of the logical positivists from the Vienna Circle, and indeed that Hutchison learned his logical positivism while lecturing in Germany. However, what is then presented as the doctrine of logical positivism is *not* the positivism of either the Vienna Circle or of their close associates in Berlin led by Reichenbach, but, roughly, the version introduced to English-speaking intellectuals by A.J. Ayer in his *Language, Truth and Logic* (1936). I will discuss the differences between these in the next section. As Michael Friedman (1999) has shown, though most philosophers know that Ayer distorted the views of the logical positivists in popularizing

⁵ No one ever says mean things about Robbins.

them, the relevant distortion is seldom fully factored out in non-specialist accounts. In particular, postwar commentators generally project back into logical positivism aspects of British empiricism which were at best a minority view among the members of the Vienna Circle. The great distal influence on the Vienna Circle, Friedman demonstrates, was Kant, not the classic British empiricists. The Austrian economists, likewise, developed their views in an intellectual environment dominated by neo-Kantians. Since Robbins was indeed heavily influenced by the Austrians – first Menger and von Mises and then Hayek – presenting him as a foil for logical positivism is not appropriate. The location of this blind spot is unsurprising in light of the fact that Kant does not occur in Caldwell's index, and is mentioned by Blaug only incidentally and tangentially, as indeed are the logical positivists themselves.

The problem with the narratives of Blaug and Caldwell goes deeper than the objection that their shared history of affinities between philosophy of science and economic methodology has large holes in it. Both write, in general, as though economic methodology were developed under tutelage of philosophers – though they implicitly exempt both the Austrians and Robbins from this by leaving out Kant. (It helps Blaug in setting up von Mises as a methodological lunatic that one might gather from the former's account that von Mises had no philosophical tutor but could have sorely benefited from one.) It is, in my view, more accurate to see the history of economic thought as largely evolving endogenously, with philosophers' doctrines only sometimes invoked when they served to buttress principles already adopted for *economics-driven* reasons. As I suggested in the opening section above, economic thought becomes 'Methodology' precisely when it is philosophy applied to economics. So Blaug and Caldwell are certainly engaged in Methodology. So was Hutchison, since he explicitly urged the importation of what he took to be logical positivist principles into economics. But those figures who have had the largest influence on economists' working understanding of their practice – Walras, Marshall, Edgeworth, Pareto, Fisher, Hayek, Keynes, Hicks, Samuelson, the Cowles Commission, Friedman, Arrow – were at most occasionally influenced in tangential ways by philosophers. In their work, methodology does not come apart from first-order economic theorizing. Robbins's description of the economist's role reflects this, and thus mirrors no *specific* doctrinal (positive) philosophy of science. Note that Robbins mentions *not one* philosopher in the *Essay*.

Despite this, I am indeed about to contrast Blaug's and Caldwell's version of the history of logical positivism with the correct history, and then to endorse Robbins's defense of economics by reference to that history. I do this not because I believe that the way to make sure we do economics well is to find the most persuasive philosophy of science and apply it. My aim is instead to use reflections on philosophy to *rationalize* Robbins's picture of economics, to show that his economics-driven motivations are not undermined by sophisticated philosophy. For two centuries philosophers of science have ranged themselves along a continuum of views that stretch from strong Kantianism at one end to radical empiricism at the other, with most actual positions blending elements of these positions in subtly different ways. Positivism in a sense broad enough to take in both the Vienna Circle and the post-war logical empiricists covers almost this whole continuum. *Any* position on the continuum, I will conclude, is an adequate setting from which to

defend economics against its populist attackers in approximately Robbins's way. Thus economists don't need to have dogs in the fights when broadly positivist philosophers of science disagree. Robbins, far from occupying a tendentious Methodological position, addresses us from a secure neutral ground; or so I aim to show.

3. Logical positivism and standard history of twentieth-century Methodology

Economic Methodologists are hardly alone in having an historically confused understanding of logical positivism. The version of that doctrine found in Blaug's and Caldwell's books also occurs in some work by leading philosophers of science (e.g. Giere 1988). I am thus not suggesting that Methodologists should be ashamed of their scholarship. However, the fact that most economists who have any view at all about the logical positivists entertain a caricature of them is worth correcting.

Caldwell (1982, Chapter 2) says relatively little about the beliefs of the 1930s logical positivists, despite devoting a chapter to them. Most of it is given over to introducing the so-called verifiability criterion of meaningfulness. According to that idea, cognitively meaningful bodies of discourse are distinguished from analytic and cognitively meaningless ones – including empty metaphysical ones – by having recognized testable empirical implications that can in principle (not necessarily in fact) be checked through observation or experiment. Caldwell recognizes that between the 1930s and the 1950s this idea underwent important modifications under pressure of internal debate. By the time those logical positivists who fled from the Nazis to America had re-branded themselves as 'logical empiricists' in the 1950s, it is correct that, as Caldwell reports, they had concluded that an individual proposition may be significant even if it has no testable consequences all by itself, but only in conjunction with other propositions. However, it is misleading to suggest, as Caldwell follows many commentators in doing, that in the 1930s the logical positivists had agreed that all individual non-analytic cognitively significant sentences must by themselves imply distinctive in-principle observations. *Some* of them thought this, but, as Friedman (1999, pp. 148-151) argues, the pre-eminent logical positivist, Rudolf Carnap, never did.

Caldwell rightly attributes to the early logical positivists the core belief that science is *unified* – that is, that if an accepted statement of any one science has logical implications that contradict those implied by an accepted statement in another science, a mistaken inference has been made somewhere, or a putative observation has been misreported or misinterpreted. Furthermore, completeness is a regulative ideal of the sciences together. Under pressure from recent philosophers of science, Nancy Cartwright (1999) and John Dupré (1993), this view has lately become controversial, as it was for different reasons in the 19th century. During most of the 20th century, however, it was accepted by nearly all philosophers of science, including those who disagreed with the logical positivists and logical empiricists about much else.

The rest of Caldwell's brief discussion of logical positivism simply sets the stage for his subsequent chapter on logical empiricism. This has two themes over and above elaboration on verificationism:

1. Theoretical terms in scientific theories – that is, terms making reference to in-principle unobservable entities such as sub-atomic particles – are legitimate objects of predication and generalization just in case statements that employ them have as deductive consequences sets of statements that make reference only to observable states of affairs. In the standard philosophical usage, theoretical terms must be *reducible*. According to Caldwell, what statements containing theoretical terms were held to be reducible *to* were statements in languages containing only ‘physical’ predicates, meaning (roughly) terms that assign individual objects and events to classes on the basis of operationally measurable properties. Caldwell says that this ambition, which he attributes to Carnap, “won out” over an alternative view defended by another leading logical positivist (Neurath) that the reductive basis should instead be statements reporting ‘raw’ sense data (e.g., ‘red here now’). As we will see, this is not quite correct.
2. Explanation of singular events or standing functional relationships by generalizations (‘scientific laws’) is deductive, and is prediction in reverse. That is, to be explained an event or relation must be deducible from a law conjoined with some set of in-principle observable initial conditions specified as potentially relevant by the law in question. Caldwell is aware of a large literature – which in fact had its most profuse flowering during the 1980s, after his book was published – on the necessary and sufficient logical properties of putative laws. It was generally recognized that an adequate account of laws must allow that many laws merely state frequency dependencies, not only to leave room for social, evolutionary and ecological sciences but also because predictions of quantum physics are irreducibly statistical.

These two doctrines, along with the verificationist theory of meaning, are characterized by Caldwell as conditioning mainstream economic methodology after World War II. The point of his title, *Beyond Positivism*, is that work by philosophers of science from the mid-1950s showed that such methodology is unduly restrictive. Thus he urges economists to break out of these narrow strictures.

Caldwell’s account is a partial representation of the common image of the logical positivists, which Friedman (1999) characterizes as follows:

The positivists – so this story goes – were concerned above all else to provide a philosophical justification of scientific knowledge from some privileged, Archimedean vantage point situated somehow outside of, above or beyond the actual (historical) sciences themselves. More specifically, they followed the lead of the logicist reduction of mathematics to logic, where the latter is also understood as fundamentally foundationalist in motivation and import. Just as the logicists attempted to justify mathematical knowledge and place it on a more secure foundation by means of a derivation from (supposedly more certain) logical knowledge, so the positivists attempted to justify empirical science and place it on a secure foundation by logically constructing the concepts of

empirical science on the basis of (supposedly more certain) data of sense. Thus formal logic furnished the foundational enterprise with the required Archimedean standpoint located outside of the actual (historical) sciences themselves, and phenomenalist reductionism, carried out rigorously using the methods of formal logic (as epitomized in Carnap's [1928] *Der logische Aufbau der Welt*), then provided the desired epistemological justification of the sciences (pp. 2-3).

This standard interpretation reads the logical positivists as radical empiricists from the beginning. According to it, they implemented the project that had been promoted by Bertrand Russell in his *Our Knowledge of the External World as a Field for Scientific Method in Philosophy* (1914), which he in turn inherited from Berkeley and Hume. What are knowable according to this programme are primitive atomic facts of one sort or another. Justification of the arcane reaches of scientific theory must then rest on strict derivation from these atoms, which modern mathematical logic was thought to make possible for the first time. However, the story goes on, Carnap's attempt to construct atomic physical states of affairs out of phenomenal atoms in the *Aufbau* failed – and this was indeed that project's most notable, though negative, lesson. Thus the ambition of empiricism was tempered, with the later Carnap and his colleagues attempting only to ground science in atomic physical observations rather than the supposedly more certain reports of sense data. The history of the evolution of logical empiricism from logical positivism is thus told as a movement from phenomenism to operationalism. This standard account of the main currents in mid-twentieth-century epistemology can be presented as isomorphic to the development of economic Methodology, also on a standard reading, with some temporal lag.

Blaug and Caldwell read Robbins as providing the final statement of a traditional neoclassical view according to which the basic postulates of economic theory are truisms furnished by direct introspection. In Robbins's specific case, what is taken to be known in this way by each person is that they subjectively rank possible and actual states of the world (or, at least, such partial states of the world as come within their perceptual purview) with respect to their ordinal preferability. The individual infers from what others report that people do this generally. An anticipatory element of the standard reading of positivism can then be identified in Robbins's rejection of the possibility of interpersonal comparisons of utility, since this seems to rest on the claim that one person can neither perceive nor construct from perceptible data another person's relative intensities of preference. It is bound to be objected that the resulting overall picture is unstable: why should we suppose that a person can know that both she and others introspectively order preferences, but can in principle infer nothing about relative interpersonal magnitudes of these preferences? If one wants to tell a story of progress, one can maintain that this tension was duly dissolved by the time of the next document of major Methodological importance in the history of mainstream economics, Samuelson's *Foundations* (1947). Here all recourse to intuitive knowledge in economics is denied, as foundations in introspection are banished with the turn to *revealed* preference. The development of Methodology, one can say, thus mirrors the evolution of positivist / empiricist philosophy of science in beginning from phenomenism and ending with

operationalism – that is, with *conceptual* reduction of the elementary theoretical aggregate of consumer theory, demand, to physically measurable sequences of choices by subjects. ‘Utility’ now becomes idle as a theoretical type – all that is required is that agents’ consumption behavior be interpretable as maximizing a function U representing a binary relation \succeq on a set X satisfying completeness and transitivity, where $U : X \rightarrow \mathfrak{R}$ represents \succeq if for all $x, y \in X$, $x \succeq y$ if and only if $U(x) \geq U(y)$. *Evidential* foundations go the other way: the basic data of consumer theory on Samuelson’s account are aggregate demand schedules.

In my own rationalization of twentieth-century Methodology in my (2005) book, I endorsed a version of the above story. I say a ‘version’ because, on the basis of the considerations raised by Friedman (1999), to be reviewed below, I pointed out that Robbins’s position was not antithetical to early logical positivism, as the standard account presupposes, and was indeed strongly resonant of key aspects of it. Linking Robbins to a version of logical positivism in turn helps to support a view of economic theory as involving no epistemological discontinuity, but as instead following a continuous line of development from Edgeworth, Pareto and Hicks through Robbins to Samuelson and beyond, which at every step involved reducing dependence on psychological hypotheses. I continue to endorse this view. However, I now think that I accepted too uncritically the assumption made by Blaug and Caldwell that Robbins must be read as asserting that people know they order their preferences by means of *introspection*. Since it is entirely untenable to imagine that contemporary economics – or any other science – has its foundations in introspection, accepting this assumption entailed my concluding that Samuelson’s operationalism *improved on* Robbins’s epistemology of economics. I amplified this mistake by not taking Friedman’s message fully enough to heart either, which led me to represent Samuelson as also a ‘better’ positivist than Robbins. As we will see, a more charitable epistemological rationalization of Robbins is possible, one on which Robbins’s and Samuelson’s foundational remarks are entirely complementary, and on which (incidentally, so far as economics is concerned) Robbins has as at least as much claim as Samuelson to unqualified membership in the positivist club.

4. Broad positivism

I will argue that Robbins and Samuelson are both broadly positivistic in the minimal philosophical commitments that should be attributed to them. So, I claim, are mainstream contemporary economists; and furthermore I think that this is a wise attitude on their part. To begin to defend these claims, I must first say what ‘broad positivism’ is. If this is not to be an exercise in tendentious labeling, then ‘broad’ positivism had better have some clearly stated set of relationships with actual, historical positivism, as well as with Robbins, Samuelson, and current practice in economics.

Friedman describes the standard account of early logical positivism, summarized in the earlier quotation from him, as “an almost total perversion of [their] actual attitude” (1999, p. 3). Their starting point, he reminds us, was “*rejection* of all ... philosophical pretensions”. By ‘pretensions’ Friedman refers to the idea that philosophy can stand

outside of or ‘above’ science – science in general, or any particular science – and ‘justify’ it. This modesty allows space for two possible attitudes to the traditional aim of metaphysics, the aim of identifying what is ‘ultimately real’. One possibility is to hold that *science* tells us directly what is ultimately real; this is the position known as ‘scientific realism’. The other possibility is to deny that questions about ‘ultimate reality’ ask anything of possible significance *once* we grant that the epistemological role of science (however precisely ‘science’ is delimited⁶) answers to no external (or Archimedean) court of appeal; science is our body of institutions and practices for collective objective inquiry and metaphysics is not part of that body. The traditional aim of metaphysics is therefore not to be taken over by science, but should be rejected as a deluded enterprise. ‘Broad positivism’ is the name I use here for the second of these attitudes.

Logical positivism, being a version of broad positivism, must of course be more specific than broad positivism. But as anyone who reads the Vienna Circle and their colleagues will quickly recognize, logical positivists disagreed among themselves. So on the one hand some doctrinal commitment must be identified if logical positivism is to signify anything; but on the other hand historical facts block any effort to go beyond certain limits of approximation. One thing all logical positivists had in common is that they believed in the philosophical importance of using formal logic to *construct* concepts out of others. More specifically, they endorsed Russell’s “supreme maxim in scientific philosophizing”: “Whenever possible, logical constructions are to be preferred to inferred entities” (Friedman 1999, p. 117). This attitude promotes conceptual economy and elegance, of course, but more importantly for the logical positivists it promotes *unity*: logical constructions make explicit the structural relationships among objects of discourse.

This point is the hinge by which Friedman justifies the procedure that historical considerations otherwise suggest, namely, understanding logical positivism as a whole mainly by reference to the development of Carnap’s thought, while acknowledging that Carnap’s opinions never constituted a ‘party line’ for the Vienna Circle. Since space requires a drastic simplification of Friedman’s complex picture of logical positivist thought – especially given that my main interest here is in economics rather than philosophy – I will carry his approach to extremes and write as if Carnap and logical positivism were synonymous.

Here, then, is the outline of the intellectual biography of Carnap according to Friedman. When Carnap moved from physics to philosophy at the time he began his doctorate, he inherited his problem space from Kant. That is, he took the aim of epistemology to be to show how scientific knowledge can be *objective*, despite the fact that, as Hume had argued, individual judgments about empirical, contingent matters of fact are subjective. Kant had tried to ground objectivity in what he called the *synthetic a priori*. This is a domain of intuitively grasped propositions in terms of which all experiential judgments are categorically framed. Grasping the content of these judgments is thus a precondition for any objective knowledge according to Kant. His most famous example of a synthetic

⁶ For a specific proposal see Ladyman and Ross 2007, chapter 1

a priori truth is that physical space has the structure of Euclidean geometry. In order to grasp Newton's laws, Kant argues, one must first conceive Euclidean space and recognize *that* this space describes the physical world. The centrality of this example in Kant's thought caused a crisis in European philosophy when alternative non-Euclidean geometries were developed and then applied by Einstein to overturn classical mechanics. First special relativity entailed rejection of Kant's claim that specifically *Euclidean* geometry is an intuitive prerequisite to physical understanding. Then general relativity more radically called into question the idea that formal geometry can be considered separately from physical (or 'applied') geometry at all.

Before we leave Kant, two further observations about his philosophy are important for what follows.

First, Kant regarded his epistemology as *anti*-metaphysical. That is, he did not think that philosophical speculation preceded or constrained the boundaries of empirical science. His perspective was 'transcendental' only in the sense of being concerned with grounds for *possibility*. He *assumed*, rather than *aimed to show that*, classical physics had discovered objective truths. Then he wondered what made this achievement possible. Thus, even with Kant – so well before we introduce any elements associated with positivism – we have entered the intellectual territory in which science is epistemologically prior to philosophy.

Second, the idea of *objectivity* for Kant, and for every major philosopher he subsequently influenced, depended on a strong distinction between psychology and logic. Psychology, for Kant, studies the mechanisms of *subjective*, partial and idiosyncratic perception and response. Logic, by contrast, concerns itself with reasoning that is disciplined by norms. It encapsulates the kinds of responses that careful observers would offer when reasoning critically together. A main source of the over-blown skepticism to which empiricism led Hume, in Kant's view, was the latter's assimilation of logic to psychology. Now, Kant knew nothing of what *we* call 'logic', that is, the science of the foundations of algorithmic computation, which the logical positivists initially followed Russell and Frege in mistakenly believing to also be the science of the foundations of mathematics. Indeed, it was precisely the new conception of logic that inspired the positivists to believe there was a straightforward way of rescuing the core ambition of Kantian epistemology from the demise of the synthetic a priori indicated by the revolution in physics.

Carnap responds to the crisis in post-Kantian thought by accepting (along with other important contemporaries, especially Poincaré) a *conventional* element in the choice of mathematical frameworks used for what we might anachronistically call 'modeling'. With respect to concern for objectivity, conventions have the useful property of being *intersubjective* – that is, being stabilized *as* conventions within a scientific community. But since it is objectivity of *empirical* knowledge in which we are interested, intersubjectivity of conventions helps only to the extent that conventions help to fix – through what the positivist philosopher Reichenbach called 'axioms of coordination' – the physical domain of reference of formal elements. Here is where Carnap finds the crucial philosophical value of formal logic. Logic – for the early, though not the later –

Carnap is *not* conventional.⁷ If logic can be shown to determine, through constructions, the types of objects that feature in different conventions *and* the types of objects that feature in empirical reports and generalizations (what Reichenbach called ‘axioms of connection’), then (i) conventions can be objectively distinguished, and (ii) the conventional part of science can be distinguished from the empirical part. Logic and the recognition of the role of convention in science replace, according to Carnap, the synthetic a priori in explaining the possibility of scientific objectivity.

This approach to philosophy is radically unlike the project of foundationalist empiricism, in at least three general respects. First, it does not attempt to found all of science on the basis of empirical observations. The logical positivists *rejected* the idea of responding to Einstein’s achievement by attempting to build an empiricist account of geometry (Friedman 1999, p. 60). Second, logical positivism is holistic rather than atomistic. The objects of philosophical study are whole mathematical frameworks and their relationships to bodies of theory and observation reports, not elementary constituents of reality. (Geometrical claims, positivists agree, cannot even be evaluated apart from physical claims [Friedman 1999, p. 84].) Third, logical positivism is not about *justifying* the truth of scientific theories by reference to a supposedly indubitable basis. This is the precise sense in which logical positivism is anti-metaphysical. The philosopher’s task according to Carnap is not to show that the types featuring in scientific theories are ‘real’ because they can be built out of atoms the philosopher has independently argued are real. Instead, the philosopher’s task becomes *part of* the task of science: elucidating, through rigorous logical constructions, the *structural* relationships amongst conventional mathematical frameworks and bodies of statements reporting and generalizing empirical findings. It is because there must *be* such relationships in order for science to be empirical in the first place that Carnap was committed, like all of his Vienna Circle colleagues, to the idea that statements that *seem to be* to about objects of experience but have no logically derivable testable consequences are to be rejected as meaningless. This is not claimed *for the sake of* providing a basis for rejecting traditional metaphysical statements as meaningless; rather, it turns out that they do have his character and so they *are* meaningless for that reason.

Carnap set out to exemplify this new role for the philosopher in his first great book, the *Aufbau* (1928 / 1967). It is on the basis of superficial readings of this work, Friedman argues, that Carnap specifically, and thus logical positivists more generally, are read as classic empiricists. For in the *Aufbau* Carnap indeed sets out to construct first physical objects, and then higher types, out of the elementary phenomena of sense experience. As Friedman carefully shows, the point of this was to demonstrate the method and power of logical construction at work, in an arena familiar to philosophers. Carnap does not view his philosophical programme as hostage to the success of this specific demonstration. Thus, having logically constructed enduring coloured patches on the perceptual manifold out of fleeting colour flashes, he is relatively casual and inexact in his construction of coloured external objects from this basis. Furthermore, he is nonchalant about the

⁷ That the later Carnap *did* decide that even logic is conventional, without this leading him to perceive any deep philosophical crisis, helps to show how far one can get from foundationalism without leaving the domain of logical positivism (let alone what I am calling ‘broad’ positivism).

question of whether the most useful language of empirical reports is a phenomenal language or a ‘thing’ language – this is taken to be a matter for a conventional judgment, to be informed pragmatically by the progress, or lack of it, of projects in logical construction. Years later he said

When I developed the system of the *Aufbau*, it actually did not matter to me which of the various forms of philosophical language I used, because to me they were merely modes of speech and not formulations of positions ... The system of concepts was constructed on a phenomenalistic basis ... However, I indicated also the possibility of constructing a total system of concepts on a physicalistic basis ... The ontological theses of the traditional doctrines of either phenomenalism or materialism remained for me entirely out of consideration (Carnap 1963, p. 18).

Friedman emphasizes that the project of the *Aufbau* is closer to Kant’s than to the classical empiricist project because it takes *structure* to be the basis of objectivity: “Scientific knowledge is objective solely in virtue of its formal or structural properties, and these properties are expressed through the ‘places’ of items of knowledge within a single unified system of knowledge” (Friedman 1999, p. 98).⁸

As Friedman stresses, Carnap’s structuralism brings him into close affinity with his leading contemporaries among *avowed* neo-Kantians, philosophers whom no one has ever regarded as classic empiricists. The most important of these figures is Ernst Cassirer. His version of structuralism is especially interesting because its later versions were directly motivated by attention to quantum mechanics, which ought to play at least as great a role as relativity theory in motivating any contemporary body of opinion on metaphysics and epistemology. As French (1999, 2000) and French and Ladyman (2003) argue, Cassirer understood individual objects in modern physics to be ‘constituted’ group-theoretically as sets of invariants under symmetry transformations. Consideration of field theory leads Cassirer to an expression of structuralism that very closely resembles Friedman’s reconstruction of Carnap’s structuralism:

The field is not a ‘thing’, it is a system of effects (*Wirkungen*), and from this system no individual element can be isolated and retained as permanent, as being ‘identical with itself’ through the course of time. The individual electron no longer has any substantiality in the sense that it *per se est et per se concipitur*; it ‘exists’ only in its relation to the field, as a ‘singular location’ in it (Cassirer 1936, p. 178).

One cannot possibly get further from the atomism of classic empiricism than that. Friedman thus maintains that at least the best-developed expression of logical positivism, Carnap’s, has more in common with neo-Kantianism than with radical empiricism. Suppose we then agree to usher Cassirer’s structuralist neo-Kantianism, along with logical positivism, into a ‘broad positivist’ tent, from which we exclude foundationalists

⁸ Philosophers will want to see this idea expressed a bit more precisely. The idea is that all scientific concepts can be expressed through purely structural definite descriptions (Friedman 1999, p. 103).

and metaphysical atomists. Who else can we bring into this tent? Carnap himself suggests extreme liberalism:

The so-called epistemological schools of realism, idealism and phenomenism agree within the field of epistemology. Construction theory represents the neutral foundation which they have in common. They diverge only in the field of metaphysics, that is to say ... only because of a transgression of their proper boundaries (Carnap 1928 / 1967, p. x).

Note that Carnap includes 'realism' here. He is referring to realism in the sense rejected as 'metaphysical' by Kantians, that is, a realism that "attempts to base objectivity on the relation of sensory data to a 'transcendent' object existing somehow 'behind' the data" and which "creates an unbridgeable gulf between thought and reality in virtue of which objective judgments are just as impossible for us as they are on a strictly empiricist or 'positivist' conception" (Friedman 1999, p. 126). It should be obvious from what has already been said as to why and how Carnap sees the empty dispute between this sort of realist and the phenomenalist dissolved by his new conception of philosophy. But what about the kind of realism that became more common in philosophy during the final quarter of the twentieth century, 'scientific realism' – the view that well-confirmed scientific conclusions should be read as direct, true claims about reality? Must broad positivism be opposed by all philosophers who consider themselves scientific realists?

A main motivation for scientific realism since its renaissance in the mid-1970s was reaction against the perceived implications of logical empiricism's holism about meaning, when *combined with* presumed atomism about the objects of scientific belief. (Ironically in light of the discussion so far, the holist aspect was then widely thought to have been a *later* logical empiricist amendment to original logical positivist atomism.) The worry was as follows. Suppose that the meaning of a term for a type of entity, such as 'electron', is determined by its role in physical theory. It then follows that when physical theory is revised in light of new discoveries, the meaning of 'electron' must change. This in turn seems to deprive us of any basis for saying that our understanding of *electrons* has improved, since we lack grounds for saying that the discarded theory and the new theory are rival accounts of the same objects. But this implies an absurd history of science. The distinctive core claim of the scientific realist is that a term such as 'electron' refers to a theory-independent constituent element of reality with whatever properties electrons *in fact* have – as opposed to whatever properties passing theoretical *descriptions* of electrons *take them* to have. This allows us to say that later physicists offered more accurate models of electrons than their predecessors.

The rise of scientific realism has inspired a revival of unabashed metaphysics in much of philosophy. Ladyman and Ross (2007), following and supplementing van Fraassen (1980), point out that, however well motivated the need for a plausible account of scientific progress may be, the grounds for rejection of strong metaphysics raised by past great philosophers, including Hume, Kant, Russell and the logical positivists, remains as much in force now as ever. Strong metaphysics is in no way supported by the positive

content of science, and presupposes the possibility of objective knowledge being achieved through exercises of pure reason and / or intuition that are independent of controlled observation and experiment. This is extravagant hubris and profoundly incompatible with the scientific attitude.

Realist worries about the basis for ascribing progress to science do indeed arise *if* the basic subjects of scientific generalizations are taken to be the kinds of *objects*, such as electrons, about which successive scientific theories have disagreed through history. However, Ladyman and Ross argue, building on earlier work by Ladyman (1998, 2002) and French and Ladyman (2003), that the basis for reading the history of science as one of radical discontinuities (along lines made popular by Kuhn and others) collapses if the content of science is understood on *structuralist* lines. New theoretical proposals taken seriously by scientists generally enrich and extend the mathematical structures developed by earlier scientists studying what they take to be the same phenomena. Apparent discontinuities at the scale of putatively self-subsistent objects tracked by particular, situated observers from limited, subjective perspectives should worry us philosophically only to the extent that scientific theories – as opposed to everyday folk ontological principles – endorse such objects. However, the progress of fundamental physics (that is, quantum physics, especially field-theoretic quantum physics) has strongly confirmed Cassirer’s basic insight on this question. The massively confirmed fact of quantum entanglement constitutes scientific refutation of the objectivity of the folk model of the world as ‘made of’ ‘little things’ that ‘occupy space’, ‘endure through time’ and change one another’s ‘accidental properties’ by banging into one another. These metaphors derived from the manipulation of everyday objects by people find no counterparts in modern mathematical physics.

We call our structuralism a kind of ‘realism’, but this is mainly to contrast it with the extreme empiricism of van Fraassen, which involves implausible denial of the possible justification of what philosophers call ‘objective modalities’. In fact, our position is very close to Carnap’s as reconstructed by Friedman. The very aspect of Carnap’s mature philosophy⁹ which, according to Friedman, makes it ultimately untenable is the one place where we introduce a novelty that has no precedent in logical positivism. To avoid the kind of objection that, according to us, makes van Fraassen’s empiricism incapable of accounting for objectivity, Carnap depended on the distinction between ‘analytic’ and ‘synthetic’ statements,¹⁰ which he was required to interpret in a way that was fatally undermined by Gödel’s incompleteness results. (For details, see Friedman 1999, chapters 7-9.¹¹) Where Carnap’s structuralism relies on analyticity to illuminate so-called ‘modal

⁹ This refers to Carnap (1934 / 1937).

¹⁰ The former are sentences held to be true (within a language) solely by virtue of the meanings of the terms they contain plus their logical structures. The latter are claims for which truth-value depends on facts stable at the so-called ‘object level’ for the language in question, which must be discovered and verified empirically.

¹¹ Friedman’s account of the downfall of logical positivism is not as exotic as this brief description might make it appear. He echoes the common view that Quine’s (1951) attack on the analytic / synthetic distinction was the deadly blow. What is novel is Friedman’s contention that Carnap is caught by Quine’s argument because of the way in which he has to conceive of analyticity if Gödel’s results are not to make his project come out circular.

structure' in science, ours appeals instead to a distinguished ('fundamental') body of physical generalizations which, if quantum theory is true, constrain every measurement taken at every scale everywhere in the universe. Should quantum theory turn out not to be true and universally applicable then our philosophy loses a premise on which it absolutely depends. But that is just the sort of state of affairs that *should* prevail in philosophy if it is antecedent to, rather than prior to, science.

We refer to our enterprise as 'naturalization' of, rather than 'rejection of', metaphysics. In the present context, however, this difference between us and other broadly positivistic philosophers is also mainly semantic. 'Metaphysics', on our account, is heavily deflated by comparison with its usual significance. We mean by it only an enterprise that aims to unify the various special sciences – another goal we emphasize in common with both Kantians and logical positivists. We argue for a *sense* in which the scientific realist is right that *fundamental physics* tells us directly what exists. But most of science (including most of physics) is not fundamental physics. Such deflated metaphysics as we allow is of no direct consequence for economics. Economists may not contradict the accepted generalizations of physics, but this has no bearing on any serious questions that preoccupy economists. Thus where special sciences like economics is concerned, our view is broadly positivistic.

This sketch of our new version of comprehensive structuralism of course cannot carry significant argumentative weight in this breathless summary form, shorn of all details and specific argument. I introduce it here merely to establish the point that broad positivism – denial of the objective significance of extra-scientific metaphysics, based on a structuralist understanding of science – remains very much a live option in contemporary philosophy of science. The reader can consult alternative versions to ours in Friedman (2001) and Maddy (2007).

Broad positivism commits philosophers to trying to explain why the scientific enterprise as a whole divides into sub-provinces – the different special sciences – by reference to distinctive mathematical structures, notwithstanding the fact that all sciences describe a common universe. By this we intend no metaphysical stipulation – we merely reflect the observation that all sciences are regulated by a principle that prohibits them from contradicting currently consensual generalizations of fundamental physics. Within the broadly positivist framework, we should expect a scientist writing about the nature of her discipline to aim at elucidating its general structural parameters – the parameters that make it the discipline it is, rather than some other discipline. If her science is mature, we should find her doing this using mathematics, the only representational technology both rich and general enough to allow for contrastive comparison of her disciplinary structure with others. We should *not* expect her to try to improve her own or anyone else's understanding of her discipline by *applying to it* a tendentious epistemological or metaphysical doctrine she has gleaned from a philosopher, based on principles that originate from outside the hard-won experience gained in *using* her discipline's structure to achieve new empirical discoveries.

With this perspective in mind, let us now return to Robbins's *Essay*.

5. Robbins as positivist

If we ask what philosophical attitude best describes Robbins's account of economics in the *Essay*, we can see that the answer is positivism, as described above. It will be sufficient for my purposes to pronounce his position as *broadly* positivistic; to try to find outright *logical* positivism would involve over-reading the text. On the other hand, I think there is solid basis for saying that Robbins is *closer to* logical positivism than was Hutchison, whose empiricism resembles Ayer's. This should not be regarded as historically mysterious or even surprising, as I suggested earlier. Robbins was strongly influenced by the Austrian economists, all of whom were extensively exposed to neo-Kantian philosophy. In his lectures on the history of economics at LSE, Robbins singled out Menger among the first marginalists as the one "who perhaps insensibly has shaped the view which that conception [of scarcity] occupies at the foundation of our subject" (Robbins 1998, p. 277), and in the *Essay* he refers to the birth of marginalism itself as "the Mengerian revolution" (Robbins 1935, p. 106). Menger's son was a member of the Vienna Circle, who debated the conventionality of logic with Carnap. Robbins spent time and lectured in Vienna, was close to von Mises and Hayek, and was the co-originator of the plans that brought the leading logical positivists, among others, out of the Nazi catastrophe. Again, the point is not to suggest that Robbins explicitly embraced (second hand) logical positivism and applied it to economics. The point is merely that the entire style of thinking – especially, in this instance, about mathematics, the nature of ideas, and the concepts of objectivity and logical priority – that Kant bequeathed to everyone educated in German-speaking universities were part of Robbins's intellectual background to an extent that was unique among major British economists.

The *negative* part of my last point – that Robbins was of the logical positivist mindset but not a positivist *philosopher* – indeed needs special emphasis. As we have seen, the positivist attitude to the elucidation of science precisely enjoins that one *avoid* trying to 'found' the structure of inquiry in a discipline on general epistemological (let alone metaphysical) principles. One important aspect of Robbins's positivism lies in the fact that all of his key premises in the *Essay* are developed from the immediate history of problem-solving in economics that he inherited. None are fundamental philosophical assumptions. Of course, if this were *all* it took to be a broad positivist, then broad positivism would be too watery a notion to be worth making much of. The positivist scientific elucidator's valuational stance must at least also conform to the characteristic positivist privileging of objectivity over subjective empirical descriptive detail. We will see that Robbins's practice in the *Essay* conforms to both of these positivist features in exemplary fashion. Finally, identifying Robbins's approach as positivistic will explain what I indicated in section 3 as the main puzzle that anyone who *does* approach the *Essay* from a philosophical point of view is bound to raise: why is it that we can be said to know with certainty that people order their preferences, yet cannot be said to know *anything* about comparative interpersonal preference intensities? This strange combination of theses seems to reflect a highly eccentric epistemology.

To show how Robbins derives the premises of his arguments from the history of problem solving in economics, some of that historical context must be introduced. In light of the standard emphasis on the role of classic British empiricism in the origins of marginalism and neoclassicism, it is useful to begin with a quotation from the marginalist founding figure who was closest to the main current of British philosophy, Jevons. “Far be it from me,” he cautions in the *Theory of Political Economy* (1871), “to say that we shall ever have the means of measuring directly the feelings of the human heart. A unit of pleasure or of pain is difficult even to conceive; but it is the amount of these feelings which is continually prompting us to buying and selling, borrowing and lending, laboring and resting, producing and consuming” (Jevons 1871, p. 13). Here is the economist who is typically credited or blamed for setting the new economics firmly on Benthamite hedonic utilitarianism – the doctrine against which Robbins was supposedly rebelling in rejecting interpersonal utility comparisons – already expressing doubt about not only the practical possibility of scaling preference intensities but about the very conceivability of such a scale.

Jevons also recognized, as had Bentham, that although ‘pleasure’ and ‘pain’ might literally be thought to denote hedonic sensations of the sort that Hume (among others) had confidently expected scientists to physically measure,¹² the ‘utility’ of which they indicate opposed valences can also be interpreted in what Bentham explicitly called a “wide and expansive” sense that encompasses all *possible* motivators. Thus, says Jevons, we can “call any motive which attracts us to a certain action pleasure and that which deters pain” (1871, p. 31). He then raises the problem that remains a favorite concern of skeptics about the psychological adequacy of neoclassically derived microeconomics: the tautology objection. If one adopts Bentham’s ‘wide and expansive sense’ of ‘utility’ then “it becomes impossible to deny that all actions are prompted by pleasure or by pain,” in which case invoking a person’s will to maximize their utility as an explanation for their actions looks empty. This prompts Jevons to draw a distinction that Robbins devotes much work in the early part of his *Essay* to undoing: dividing pleasures into “higher” and “lower” categories, in which the former include those that involve moralized or altruistic motivations, while the latter are restricted to the satisfaction of self-focused “material” sensations. The aspect of behavior concerned with such material and individually self-interested well-being is then taken by Jevons to be the proper domain of the economist. Marshall followed him in this view. Given how hard Robbins works in the *Essay* to *restore* the ‘wide and expansive’ interpretation of ‘utility’, this respect in which Robbins moved to *correct* a key foundation stone of methodological individualism in economics, his Austrian affinities notwithstanding, is worth noting for later consideration. We should also note that Robbins was not the first among the marginalists to break with Jevons and Marshall on the hedonic and materialist interpretation of utility. Wicksteed (1910) urged that the scope of motivations encompassed by the principle of marginal-utility maximization includes “all the heterogeneous impulses of desire or aversion which appeal to any individual, whether material or spiritual, personal or communal, present or future, actual or ideal” (*ibid*, p. 32).

¹² See Ross (1991).

This widening *and distancing from psychology* of the *general* concept of purposeful motivation is already a major step away from British empiricism in a Kantian direction. Kant and philosophers influenced by him strongly separate motivation in the logical sense from motivation in the psychological sense. It is the former, normative, form of motivation that they take to be relevant to the objectivity of science. Thus for such philosophers the domain of ‘action in general’ is an appropriate object of analysis distinct from a psychologist’s study of the pursuit of any particular end, such as material wealth. The former domain is a suitable candidate for the application of a body of systematic logical relations, whereas activity aimed at satisfaction of material wants is a rationally arbitrary concatenation yoked together by reference to practical human purposes. Economists since Robbins’s generation have generally – at least until the recent vogue for behavioral economics – adopted the perspective of the Kantians more than that of the empiricists in this regard. It is the basis of the point, now often made derisively by anti-economists and heterodox economists, that mainstream economics models ‘rational agents’ rather than ‘real people’.

In Robbins the distinction that a contemporary philosopher would draw by reference to ‘psychology’ versus ‘logic’ is made instead using unsystematic reference to two different senses of the word ‘psychology’. At one point, less confusingly, he resorts to German for a semantic resource and refers to psychology in the empiricist’s sense as “*Fach-Psychologie*”, of which he says that psychological hedonism is one “brand” (Robbins 1935, p. 85). He regrets that certain “English” economists “did in fact claim the authority of the doctrines of psychological hedonism as sanctions for their propositions” but happily “[t]his was not true of the Austrians. The Mengerian tables were constructed in terms which begged no psychological questions” (p. 84). The ‘English’ flirtation with *Fach-Psychologie* is then said to be merely that: “trimmings” and “*ex post facto* apologia” around a core of “logic” about which the Austrians were clear. We need hardly wonder where Robbins would stand on the currently incessant announcements of a ‘paradigm shift’ in economics driven by experiments showing that people are not ‘rational’, when he says:

The borderlands of Economics are the happy hunting-ground of minds averse to the effort of exact thought, and, in these ambiguous regions, in recent years, endless time has been devoted to attacks on the alleged psychological assumptions of Economic Science. Psychology, it is said, advances very rapidly. If, therefore, Economics rests upon particular psychological doctrines, there is no task more ready to hand than every five years or so to write sharp polemics showing that, since psychology has changed its fashion, Economics needs “re-writing from the foundations upwards”. As might be expected, the opportunity has not been neglected (pp. 83-84).

Thus when Robbins says, just a few pages later, that “if we are to do our job as economists, if we are to provide a sufficient explanation of matters which every definition of our subject matter necessarily covers, we must include psychological elements” (p. 89), we must understand ‘psychological elements’ in terms of something

other than *Fach-Psychologie*. Unfortunately, at this point Robbins is far from precise. His favourable reference to Max Weber here has directed most commentators (including Blaug and Caldwell) to attribute to him the idea that economics depends upon *verstehen*. This does not seem to me to be the best-supported interpretation. Robbins never uses the word, though he obviously would have been familiar with it. *Verstehen* is an epistemological method, and in this part of the *Essay* Robbins is not discussing epistemology. (He never discusses epistemology, except incidentally, in the *Essay*.) Chapter IV, where the attack on incorporation of *Fach-Psychologie* and the citation of Weber occur, concerns the scope and subject matter of economic generalizations. One might almost think Robbins's focus was ontological, since in the paragraph in question he refers to "links in the causal chain which are psychical, not physical" (p. 90). Yet if there is scant evidence for attributing commitment to the method of *Verstehen* to Robbins, there is none whatsoever for imagining that he hankers after metaphysical dualism. So how are we to understand his insistence on the need for a psychological element in economics, in the same book which a few chapters earlier advises economists to *ignore* psychology?

I suggest that knowledge of the neo-Kantian intellectual environment of the Austrians renders these remarks of Robbins easy to interpret. What he is talking about here is *logic* in the sense that a Kantian philosopher *prior to* Frege, Russell and the logical positivists would have used the term. We might speculate that Robbins does not *say* 'logic' here because he knows that logic now means something more exact and technical in nature, but isn't sure what to say instead. And well he might not know, because his contemporaries among the broadly positivistic philosophers were divided and unsure of themselves in this terrain also. Cassirer, for example, still used 'Psychologie' in contexts where, as an *avowed* neo-Kantian, he certainly did not mean *Fach-Psychologie*. Carnap and other logical positivists use a variety of German words marking subtle distinctions, which their English translators tend perforce to leave in German. 'Gedanke' is probably a good rendering of what Robbins, not being a philosopher, is getting at; but its normal translation as 'thought' would hardly have aided the clarity of his expression in English. He might have been well advised to use 'reason'. That this semantic association was near the front of his mind as he wrote these passages is suggested by the fact that his very next topic, that of the succeeding paragraph, is the extent to which the generalizations of economics depend upon "the assumption of completely rational conduct". Although by this he clearly intends reference to the *contemporary* economist's sense of 'rational', he refers to this as "a more general psychological assumption". This use of 'psychological' does *not* accord with contemporary usage. I suggest it is exactly the alternative sense of 'Psychologie' – alternative, that is, to *Fach-Psychologie* – intended in the Kantian tradition by Cassirer.

The decisive clue that the rationalization of this part of the *Essay* in terms of neo-Kantian philosophy is correct lies, I think, in Robbins's gloss of the 'psychological elements' with which the economist cannot dispense as "the conception of purposive conduct" (p. 90). He later says "it is arguable that if behaviour is not conceived of as purposive, then the conception of the means-end relationships which economics studies has no meaning. So if there were no purposive action, it could be argued that there were no economic

phenomena” (p. 93). In a footnote, he attributes this view to von Mises. What Robbins is mainly saying here, in plain modern terms, is that economics is essentially about ‘agency’. ‘Agency’ is necessarily *rational*, in precisely the sense of ‘purposive’. It does *not* necessarily involve sophisticated computational processing or self-consciousness – a negative point which the greatest of the Austrians, Hayek, regarded as especially important (see Hayek 1952). This is why, I have argued at length elsewhere (Ross 2005), economic theory has turned out to apply not only readily but surpassingly well to non-human animals,¹³ but not to rocks or planets. The former have conditions of better and worse flourishing that motivate *actions* (consciously or not) whereas no state of affairs is better than another for a rock. Thus we can construct utility functions for, and assign opportunity costs to the behaviours of, animals (and plants, and, in the new domain of neuroeconomics, brain cells), but not to rocks. It would thus have been more consistent with Robbins’s conception of the domain of economics, I argue, had he dropped the restriction in his definition of the scope of economics to ‘human’ behaviour. But the science of ethology did not yet exist by name when Robbins wrote, and the view of animals as deterministic robots was then still widespread.

Robbins is again at one with neo-Kantian philosophers in thinking that rational agency is constitutive of purposiveness, but does not require or entail “perfect constancy” – the sense of ‘rational’ that is typically at issue when contemporary critics of mainstream economics attack ‘rational economic man’. In this connection, Robbins gropes for a distinction we might now draw as that between an agent with perfectly rational expectations and an agent whose preferences merely do not cycle. The latter, he notes, is excluded in equilibrium lest arbitrage – money pumps, we would say here – be possible (p. 92). The former, he suggests, is an idealization with which the economist interested in limiting conditions begins, and which is relaxed in applications (p. 94).

Robbins’s treatment of these matters, which still cause difficulty in contemporary economic theory, cannot be said to be handled by him in a completely satisfactory way. He might have noted that both the requirement of acyclicity in equilibrium and the idealization of perfectly rational expectations are both properties of aggregate states of an economy (notwithstanding their later, gratuitous and partly ideologically inspired, reductive interpretation in the microfoundations literature). It seems to me that we should not regard them as psychological elements even in Cassirer’s sense of ‘psychological’; they are rather system-scale properties that need not decompose. But then of course we confront the general problem that relaxing these system-scale idealizations does not generally work in economics in the way in which, as Robbins suggests, it works in physics: in economics systems not *at* equilibrium don’t reliably tend to be *near* equilibrium in any well-defined sense.

Robbins might be accused – as I accuse him in my 2005 book (p. 94) – of waffling somewhat on the importance of acyclicity of preference in the individual agent. It is clear that he aims to minimize the extent to which economic logic is *committed* to this property. “In so far,” he says, “as the term rational is taken to mean merely ‘consistent’, then it is true that an assumption of this sort does enter into certain analytical

¹³ See, for evidence, the papers in Noë, van Hoof and Hammerstein (2001).

constructions” (p. 91). This can surely be said to be a considerable, if not whopping, understatement of the truth. Earlier in the *Essay* Robbins has enshrined the *fact* that preferences are normally ordered as one of the two foundational assumptions (along with the ubiquity of scarcity) for the empirical significance of economics. But acyclicity is part of what it means to say that preferences are ordered. Robbins makes clear that he thinks that inconsistent preferences are common but that consistent ones characterize *enough* of a typical person’s behaviour for economic analysis to gain purchase. This must be regarded as an incomplete account if the conception of economic analysis as consisting in deduction from basic postulates, which Robbins also asserts (pp. 75-76), is not to be threatened. ‘People have *many* consistently ordered preferences’ is not a postulate from which any precise claims about aggregate demand characteristics can be derived.

I argue in my 2005 book that there is a straightforward way out of this problem we can urge on Robbins’s behalf, which is to acknowledge (as is obvious anyway) that an entire person’s biography does not map onto one enduring economic agent. A person is instead a succession of economic agents, changing ‘agent identity’ when her tastes change. It may be objected that this implicates economics in *ad hoc* ontological principles, allowing agents to be multiplied without limit merely to make a certain style of formal analysis tractable. This objection is misplaced against a broadly positivist conception of economics, however. For the positivist, a scientific discipline is distinguished from others partly by identifying its conventional analytic categories. Psychological generalizations apply to a person named Lionel Robbins. Biological generalizations apply to another entity tracked by many of the same perceptible markers as Robbins, but as an instance of the species *H. sapiens* rather than as a person. Economic generalizations apply to an agent maximizing its present utility given a consistent preference ordering. We might casually say that the entity in question at time *t* is Lionel Robbins executing the purchase of some shares. But we should not assume that this entity will still be found hanging around the LSE at *t* plus five years’ time merely because the person Lionel Robbins will be found there. Agent identification is governed by pragmatic considerations, not metaphysics; thus we can be practical concerning how sensitive models must be to small variations in preference profiles. As long as the person we are modeling keeps preferring a secure but lower-return asset to a riskier but higher-returning one (for fixed prices) during the time he is wondering how to dispose of a few months’ income, that is generally good enough for the economist’s purposes. If he grows permanently fed up with brandy, which he formerly enjoyed, during the same period, we might simply decide to ignore this fact because we deem it unimportant to our analytical projects. We must *not* ignore it if we were interested in demand for beverages and we think Robbins is *psychologically* representative of other people or is part of a small cartel of buyers.

The conventional nature of the concept of agency does *not* undermine objective judgments about an agent’s expected consumption behaviour in response to a given price change once that agent has been identified (by identification of a utility function). As we saw in section 4, this is the basic point of the broadly positivist attitude to science. It might be thought that this attitude conflicts with the following passage from the *Essay*: “[I]t does not follow in the least that [economic] generalizations have a ‘merely formal’ status – that they are ‘scholastic’ deductions from arbitrarily established definitions ... It

is true that we deduce much from definitions. But it is not true that the definitions are arbitrary” (p. 105). That ‘economic agent’ has the conceptual structure it does as a function of its role in economic argument patterns and generalizations, and that we have degrees of freedom over where to apply it *in advance of* a specific problem context does not imply that facts do not determine where the agents are *after* a problem has been identified. Some models capture the economic influences actually at work in a circumstance and others do not, or do so less completely. As Robbins says “the *validity* of a particular theory is a matter of its logical derivation from the general assumptions which it makes. But its *applicability* to a given situation depends upon the extent to which its concepts actually reflect the forces operating in that situation” (pp. 116-117).

In saying that the bounds of agency are set by convention, the positivist does *not* say that agents ‘exist only in theory’ or ‘are merely theoretical entities’. Claims about the ontological status of a type, except where these are relative to a specific model, are metaphysical claims. These are just the sorts of claims, I am arguing, that Robbins should not be interpreted as making. There is no need for any economist to try to come to a view on such matters. Indeed we might go even further and suggest that she may make herself a less effective economist if she does so, because if the resulting philosophical opinion makes any difference to her practice at all, it must make her less flexible as a model builder.

In section 4 I partly identified broad positivism with structuralism. This applies well to Robbins, especially as an aid to understanding his polemic against historicism and institutionalism in chapters IV and V of the *Essay*. His problem with institutionalism is not that institutionalists mention institutions. Most contemporary development economists believe in a sort of Gresham’s law of institutions to the effect that institutions which encourage rent-seeking tend to drive out efficiency-promoting ones. I can find nothing in Robbins’s *Essay* that would bid us to reject such a generalization about institutions and their economic effects. The basis for Robbins’s critique of the historicism of his time is that these historicists deny themselves a stable basis for sorting causally relevant factors in a situation into sets of endogenous and exogenous factors. In consequence they produce historical descriptions instead of generalizations, and deprive themselves of bases for projection of their knowledge to new instances. The economist is directed by Robbins to search for the aspects of a situation which represent *structural* parameters featuring in economic “laws”, and which thereby allow deduction of further structural parameters. It is this perspective that licenses the following claim:

Economic laws describe inevitable implications. If the data they postulate are given, then the consequences they predict necessarily follow. In this sense they are on the same footing as other scientific laws, and as little capable of “suspension”. If, in a given situation, the facts are of a certain order, we are warranted in deducing with complete certainty that other facts which it enables us to describe are also present ... If the “given situation” conforms to a certain pattern, certain other features must also be present, for their presence is “deducible” from the pattern originally postulated (pp. 121-122).

This level of authority in analysis would not be remotely plausible, as Robbins explains, if the economist took herself to be responsible for predicting all or even most details of social states of affairs (including stereotypically ‘economic’ details such as exact prices or rates of output). The domain of the economist is a carefully circumscribed network of general structural relationships that apply whenever an interacting group of agents confront scarcity in the means to the ends over which they each have ordered preferences.

On the basis of the evidence mustered to this point, I claim that Robbins exemplifies the philosophical minimalism of the broad positivist. The basic implication of this so far as methodology is concerned is that metaphysical opinions are irrelevant to the conduct of economics. The economist travels about with a toolkit full of structural relations which, when applied to situations in which people (or, I add, other entities to which the concept of agency can be applied) confront scarcity with respect to the means to their ends, allow her to deduce further structural relations. General agreement among economists about the representation and nature of these structural relationships, and about conditions for their applicability, provide the basis for objectivity in economics. Economists of course routinely disagree amongst themselves. Such disagreement, if it is genuinely *economic* disagreement (rather than political or ethical disagreement in disguise) should always turn out to be disagreement over whether some instances of relevant structural parameters have been left out of the model of the situation at issue. Are there agents with influence who are not being accounted for? Do some agents have incentives that have not been noticed? Will different consumption patterns with respect to the scarce resources at issue influence patterns of scarcity in other resources that have not been foreseen? The contemporary economist would raise other potential questions, unavailable to Robbins, about distributions of asymmetrical information among the agents, and about constraints on the agents’ own modeling of the expectations of one another. Current economists might in addition have methodological disagreements – over, for example, whether one econometric test or another is more likely to reveal robust dependencies among variables. None of these disagreements are philosophical.

One major task remains. I said earlier that understanding Robbins’s broad positivist stance can help us make sense of his curious conjunction of views on what can and cannot be known about preferences. According to him, we can know that people order them but cannot know anything about their relative intensities. My reading of Robbins motivates searching for an *economic* justification of these views that does not appeal to one or another tendentious doctrine of general epistemology. To this search I turn in the next section.

6. Robbins on properties of preference relations

Let us begin by setting out the problem that Robbins raises for us in interpreting his views on the properties of preference sets.

In such history of economic thought as is influenced by standard Methodology, in which philosophy is assigned an influential role,¹⁴ the rejection of hedonistic sensationalism and the campaign for ordinalism in the interpretation of utility functions are typically run together as two aspects of the same view. We have seen how Robbins's broad positivism rationalizes anti-sensationalism: the latter is a thesis about specific psychological causation, a topic for *Fach-psychologie*, which the economist is advised not to incorporate into her models. However, Robbins seems then to claim that the possibility of economic analysis rests on one *general* psychological fact, namely, people's introspective awareness of processes of deliberative choice phenomena (pp. 75-76). Allowing that this *one* datum from psychology, however general, is *foundational* for economics while *all other* psychological phenomena are to be rigorously excluded as irrelevant seems *prima facie* surprising, and to call for philosophical argument which Robbins does not provide. Furthermore, the brief basis Robbins's *does* offer for exclusion of relative preference intensities appears to be of a behaviourist and operationalist character:

[S]uppose that we differed about the satisfaction derived by A from an income of £1,000, and the satisfaction derived by B from an income of twice that magnitude. Asking them would provide no solution. Supposing they differed. A might urge that he had more satisfaction than B at the margin. While B might urge that, on the contrary, he had more satisfaction than A. We do not need to be slavish behaviourists to realize that here is no scientific evidence. *There is no means of testing the magnitude of A's satisfaction as compared to B's.* If we tested the state of their blood-streams, that would be a test of blood, not satisfaction. Introspection does not enable A to know what is going on in B's mind, nor B to measure what is going on in A's. There is no way of comparing the satisfactions of different people (pp. 139-140).

We might well agree that in the imagined circumstances *asking* A and B to compare their preference intensities would be an ill-advised procedure. Their evidential circumstances are obviously asymmetrical between themselves and the other *if* introspection is admitted as a viable source of evidence in the first place, as Robbins seems to allow. Furthermore, A's and B's claims will be undermined by moral hazard if redistributive policies or considerations of social equality are operative (something Robbins hints is relevant here). But Robbins also rules out the possible relevance of evidence gathered independently of A's and B's avowals. If some property of the blood-stream were a reliable indicator of satisfaction, then why would a 'test of blood' *not* be an indirect 'test of satisfaction'? (We may indeed soon have the capacity to use neuroimaging probes to measure intrapersonal emotional states. Whether we will ever be able to justify comparisons of these measurements on an inter-personal scale is much less clear.)

Worse, earlier in the *Essay*, when defending the economist's appeal to everyone's experience of ordering their preferences, Robbins waves away operationalist scruples:

¹⁴ This is a minority of that history – a fact of which we should be glad.

In recent years ... partly as a result of the influence of Behaviourism, partly as a result of a desire to secure the maximum possible austerity in analytical exposition, there have arisen voices urging that this framework of subjectivity should be discarded. Scientific method, it is urged, demands that we should leave out of account anything which is incapable of direct observation. We may take account of demand as it shows itself in observable behaviour in the market. But beyond this we may not go. Valuation is a subjective process. We cannot *observe* valuation. It is therefore out of place in a scientific explanation. Our theoretical constructions must assume observable data ... It is an attitude which is very frequent among those economists who have come under the influence of Behaviourist psychology or who are terrified of attack from exponents of this queer cult.

At first sight this seems very plausible. The argument that we should do nothing that is not done in the physical sciences is very seductive. But it is doubtful whether it is really justified (p. 87).

No argument is given for the final assertion in this passage; Robbins merely asserts that “we do in fact *understand* terms such as choice, indifference, preference and the like in terms of inner experience” (pp. 87-88).

Once it is imagined that people phenomenally experience their preferences as ordered, it is then peculiar to suppose that they experience them as *merely* ordered. Hume, by contrasting example, clearly thought that we infer our preference orderings *from* our phenomenal awareness of differing levels of ‘vivacity’ in our passions for outcomes. In maintaining this doctrine, Hume surely speaks for folk psychology. In the second passage above Robbins seems to baldly appeal, over the heads of scientific psychologists, to folk psychology – while also appearing to defend a view of *economic* psychology, one we might dub ‘psychological’ ordinalism, which would perplex the folk and scientists alike.

I do not think there is any way to avoid finding Robbins guilty of at least careless writing in fostering this tension. If some degree of behaviourism is ‘slavish’ this implies that some lesser degree – the degree to which we can have recourse in justifying rejection of interpersonal comparisons – apparently is not. Yet in the earlier passage the most basic of behaviourist commitments, to the conviction that *introspection* is *never* a valid source of scientific evidence, is dismissed by implication as the dogma of a “queer cult”. No one can write two such things in the same book, without further explanation, and reasonably expect readers to clearly understand what he believes.

In showing how to best rationalize these apparently conflicting ideas about knowledge of preference structures – that is, to read the text in a way maximally charitable to Robbins’s consistency as a thinker – I will need to convict him of one further semantic slip. My interpretation will require us to assume that the inclusion of the word ‘inner’ in the statement “we do in fact *understand* terms such as choice, indifference, preference and the like in terms of inner experience” is a mistake. I will suggest that it is an explicable

mistake because recognizing it as such depends on subtle philosophical distinctions with which it would not be reasonable to expect Robbins to have been acquainted.

Before we return to this, we must locate the *economic* motivation for Robbins's ordinalism. The key comes in his earlier argument that, appearances notwithstanding, economists are not concerned with measuring quantities:

[T]he valuations which the price system expresses are not quantities at all. They are arrangements in a certain order. To assume that the scale of relative prices measures any quantity at all save quantities of money is quite unnecessary. Value is a relation, not a measurement (p. 56).

To this Robbins appends the following footnote:

Recognition of the ordinal nature of the valuations implied in price is fundamental. It is difficult to overstress its importance. With one slash of Occam's Razor, it extrudes for ever from economic analysis the last vestiges of psychological hedonism. The conception is implicit in Menger's use of the term *Bedeutung* in his statement of the Theory of Value, but the main credit for its explicit statement and subsequent elaboration is due to subsequent writers. See especially Cuhel, *Zur Lehre von den Bedürfnisse*, pp. 186-216; Pareto, *Manuel d'Economie Politique*, pp. 540-2; and Hicks and Allen, *Reconsideration of the Theory of Value* (*Economica*, 1934, pp. 51-76). In this important article it is shown how the most refined conceptions of the theory of value, complementarity, substitutability, etc., may be developed without recourse to the notion of a determinate utility function (p. 56).

This is among the most revealing passages in the book. It is a footnote in part because it is a citation of specific sources, but in part because Robbins is interrupting his discourse to the general reader and addressing fellow economists. Crediting Menger with the original insight, Robbins lists the recent highlights in what he sees as progress in separating economics from any interpretation in terms of folk psychology. As we will see shortly, the final sentence indicates Robbins's stand on the contested question of how to relate the traditional economic concept of utility to the newly developed analytic framework of indifference curves. Like Samuelson just a few years after him, Robbins would prefer to eliminate 'utility'.

I suggest that many of the main difficulties in interpretation of Robbins's treatment of preference relations in the *Essay* stem from the fact that he did not come up with (perhaps did not try to come up with) a way of explaining the recent technical developments in economics to his non-specialist readership. He therefore fell back on prosaic formulations from the Austrians, which, in his opinion, captured at least the implications, if not the primary motivations, of these technical developments. The main text in the passages we have been following resumes thus:

[I]t follows that the addition of prices or individual incomes to form social aggregates is an operation with a very limited meaning. As quantities of money expended, particular prices and particular incomes are capable of addition, and the total arrived at has a definite monetary significance. But as expressions of an order of preference, a relative scale, they are incapable of addition. Their aggregate has no meaning. They are only significant in relation to each other. Estimates of the social income may have a quite definite meaning for monetary theory. But beyond this they have only *conventional* significance (pp. 56-57).

This is entirely sufficient to explain why Robbins would reject the argument for income redistribution from diminishing marginal utility that is used as the motivation for his attack on interpersonal utility comparisons in Chapter VI. That argument was no doubt seen by him as a more dangerous application of a fallacy to a policy debate, and so a more important focus when addressing a popular audience. At the policy level, A's and B's incentives to exaggerate the downward slopes of their utility curves seem like persuasive and important considerations; and then a pinch of the kind of empiricism running so strongly in 1930s British intellectual currents¹⁵ might as well be thrown into the kindling. But the decisive basis for rejecting interpersonal comparisons had emerged from *economics*, not philosophy, and had been stated several chapters earlier for the cognoscenti to see. Aggregate demand buries information about opportunity costs. But opportunity cost, as Robbins explained back in Chapter I, is the very subject matter of economics. Now, aggregate demand, being expressed in monetary expenditure, is at least meaningful if monetary expenditure is what interests us. However, there is no analogous thing to be said about aggregate *utility*; that is no quantity at all. But 'utility' tends very naturally to be interpreted as denoting a quantity, and the first generation of British marginalists – but not Menger – had unfortunately encouraged this interpretation. So, Robbins says in his footnote, better to have done with 'utility' altogether.

Stated precisely, ordinalism of the sort Robbins defends is the thesis that objective functions – which, notwithstanding the apparent preference of Robbins and the clear later preference of Samuelson, continue to this day to be called 'utility functions' – are defined only by reference to properties preserved under monotonically increasing transformations. Diminishing marginal utility is not such a property – which is just why the mainstream economists of the 1930s welcomed its apparent elimination by Hicks and Allen in the 1934 paper to which Robbins refers in his footnote. Their replacement property, diminishing marginal substitutability, guarantees convexity of demand curves by supposing that agents will exchange less of any commodity x for another commodity y as their stock of x increases,¹⁶ but makes no reference to any sensationalistic or other basis in *Fach-Psychologie*.

¹⁵ This was mainly due to Russell. But Ayer had already seen his moment and was about to publish *Language, Truth and Logic*.

¹⁶ Hicks and Allen were aware that the principle as stated cannot be perfectly general; it typically fails in the case of complementary goods, such as gin and vermouth for martini drinkers. But if one assumes, as they did, that the basis of the principle is empirical rather than logical, there can be no decisive objection to simply treating this as a contingent set of limiting cases.

Mandler (1999, pp. 85-96) usefully distinguishes between the diminishing marginal utility principle as Jevons had understood it and the weaker property Mandler calls 'psychological concavity'. The former is the thesis that agents are introspectively aware of the *rates* at which the marginal utilities of particular commodities diminish on the margins, while psychological concavity denotes the property of *mere* awareness *that* marginal utility diminishes. Mandler defines a model of psychological concavity as follows: "At any point x , the set of psychologically accurate utility representations of preference on any line intersecting x is nonempty and consists of all of the concave utility representations of the agent's preferences on that line. In other words, agents experience diminishing marginal utility in all directions but no further nonordinal psychological reactions; on any line, any concave function representing the agent's preferences is psychologically accurate" (1999, p. 87). This specification allows assessment of the formal relationship between psychological concavity and diminishing marginal utility. Mandler shows that the set of utility-function transformations respecting Jevonsian diminishing marginal utility is a proper subset of those respecting psychological concavity, so the latter is a weaker assumption. However, psychological concavity is still not strictly ordinal.

In my 2005 book I argued (pp. 98-99) that Robbins *should have* embraced psychological concavity, which allows the possibility of *qualitative* awareness by agents of diminishing marginal utility, but in Chapter IV he goes unnecessarily far (by his own lights) and effectively endorses the purely behavioural principle of diminishing marginal substitutability. He says in a footnote that despite wholly approving of Hicks's and Allen's accomplishment, he still "prefer[s] the established terminology" – i.e., diminishing marginal utility – but gives no reason for this preference. Might he have thought that the difference between 'diminishing marginal substitutability' and 'diminishing marginal utility' was *merely* 'terminological'? That would be a very natural thing for a logical positivist to think. And I suggest that it is a quite natural thing for *Robbins* to think, because although the two principles have quite different implications for the relationship between economics and psychology – with psychological concavity having yet other implications – they make no difference for *economic* analysis if the data for that analysis are demand schedules rather than reported preferences.

That is exactly what the leading economic theorists of the 1930s and afterwards thought their data ought to be. They did not think this under the direct influence of philosophers, but their ambition for a separate and systematic economics is, as it happens, just what both Kantians and logical positivists would have regarded as equivalent to pursuit of 'objective' economics. Since broad positivists claim to derive their philosophical principles *from* science instead of deriving them independently and imposing them *on* science, a broad positivist should welcome evidence that the evolution of objective economics had been continuous, and endogenously driven. There is no shortage of such evidence. The achievement of Hicks and Allen that Robbins welcomes marked a milestone in a long development that began as early as 1881 with Edgeworth, was substantively completed by Samuelson in 1947, and was technically finished by Debreu in 1959.

My saying (following Mandler) that the process ‘began’ with Edgeworth alludes to his introduction of indifference curves to represent marginal analysis. It was, however, Fisher (1892) who made more than a mere representational device of the indifference curve. In particular, Fisher eliminated all assumptions about cardinal utility *beyond* the indifference judgment itself, which in his treatment becomes primitive. The indifference curve assumes comparison of *signs* of marginal utility (i.e., measurement of relative utility is unique up to monotone transformation), but presumes no measurements of any quantitative sums or totals of utilities. Fisher showed that relative price-levels at equilibria – points where agents could not improve their satisfaction by shifting their consumption – can be determined strictly by the gradients of indifference curves. Therefore, if we can derive families of indifference curves for all consumers and all consumption bundles, then we can do our economic analysis without having to know anything at all about cardinal magnitudes. Pareto (1909 / 1971) took this analysis one step further, arguing that since indifference curves can be constructed on the basis of sequences of observed choices by agents, we need not begin microeconomic analyses from *any* independent measurements of utility, if utility is interpreted as some sort of psychological aspect or coefficient.

As Mandler (1999) demonstrates, neither Fisher nor Pareto was consistently anti-realist in their attitude toward utility as a psychological force. Fisher’s specific analysis presupposed that the utility an agent derives from a particular commodity is often meaningfully separable from the utility she derives from other commodities; and if utilities can be separated then they are not yet formally redundant.¹⁷ Pareto makes the same move, and, at at least one point, offers a cardinalist interpretation of the meaning of indifference indices (Mandler 1999, p. 121). This is the point of entry for Hicks and Allen (1934).

Hicks and Allen begin from Pareto’s insight that preference-maps sufficient for prediction of consumption can be based on indifference curves that need not themselves be derived from utility functions. Since, following Fisher, indifference curves incorporate no presumption of cardinal comparability beyond primitive indifference judgments, any analytic use of utility functions built only out of the elements necessary and sufficient for the construction of indifference curves could be interpreted as harmless from the anti-cardinalist perspective. Furthermore, Hicks and Allen showed that convexity of the demand function does not require that diminishing marginal utility be given a psychological interpretation. It requires only the behavioural property of diminishing marginal rates of substitution, that is, that the amount of good x an agent will exchange for a marginal increment of her stock of good y increases with her stock of x . Note that if this is treated as an assumption instead of an empirical observation, it can be justified (to the extent that it *is* generally justified) by a purely economic (as opposed to psychological) argument: its violation is consistent with the possibility that an agent could maximize her welfare by consuming only one commodity.

¹⁷ Different psychological interpretations of this are possible; see Mandler (1999, 117-120).

From here the motivation for Samuelson's project of constructing revealed preference theory is immediate. I have argued elsewhere (Ross forthcoming) that this motivation was buttressed by the Keynesian encouragement – encouragement accepted by Hicks and Samuelson, along with Robbins – of treating aggregate demand schedules as basic data. I think it has been under-appreciated (though see Davis 2003) that in Samuelson the *agent* disappears altogether – the word occurs nowhere in his *Foundations* – and is replaced for modeling purposes by 'floating' preference orderings that can in principle be mapped onto any part of nature where (looking back to Robbins) 'purposiveness' might reside. The excess demand literature of the 1970s showed that economists were not in fact ready to let the attachment of agency to individuals slip away – otherwise that literature would not have been regarded as exposing a *problem* – but I contend that this concern derived from the relationship between demand theory and welfare economics, not from ontological anxieties about 'what agents are', or epistemological issues about how to discover their properties.

Of course we have now got ahead of Robbins. But it is necessary to follow the story a few steps past him to appreciate that he stands in the middle-to-late stages of a steady course of development internal to economics, which a broad positivist would gloss as aiming at increased disciplinary objectivity. I now believe I made a mistake in my 2005 treatment of this development. Because positivism among philosophers drifted from a Kantian to an empiricist emphasis between the 1930s and the 1950s, I took this direction to constitute progress *by broadly positivist lights*. I therefore read Robbins as a confused precursor to Samuelson. But this implicitly treats the historical drift of opinion among philosophers as though it carries persuasive weight of its own, which is *not* an attitude fully in conformity to broad positivism. Let me therefore now offer an amended rationalization of Robbins's stance on preference relations that is clearer about giving priority to science (that is, that does not privilege empiricism over neo-Kantianism for reasons that come from outside of economics).

If Robbins is read as being committed to the idea that the epistemological basis of demand and consumer theory must be introspection, then my earlier verdict would be unavoidable: we would have to say that Robbins let the philosophical tail wag the economic dog. I do not think the text requires this reading, however. *Given* the economic context, Robbins assumes that relative differences in preference intensities must allude to hypothesized quantities. Thus, in light of the direction in which he sees that progress in economic analysis had been going from Edgeworth through Hicks and Allen – and in light of where he thinks the Austrians had been all along – he must reject cardinal preference rankings as relevant to economics. In my book I wondered why Robbins did not, in light of his belief in ordinal psychological rankings, embrace Mandler's psychological concavity construct. I suggested that it simply did not occur to him. But this argument presupposed that he thought he had a good reason to hold on to introspective epistemological foundations if possible. Since I took it, and continue to take it, that there was in fact no such good reason, this yielded a Robbins who was muddled by comparison with Samuelson.

We are only led to think that Robbins was committed to introspective foundations for economics, however, if we fail to draw the distinctions between *Fach-Psychologie* and ‘psychology’ in Cassirer’s sense that were explained in section 5. Robbins encourages us to fail to draw this distinction by inserting the word ‘inner’ into his claim that “we do in fact *understand* terms such as choice, indifference, preference and the like in terms of inner experience.” Suppose, however, that this word were excised from the remark. Recall that several of Robbins’s other comments on the ‘psychology’ that is presupposed by economics suggest that this is merely the recognition by agents *that* they are agents – that is, ‘have purposes’ – or, better still in the context of Robbins, *have ends*. For Kantians and neo-Kantians this recognition, though a priori, is synthetic; although it is a condition for a certain sort of experience, it occurs *in* experience, not prior to it (as it does in the thought of, for example, Descartes). A logical positivist would have treated the modeling of economic behaviour in terms of agency as conventional at the formal level, but non-optional once ‘economic behaviour’ was defined in terms of means to the achievement of ends and in terms of scarcity of those means. The Austrians, who influenced Robbins, were in turn influenced by neo-Kantians, not by logical positivists. Notice, however, that it makes no difference to anything else Robbins says *about economics* in the *Essay* whether we attach a neo-Kantian or a logical empiricist interpretation to the ‘psychological elements’ as long as we do not reference these to *Fach-Psychologie*. Any *broadly* positivist framework will do; inside that conceptual space, the considerations that drive argumentation in the *Essay* all come from economics.

I am *not* here claiming that if we could bring Robbins back from the grave and ask him: “Did you really mean to put ‘inner’ in that sentence?” he would recognize a casual error and be moved to retract it. Robbins was not a philosopher. The subtle philosophical distinction between ‘experience’ and ‘inner experience’ is not even clear to many philosophers if their starting point is classic British empiricism. Ayer, for example, obscures the distinction, and it surely is altogether absent in Hutchison’s work on Methodology. It is likewise missing in Caldwell, who appreciates the importance of the Austrians for Robbins but misses the affinities between the philosophical background of the Austrians (i.e., Kantianism and neo-Kantianism) and logical positivism. My claims are instead as follows. Robbins interpreted economic theory as presupposing agency and as abjuring reference to hypothetical perceptions of relative psychological preference intensities. That was an accurate interpretation of the theory as he found it. In the *Essay* he tried to articulate these assumptions in non-technical terms. Here he borrowed repeatedly from the Austrians, who were more philosophically sophisticated than he was. Thus he lost some careful distinctions they might have made. But this matters only to a reader with a philosophical agenda, that is, one who wants to know whether Robbins was an empiricist or an a priorist. He was not really either. Like most scientists who prefer to leave philosophy to philosophers wherever possible, he took a broadly positivist attitude to his discipline. In the *Essay* he does not *justify* his discipline on the basis of philosophical assumptions; he instead *describes* what is distinctive about it, including what distinguishes it from psychology.

7. Conclusion

I started by offering a personal judgment that Robbins's *Essay* remains the gold standard for descriptions of the 'mainstream economic attitude'. In light of what I have subsequently argued, the reader will infer that this reflects endorsement of broad positivism. It may have occurred to some readers that the current wave of 'behavioural economics' challenges the form of separation between economics and psychology that I have said constitutes the basis for economic 'objectivity' in a broadly positivist framework. How then can my attitude toward the critical behaviourist economists be consistent? Must I not, in saying that Robbins is still the gold standard for identifying the basis of economic objectivity, be implicitly criticizing a recent body of economics on philosophical grounds? Yet isn't that exactly what the broadly positivist attitude forbids?

I intend no philosophical criticism of the scientific value of the products of behavioural economics. Of course this does not involve my swearing off all *economic* criticism of these products, but scientific criticism must always proceed one specific product at a time. And in any case I am thoroughly persuaded by many empirical findings of behavioural economists: framing effects are ubiquitous in human choice, people are dreadful statistical reasoners, and if they don't establish and rely on institutional safeguards they tend to meliorate instead of maximize, thus reversing their preferences and making themselves vulnerable to money pumps and other forms of manipulation by more sophisticated agents. What I join Robbins in rejecting is the often-heard claim that these *psychological* discoveries must prompt "re-writing of economics from the foundations up". That sort of claim does rest on any discovery *in economics*; it instead asserts some implicit (or occasionally explicit) *philosophy*. The philosophy it asserts is a very simple-minded realism to the effect that a science must take found objects of everyday experience – in this case, individual human organisms – as its *direct* subject matter.

Scientific disciplines study distinct structural networks of functional and causal relationships. Economics studies relationships between scarce resources and achievement of ends towards which agents put such resources. People implement agency, though no whole person is ever just one agent. Other kinds of entity, including firms and groups of neurons, also implement agency and face scarcity; so economic generalizations apply to them too. Human economic behaviour is heavily mediated by human psychology, and if we want to predict the activities of people we must therefore consider psychological factors (plus physical and chemical and evolutionary factors) in addition to economic ones. A precondition for consistently applying these factors in careful *conjunction* is keeping them *distinguished*. Economists should read Robbins's *Essay* to learn that there is a recurring need to be reminded of this point, which was often overlooked in the 1930s just as it is often overlooked today. They should not read the *Essay* for methodology – that is, to be told how to go about doing economics; for that there are textbooks.

Note finally that in saying that Robbins gives an accurate account of the distinctive economic attitude, one need not endorse the *content* of 1930s economics as against the economics of the early twenty-first century. I do not think that interpersonal comparisons of utility make no sense; I think we can infer from the equilibria of bargaining games a significant, *but non-psychological*, construct that is well described by that label (see

Binmore 1998). I thus think that there is such a field as welfare economics. In light of all we have learned over the past seven decades, it would indeed be ludicrous to hold up Robbins's *Essay* as the gold standard for (first order) *economics*. It is the gold standard for how to think *about* economics without going over into philosophy to do so.

References

Ayer, A.J. (1936). *Language, Truth and Logic*. London: Gollancz.

Binmore, K. (1998). *Game Theory and the Social Contract, Volume Two: Just Playing*. Cambridge, MA: MIT Press.

Blaug, M. (1980). *The Methodology of Economics*. Cambridge: Cambridge University Press.

Caldwell, B. (1982). *Beyond Positivism*. London: Unwin Hyman.

Carnap, R. (1928 / 1967). *Der logische Aufbau der Welt*. Berlin: Weltkreis. Translated by R. George as *The Logical Structure of the World*. Berkeley: University of California Press.

Carnap, R. (1934 / 1937). *Logische Syntax der Sprache*. Vienna: Springer. Translated by A. Smeaton as *The Logical Syntax of Language*. London: Kegan Paul.

Carnap, R. (1963). Intellectual autobiography. In P. Schilpp, ed., *The Philosophy of Rudolf Carnap*, pp. 1-84. LaSalle: Open Court.

Cartwright, N. (1999). *The Dappled World*. Cambridge: Cambridge University Press.

Cassirer, E. (1936 [1956]). *Determinism and Indeterminism in Modern Physics*. New Haven: Yale University Press.

Coleman, W. (2002). *Economics and its Enemies*. Houndmills, Basingstoke: Palgrave Macmillan.

Davis, J. (2003). *The Theory of the Individual in Economics*. London: Routledge.

Debreu, G. (1959). *Theory of Value*. New York: Wiley.

Dupré, J. (1993). *The Disorder of Things*. Cambridge, MA: Harvard University Press.

Edgeworth, F. (1881 / 1932). *Mathematical Psychics*. London: London School of Economics.

Fisher, I. (1892). *Mathematical Investigations in the Theory of Value and Price*. New Haven: Yale University Press.

- French, S. (1999). Models and mathematics in physics: The role of group theory. In J. Butterfield and C. Pagonis, eds., *From Physics to Philosophy*, pp. 187-207. Cambridge: Cambridge University Press.
- French, S. (2000). The reasonable effectiveness of mathematics: Partial structures and the application of group theory to physics. *Synthese* 125: 103-120.
- French, S. and Ladyman, J. (2003). Remodelling structural realism: Quantum physics and the metaphysics of structure. *Synthese* 136: 31-56.
- Friedman, M. (1999). *Reconsidering Logical Positivism*. Cambridge: Cambridge University Press.
- Friedman, M. (2001). *Dynamics of Reason*. Stanford: Stanford University Press.
- Giere, R. (1988). *Understanding Science*. Chicago: University of Chicago Press.
- Hayashi, F. (2000). *Econometrics*. Princeton: Princeton University Press.
- Hayek, F. (1952). *The Sensory Order*. Chicago: University of Chicago Press.
- Hicks, J., and Allen, R. (1934). A reconsideration of the theory of value. *Economica* 1: 52-76, 196-219.
- Hutchison, T. (1938). *The Significance and Basic Postulates of Economic Theory*. London: Macmillan.
- Jevons, W.S. (1871). *The Theory of Political Economy*. London: Macmillan.
- Ladyman, J. (1998). What is structural realism? *Studies in History and Philosophy of Science* 29: 409-424.
- Ladyman, J. (2002). Science, metaphysics and structural realism. *Philosophica* 67: 57-76.
- Ladyman, J., and Ross, D. (2007). *Every Thing Must Go*. Oxford: Oxford University Press.
- Maddy, P. (2007). *Second Philosophy*. Oxford: Oxford University Press.
- Mandler, M. (1999). *Dilemmas in Economic Theory*. Oxford: Oxford University Press.
- Noë, R., van Hoof, J., and Hammerstein, P., eds. (2001). *Economics in Nature*. Cambridge: Cambridge University Press.
- Pareto, V. (1909 / 1971). *Manual of Political Economy*. New York: Augustus Kelley.

- Quine, W.V.O (1951). Two dogmas of empiricism. *Philosophical Review* 60: 20-43.
- Robbins, L. (1935). *An Essay on the Nature and Significance of Economic Science*. 2nd edition. London: Macmillan.
- Robbins, L. (1998). *A History of Economic Thought*. Princeton: Princeton University Press.
- Ross, D. (1991). Hume, resemblance and the foundations of psychology. *History of Philosophy Quarterly* 8: 343-356.
- Ross, D. (2005). *Economic Theory and Cognitive Science: Microexplanation*. Cambridge, MA: MIT Press.
- Ross, D. (forthcoming). The economic agent: not human, but important. In U. Maki, ed., *Handbook of the Philosophy of Science Volume 13: Economics*. Amsterdam: Elsevier.
- Russell, B. (1914). *Our Knowledge of the External World as a Field for Scientific Method in Philosophy*. London: Allen and Unwin.
- Samuelson, P. (1947). *Foundations of Economic Analysis*. Cambridge, MA: Harvard University Press.
- van Fraassen, B. (1980). *The Scientific Image*. Oxford: Oxford University Press.
- Wicksteed, P. (1910). *The Common Sense of Political Economy*. London: Macmillan.