ON THE UNDERDETERMINATION OF SPACETIME THEORIES

Candidate Number: 14246
Degree Course: MSc in Philosophy & History of Science
Year: 2005-6

On the Underdetermination of Spacetime Theories

If not for the extended family of problems falling under the rubric 'underdetermination of *x* by *y*', the philosophy of science would inevitable by a more tranquil discipline. For underdetermination, in its many guises, is a philosophical commonplace: where *x* is any statement of future fact and *y* is presently available evidence, the underdetermination thesis is supported by Hume's original skeptical argument and Goodman's subsequent 'grue paradox' (Hume's recognition that constant conjunction/necessary connection are empirically equivalent alternatives raises further problems for accounts of causality and laws of nature); extending *x* to metaphysics and *y* to include any evidence whatsoever, there are Berkeley's doubts about an external world and Descarte's evil demon or the brain-in-a-vat; restricting *x* to scientific theories, the 'problem of empirically equivalent theories' (to use Magnus' (2005) term) and the Duhem-Quine thesis pose serious obstacles to scientific confirmation and a realist interpretation of science. Now Cartesian skepticism and the Problem of Induction have haunted the curious mind for centuries, casting the shadow of doubt on both scientific and nonscientific knowledge, however, only by the mid-twentieth century was *scientific* underdetermination (the underdetermination of theories by evidence) perceived as a distinct epistemic problem facing scientific inquiry. Begetting such antirealisms as Reichenbach's conventionalism and van Fraassen's constructive empiricism, the argument from scientific under-determination challenges our belief in the truth of a theory given the existence of empirically equivalent alternatives and the inability to test theories in isolation.

This essay is an attempt to dissect the scientific underdetermination project as it relates to Newtonian and relativistic spacetime physics. After considering general arguments by philosophers of science which purport to show that *all* theory choice is underdetermined (section 1), we narrow our lens on concrete examples from physics, probing the exact nature and extent of the underdetermination of spacetime theories (section 2) and whether this underdetermination is really something to worry about (sections 3 and 4). Examples of 'empirically equivalent' physical theories abound: Newtonian mechanics where the center of mass of the universe is at rest or in uniform motion relative to absolute space (TN(*0*) vs. TN(*v*)), versions of general relativity (GTR) with differing global topological features, standard GTR vs. relativistic theories of gravitation in flat spacetime vs. gauge theoretic formulations of gravity, Einstein's *Entwurf* theory and Nordström's Lorentz-covariant gravitation theory circa 1913, GTR and Brans-Dicke theory in the 1960s and early-1970s. Some theory pairings clearly seem more significant than others in establishing the force of scientific underdetermination. Retrospectively, the equivalence of *Entwurf* and Nordström's theory was of fleeting relevance, a historical footnote, as by November 1915, Einstein had already abandoned the problematic *Entwurf* and painstakingly discovered GTR's final field equations. The antirealist's only comfort here is a brief period of transient underdetermination.[1] TN(*0*) vs. TN(*v*) also appears a banal example as not only is Newtonian mechanics empirically inadequate, but the relativization of non-accelerated motions in Galilean spacetime (*i.e.*, the elimination of absolute space) is generally thought to reflect our deeper understanding of the kinematics of moving bodies.

Though alternative formulations of GTR incorporating all sorts of disparate spacetime structures (from scalar tensors to fiber bundles) suggest a more serious epistemic predicament, we contend that even these examples are insufficient to ground a virulent, non-transient underdetermination problem in spacetime physics. So long as physicists pursue a unified theory, alleged proofs and examples of non-transient underdetermination are wishful thinking – wishes that the competing features of currently empirically equivalent theories will carry into a complete theory of quantum gravity, or wishes that quantum gravity will itself be explained by a new host of empirically equivalent rivals. In particular, we argue in favor of the following theses:

- attempts by philosophers to establish the scientific underdetermination thesis from general arguments collapse into Cartesian skepticism (Stanford, 2001) or lead to only a transient epistemic predicament, failing to show that scientific underdetermination is a *distinct* problem that cannot be evaded through the further progress of science;
- in the absence of a quantum gravity theory, even the 'best' examples of empirically equivalent physical theories support only a transient underdetermination problem;
- actual examples from the history of relativistic physics point to a beneficial role for transient underdetermination in identifying research areas that may be fruitful for questions of scientific progress and motivating greater experimental precision.

1 | ‘Systems of the World’: Science or Fantasy

There are two roads to underdetermination. One often winds through the cosmos or quantum realm, leading from specific examples of underdetermined theory choice to general conclusions (and on this road we shall travel at length in subsequent sections of this essay). Along the other, the skeptic reflects on the nature and methodology of science and grounds the underdetermination thesis in broad philosophical considerations and a proof of the existence of empirically equivalent alternatives to any scientific theory. If two incompatible theories *T* and *T’* tell us the same things about the observable world, the skeptical philosopher reasons, why should we believe one theory over the other? Together with the Duhem-Quine thesis, this epistemic dilemma undermines our confidence in scientific knowledge.

In its general form, the problem of empirically equivalent theories runs as follows (Ladyman 2002: 174):

(i) for all *T*, there exist an infinite number of empirically equivalent but incompatible rival theories *T’*, *T’’*, *T’’’* …;

(ii) if *T* and *T’* are *empirically* equivalent, then *T* and *T’* are *evidentially* equivalent;

(iii) no evidence can support *T* more than *T’*, *T’’*, *T’’’* …, so theory choice is radically underdetermined.

Epistemic equivalence reduces to semantic equivalence – as the *T*s share the same observational consequences (and hence rise or fall together in their confrontation with the world), the competing theories are held on an epistemic par. As it is not altogether clear that every theory has an empirically equivalent alternative, let alone an infinite number of them, the underdetermination thesis also requires a proof, usually in the form of a construction algorithm, that such rival theories exist.

Both pillars on which this traditional argument for underdetermination stands – the existence of competing *T*s and the inference from empirical equivalence to epistemic parity – stand on tottery foundations. Putting algorithmic proofs aside for the moment, Laudan and Leplin (1991) have questioned the permanence of empirically equivalent theory pairings in the face of an ever-shifting science. The dependence of the very notion of 'empirical equivalence' on a sharp, static cleavage between observable/unobservable phenomena proves to be an Achilles heel as even if such a dichotomy exists, Laudan and Leplin contend that it changes over time (piggybacking on Maxwell's (1962) argument) and with it the empirical equivalence of scientific theories. Consequently, premise (i) is "both contextual and defeasible" (Laudan and Leplin 1991: 454) as judgments of empirical equivalence must be relativized to a particular state of science and technology. Laudan and Leplin do not deny that the problem of empirically equivalent theories may ultimately establish an unavoidable epistemic limitation to scientific inquiry but rather feel that so long as the rival *T*s are partial theories, scientific underdetermination is still an open concern, itself subject to empirical investigation. As science steadily marches on and the range of observable phenomena swells, the present empirical equivalence of theories justifies only a transient form of underdetermination.

Even if we grant that empirically equivalent rivals are ubiquitous to science and timeless, the move from empirical equivalence to evidential equivalence is also controversial. Of course, what discussion of underdetermination would be complete without mentioning the role of 'superempirical virtues' – such as simplicity, parsimony, explanatory power and fruitfulness – in theory choice. Incorporating such criteria in scientific standards stretches a theory's evidential support beyond mere conformity to experience, shattering the link between empirical and evidential equivalence (*i.e.*, premise (ii) fails). But now the problem of empirically equivalent theories is replaced with fresh difficulties as the superempirically-bent philosopher must still show that these virtues are signs of truth and not just pragmatic and outline how to deal with circumstances where superempirical considerations pull in opposing directions. Some realist accounts go at least some way in addressing this first concern by considering how scientific theories are confirmed: on Glymour's (1977) 'bootstrapping' approach to confirmation, a theory can be better tested than an empirically equivalent rival by the same body of evidence as the manner in which evidence bears on the individual hypotheses that compose a theoretical system is crucial for confirmation (*e.g.*, a theory is better confirmed if evidence tests a small number of central hypotheses repeatedly rather than subjecting a larger group of hypotheses to fewer tests); Friedman (1983) feels we should prefer theories with greater 'unifying power' as such theories are better tested given their ability to meet the world along multiple direct and indirect routes. The common theme here is that taking empirical adequacy as the sole criterion of a theory's worth ignores salient features of the scientific confirmation process that may affect our confidence in the truth of theory.

Sagging under the weight of these objections to the evidential equivalence of *T*, *T'*, *T''*…, the problem of empirically equivalent theories, applied to partial science, is not a convincing argument for scientific underdetermination. The skeptic, however, has other tricks up his sleeve as he can still (1) invoke the Duhem-Quine thesis (Duhem, 1904-5; Quine, 1951), or (2) consider empirically equivalent total sciences (Hoefer and Rosenberg, 1994). According to the Duhem-Quine thesis, only entire theoretical systems (groups of theories and auxiliary hypotheses for Duhem; the whole of science for Quine) are empirically tested so we cannot speak of the observational consequences of a theory in isolation. Again we encounter an epistemic dilemma as faced with an empirical inconsistency, which part of science do we revise or discard: theory? auxiliaries and boundary conditions? perhaps even vast tracts of mathematics or classical logic itself? Philosophers tend to agree with Duhem and Quine that the 'unit of empirical significance' (in Quine's terms) is often larger than a single theoretical statement, requiring *le bon sens* (Duhem) or pragmatic considerations (Quine) in making scientific judgments. Nonetheless, few accept full-blown Quinean holism as "no science would be possible if everything were up for grabs at the same time" (Magnus 2005: 841) and do we really test hypotheses about monkey mating habits by observing planetary orbits? The realist might also add that the Duhem-Quine thesis is, in principle, only a problem at the local level, another source of transient underdetermination, as given an empirically adequate total science whose parts combine to provide a complete, accurate picture of the world, the discrimination between theory, auxiliaries, boundary conditions, etc. is no longer necessary.

Ironically, elevating the discussion to the level of total science or ‘systems of the world’ (using Quine’s term again) only avoids the troubling epistemic implications of the Duhem-Quine thesis at the cost of resuscitating the problem of empirically equivalent theories (*i.e.*, there is a tradeoff between (1) and (2)). Taking *T*, *T’*, *T’’* … to be competing systems of the world, the slings and arrows of Laudan and Leplin now miss their mark as their arguments assume empirically equivalent alternatives are partial theories (Hoefer and Rosenberg 1994: 596-601). Moreover, as systems of the world account for *everything*, it also seems harder to justify an epistemic role for superempirical virtues in the choice between total sciences. Consider a simple, elegant total science *T* and its chaotic ‘dappled’ twin *T’*.[2] Whereas we might have attempted to vindicate our preference for *T* over *T’* with some type of historical-inductive argument – *e.g.*, the historical record shows that ‘super-empirically superior’ theories are typically empirically successful while more complex, vulgar theories eventually prove inadequate[3] – such an argument no longer works as the impish *T’* does everything *T* can possibly do. As already mentioned, other evidential reasons may still exist for choosing *T* over *T’* but until someone spells these out in a convincing fashion, underdetermination does appear to be an acute problem at the global level:

> If there can be two genuine, incompatible theories that are equally well confirmed by all possible evidence, we face an epistemic dilemma that we cannot hope to evade through the further progress of science, or extension of theory to cover new domains. (Hoefer and Rosenberg 1994: 592)

Recognizing then that the key to underdetermination lies in rival systems of the world, the scientific underdetermination thesis hinges on whether such competing total sciences exist. Though it is far from obvious, let us be generous and concede that science

T will one day account for all of nature's quirks and capacities and ask whether an empirically equivalent rival science *T'* is viable. The easy response here is to wait and see but we are then left with only a promissory note for *T'*, a hypothetical epistemic predicament which is wholly contingent on future science. Alternatively, the literature provides a wealth of interesting, often wonderfully ad-hoc algorithms for generating empirical equivalents: a trivial instrumentalist version takes $T_1' = O \cap O' \cap O'' \cap \ldots$ where O, O', O'', …, are the observational consequences of *T*; Kukla (in Ladyman 2002: 179) suggests T_2' = '*T* whenever somebody is observing something or ~*T* when there is no observation'; Stanford (2001: S2) presents the sci-fi hypothesis 'the Makers' (or 'the Matrix' in popular culture) where T_3' asserts that nature is really an elaborate computer simulation. Some authors feel these fanciful constructions are real theories and should be taken seriously (Ladyman 2002: 179-80; Kukla mentioned in Stanford 2001: S2-3), others reject them outright (Laudan and Leplin 1991: 456-7; Hoefer and Rosenberg 1994: 603-4) while Stanford all but avoids this question altogether:

> *[W]hether or not* such farfetched scenarios are real theories they amount to no more than a salient presentation of the possibility of radical or Cartesian skepticism. While many contemporary philosophers are inclined to grant the irrefutability of such skepticism, underdetermination was supposed to represent a distinct and important problem, arising perspicuously in the context of scientific theorizing about inaccessible domains of nature and troubling *even those who never hoped to defend their scientific beliefs to the truly radical skeptic*. Thus, if Cartesian fantasies are the only reasons we can give for taking underdetermination seriously, then there simply *is no* distinctive problem of scientific underdetermination to worry about, for the worry *just is* the familiar specter of radical skepticism. (Stanford 2001: S3)

As we took pains to emphasize in the opening passage of this essay, the schema 'underdetermination of *x* by *y*' can be filled in by various *x* and *y*, resulting in a plurality of epistemic concerns. Sensitive to these distinctions and the skeptic's tendency to confuse them, Stanford insightfully argues that insofar as the underdetermination of theory is perceived as separate from Cartesian skepticism (or the Problem of Induction for that matter), the scientific underdetermination thesis cannot be secured through Cartesian considerations (or Humean/Goodmanesque arguments) – there is a category mistake. The punch-line is that since the only non-trivial proofs of the existence of rival systems of the world *T'* invoke Cartesian fantasies, the skeptic, for all his clever algorithms, has yet to show that scientific underdetermination is a distinct, non-transient problem for global science.

The high road to underdetermination by way of general philosophy, cobbled with evidential considerations and Cartesian algorithms, is thus a road of dead ends and open promises. At best, the skeptical philosopher salvages some form of transient underdetermination if he takes the Duhem-Quine fork or ends his travels at the level of partial science. Turning our backs then on these broad philosophical arguments, let us take the other route and look to specific examples of underdetermined spacetime theories.

2 | Kinematic Shifts, Global Cosmologies & Curved Space, Oh My!

The center of mass of the universe shuttles along in uniform motion relative to absolute space. Light rays bend in the vicinity of the Earth's gravitational field, following their natural course through the curved fabric of spacetime. An apple falls from a tree to the ground in flat spacetime, acted on by 'universal forces' whose effects, though negligible to the falling apple, ensure that cosmic trajectories agree with the predictions of general relativity theory. The reader familiar with the philosophy of science literature on the underdetermination of spacetime theories should recognize these examples immediately: the first is the celebrated TN(v) case from Newtonian mechanics, the second is just standard general relativity while the third derives from Reichenbach's (1958) famous strategy for generating empirically equivalent rivals to GTR, or more generally, for demonstrating the conventionality of spacetime geometry. In this section, we investigate such examples of empirical equivalence in physics and ask whether they are sufficient to ground a vicious, non-transient underdetermination problem for science.

Let us begin with TN(*v*), where TN = Newton's laws of motion + law of gravitation and *v* is the hypothesis that the center of mass of the universe has constant absolute velocity *v* (so according to TN(*0*), the center of mass is at absolute rest). The TN(*v*) case goes all the way back to the 1715-6 Leibniz-Clarke debate and Leibniz's original kinematic shift argument against absolute space. Instantaneously change the velocity of everything in the universe by a fixed amount relative to absolute space, Leibniz challenged, and there is no observable difference so by the Principle of Identity of Indiscernables, absolute space is only a fiction. In modern terms, Leibniz recognized that the existence of absolute space and Galilean invariance of TN imply an infinite number of incompatible theories (*e.g.*, TN(*v**) and TN(*0*)) underdetermined by all available evidence. By eliminating absolute space, we evade this epistemic dilemma as the absolute motions *v* are now meaningless.

Many philosophers regard TN(*v*) as a trivial harmless example of scientific underdetermination. Hoefer and Rosenberg (1994: 599) concede that TN(*0*) and TN(*v**) are empirically equivalent theories but the inability for their common core TN to account for the deflection of light in a gravitational field or anomalies in the perihelion advance of

Mercury "makes their empirical equivalence an uninteresting issue for working scientists." Friedman, like Leibniz, argues that the notion of absolute space, and with it the underdetermination of TN(*0*) and TN(*v**), should be rejected. But while Leibniz appealed to observational indistinguishability, Friedman (1983: 112, 248-9) condemns absolute space for its 'theoretical dispensability' and lack of unifying power in the context of Newtonian gravitation theory. Stanford, by contrast, does not even take TN(*0*) and TN(*v**) to be distinct competing theories but rather a single theory conjoined with alternative factual claims about the world. Our problem is this: we have TN + *v* where TN itself implies that the factual claim *v* is unobservable so the realist must reconcile their disbelief in the claim *v* with the unfortunate circumstance that the conjunction TN + *v* is immune to empirical refutation if TN is indeed true. Again, Stanford feels scientific underdetermination reduces to 'a familiar philosophical chestnut', this time the problem of irrelevant conjunction in confirmation theory: if evidence E confirms theory T, then E also confirms $T + X$ where X is anything we please that does not interfere with the connection between T and E. Stanford (2001: S5) concludes: "Like Cartesian skepticism, this problem is philosophically serious (indeed, it requires some solution), but it cannot be our only reason for taking underdetermination seriously without simply collapsing the latter problem into the former."

The responses to the TN(*v*) case should seem familiar. Heeding our discussion in section 1, three basic strategies for avoiding the strong non-transient underdetermination thesis re-emerge: (S1) show that underdetermined theory choice is not a distinct problem for scientific inquiry by collapsing scientific underdetermination to other philosophical riddles (Stanford's TN + *v* and the problem of irrelevant conjunction); (S2) repudiate the link from empirical equivalence to evidential equivalence (this is implicit in Friedman's critique where TN(*v*) is evidentially inferior to Newtonian theory without absolute space given TN(*v*)'s superfluous structure); or (S3) argue that examples of underdetermined partial theories cannot secure the underdetermination thesis as scientific under-determination is only a potential problem for total science (Hoefer and Rosenberg's quick dismissal of TN(*v*) as TN is an empirically inadequate partial theory in light of general relativity's predictive successes). What should already be clear then from this Newtonian example is that the attempt to ground scientific underdetermination in particular cases of empirically equivalent physical theories is not isolated from broad philosophical considerations. Though we earlier laid out two separate roads to underdetermination – one through general philosophy, one from specific examples – these routes now merge as arguments from the general debate over scientific underdetermination carry over into our current localized discussion of spacetime physics.

In other alleged examples of physical theories where underdetermination rears its ugly head, from different models of GTR to alternative relativistic theories of gravitation, (S1)-(S3) may again show underdetermination to be a mere will-o'-the-wisp. Consider the underdetermination of global cosmology due to the constraints of relativity theory. As evidence available to an observer consists solely of events lying within the observer's past light cone, "even idealized observers who live forever may be unable to empirically distinguish hypotheses about global topological features of some of the cosmological models allowed by Einstein's field equations for gravitation." (Earman 1993: 31) The idea here is that underdetermination results from a natural restriction on our acquisition of empirical evidence through causal interactions with the world. Consequently, hypotheses about spacetime events and structures falling outside this domain of observability are underdetermined by evidence available to even 'idealized observers'. Now one possible response here is to follow (S1), interpret our subject matter as GTR + factual claims about global topology, take GTR as a local field theory and once more let the problem of irrelevant conjunction play the scapegoat (Stanford 2001: S5). Alternatively, we can pursue (S2) and claim some evidential basis for preferring a particular global cosmology, such as the globally closed Einstein cosmos, over other solutions to GTR's field equations. A final escape route, we can also adopt (S3) and argue that the matter ultimately rests on what a complete theory of quantum gravity tells us about the far reaches of the universe.

Reichenbach's (1958) conventionality of geometry thesis similarly fails to show scientific underdetermination is a distinct problem in spacetime physics, fairing no better with respect to (S1) and (S2) (we discuss (S3) later). Recall that for Reichenbach, empirical observations inherently combine both geometric and physical influences so geometry maintains an independence from empirical fact and can thus be chosen at whim, providing compensating adjustments are made in physical theory (*i.e.*, we account for the effects of unobservable 'universal forces'):

> Theorem θ: Given a geometry G' to which the measuring instruments conform, we can imagine a universal force F which affects the instruments in such a way that the actual geometry is an arbitrary geometry G, while the observed deviation from G is due to a universal deformation of the measuring instruments. (Reichenbach: 33)

The crucial point, for our present purposes, is that since universal forces cannot be observed, the theoretical combinations ‘G'’ and ‘$G + F$’ are underdetermined by all available evidence. If G'' is the curved spacetime of GTR and G is Minkowskian geometry, we may nevertheless favor ‘$G + F$’ for the mathematical simplicity and familiarity of flat spacetime, or alternatively, set $F = 0$ by the principle of parsimony as F lacks unifying power (Friedman: 300) but we are still left with the old problem of demonstrating that such evidential considerations are epistemic and not solely pragmatic (so (S2) is a live option but needs to be worked out). A stronger argument against Reichenbach's case for underdetermination from the empirical equivalence of ‘G'’ and ‘$G + F$’, however, lies in Stanford's (2001: S6) recognition that "in the end, [universal forces] are no better than ‘phantom effects’ and we are left with just another skeptical fantasy." Akin to the colorful algorithms for generating empirically equivalent systems of the world, universal forces underwrite another instance of (S1) as we may just as well adopt any arbitrary geometry and imagine that the perpetual sneezing of a sick demon alters the trajectories of bodies in spacetime so that they agree with observation. Reichenbach has only given us another reason to be Cartesian skeptics; there is nothing new here.

In the end then, the skeptic's motley crew of kinematic shifts, global cosmologies and universal forces prove insufficient to ground a vicious, non-transient underdetermination problem for theoretical physics. Nonetheless, the skeptic need not surrender. For by restricting Reichenbach's analysis, allowing geometry to vary but turning a blind eye to universal forces (besides gravitation), we arrive, without further ado, at arguably the most significant case of alleged scientific underdetermination in spacetime physics, a single question: is spacetime curved? Answering the question in the negative, one describes gravitation in terms of a flat manifold where gravitational force acts on rigid bodies, knocking them off geodesics of the manifold. Otherwise, one 'geometrizes away' (in Friedman's terms) gravitational force by incorporating the gravitational field into the now-variably curved affine structure of spacetime, associating gravity with the effect of spacetime curvature and allowing gravitational trajectories to follow geodesics like well-behaved schoolchildren. As the first approach is found in Newtonian gravitation theory while the latter is realized in general relativity, the historical record is generally taken to support the conclusion that spacetime is indeed curved.

This conclusion is hardly certain. With subsequent formulation of a curved-space version of Newtonian gravitation theory (Friedman 1983: §III.4; see also Glymour 1977: 238-42) and flat-space relativistic theories of gravitation (Gupta, 1952, 1957; Thirring, 1959, 1961; as discussed in Lyre and Eynck, 2001), underdetermination threatens the affine geometry of spacetime in *both* relativistic and non-relativistic contexts. To simplify slightly, flat relativistic gravitation theories split GTR's metric field $g_{\mu\upsilon}$ into a fixed Minkowskian background metric $\eta_{\mu\upsilon}$ and a dynamical gravitational field tensor $F_{\mu\upsilon}$ dependent on the mass distribution of the universe: $g'_{\mu\upsilon} = \eta_{\mu\upsilon} + \kappa F_{\mu\upsilon}$ (Lyre and Eynck 2001: 4). As in Newton's theory, gravity is conceived as a force, causing the trajectories of freely-falling bodies to deviate from the Minkowskian geodesics. But by adjusting the coupling constant κ, the observational predictions of flat-space relativity converge to those of GTR so the choice between $g_{\mu\upsilon}$ and $g'_{\mu\upsilon}$ is underdetermined by all available evidence. Note that flat gravitation theories may still differ from GTR in their global cosmological features as flat theories, for example, preclude a globally closed spacetime but as we have already mentioned, the underdetermination of global cosmologies is not really something to worry about.

Alternatively, the underdetermination of the affine structure of spacetime can be motivated entirely within Newtonian gravitation theory. Let TN(φ) represent a generally covariant version of Newton's original theory featuring a flat affine connection D and gravitational potential φ. Friedman (1983: §III.4) argues that since we cannot locally distinguish an acceleration from a gravitational potential, an empirically equivalent rival to TN(φ), call it TN(ψ), can be constructed with a new (potentially true) gravitational potential ψ and flat connection D' which exactly offsets differences in the effects of φ and ψ on spacetime trajectories. When $\psi = 0$, we have Einstein's famous elevator dangling in space: trapped inside and feeling pulled towards the floor, one cannot discern (ignoring tidal forces) whether the elevator is stationary in a gravitational field (TN(φ)) or accelerating upward through gravity-free space (TN(ψ)). Analogous to the earlier TN(v) example which called, at least to Leibniz, for the elimination of absolute space, we can avoid the underdetermination of TN(φ) and TN(ψ) by changing our ontological commitments, replacing D and φ (or equivalently D' and ψ) with a dynamic affine connection and thus allowing the spacetime of Newtonian physics to curve (see Friedman 1983: §III.4 for a nice covariant presentation of the 'Curved Newton' theory). But now, the skeptic may reasonably counter, we have simply replaced one under-determination predicament with another as the standard ‘Flat Newton’ and new ‘Curved Newton’ theories are also observationally indistinguishable.

At first glance, the underdetermination of affine geometry in both relativistic and non-relativistic theories of gravitation appears only a special case of Reichenbach’s ‘G'’ and ‘$G + F$’. In the relativistic realm, this immediately follows from our discussion of flat relativistic gravitation theories if we take $G' = g_{\mu\upsilon}$, $G = \eta_{\mu\upsilon}$ and $F = F_{\mu\upsilon}$. Glymour (1977: 245) also appreciates this point with respect to the rival Flat/Curved Newtonian theories, writing: “it is clear that [Flat Newton] is just [Curved Newton] *with a universal force*. The universal force in that case is $-g^{ir}\varphi_{ir}$, *i.e.*, just the gravitational force that enters the equation of motion of [Flat Newton].” Accordingly, the malignancy of underdetermined affine structure depends on whether gravitation can be taken as a universal force (or ‘universal field’ in the relativistic case) in Reichenbach’s sense. If so, our previous objection to Reichenbach’s conventionalist thesis still goes through and our present epistemic predicament may be shrugged off as another instance of Cartesian skepticism. If not, (S1) is no longer a viable option and we have genuine examples of non-skeptical and distinct empirically equivalent theories on our hands. So is gravitation, conceived as *gravitational force*, a ‘universal force’? I think not. Though gravitational force does meet Reichenbach’s *prima facie* criteria for a universal force – it affects all materials in the same way and no insulating walls are possible – the historical development of spacetime physics makes it painstakingly clear that treating gravity as a fundamental force is not a mere skeptical fantasy invoked to exploit our sensory limitations. Consider that for over two hundred years after the *Principia*, Newton’s ideas on gravity were universally accepted in the scientific community until Einstein's theory of general relativity revolutionized our conception of gravity as the effect of spacetime curvature. Sure, a GTR realist might say that Newton was wrong, the beliefs of the scientific community are irrelevant and we now recognize gravitational force to be the fantasy it has always been. But then again, Hume's ghost reminds us that these same comments may likewise apply to the rather far-fetched notion of 'curved space' in the future.

Abandoning (S1) then and our hopes of exposing our current epistemic dilemma as Descarte's demon dressed in new clothes, we might still follow (S2) and instead of crying 'universal force', ask whether evidential considerations beyond empirical adequacy justify our choice between a flat or curved spacetime geometry. There is the usual panoply of responses: Friedman (1983: 96-7, 121-2) favors Curved Newton given the indistinguishability of inertial frames from arbitrary Galilean frames and, correspondingly, the theoretical indeterminacy of φ and D in the field equations of Flat Newton; the flat connection D is also rejected for its lack of unifying power, both within gravitation theory and in the unification of gravitation theory with electrodynamics (Friedman 1983: 290-1); on Glymour's account (1977: 238-44), Curved Newton is better tested than Flat Newton as the indeterminacy of φ and D implies that many of the equations of the standard Newtonian theory cannot be instantiated so, by the principle of parsimony, Flat Newton should be discarded (the same argument applies for the choice between standard GTR and flat-space alternatives as the indeterminacy of $F_{\mu\upsilon}$ and $\eta_{\mu\upsilon}$ implies that field equations including these objects do not confront the world). But perhaps the strongest evidence supporting spacetime curvature is the Principle of

Equivalence which states that inertial and gravitational mass are equal ($m_i = m_g$ is actually the weakest of three versions of the principle; see Friedman 1983: §V.4 for details). Suggesting that gravity is not associated with rigid bodies but with spacetime itself and allowing us to geometrize away gravitational force in a non-flat affine structure, the equivalence of inertial and gravitational mass is an extraordinary physical fact, confirmed by Eötvös' famous 1889/1908 torsion balance experiments and considered by Einstein "the most fundamental property of gravitation." (Einstein (1933) in Norton 1989a: 8) While the principle also holds in flat-space relativistic gravitation theories, the variable curvature of spacetime is nonetheless an intuitive, compelling explanation of why all bodies fall the same way in a gravitational field.

It must be emphasized (yet again) that given the current state of the philosophy of science, the strategy (S2), though a live option for the realist, still does not dispel the scientific underdetermination threat. Curved space may be an 'intuitive, compelling explanation' of $m_i = m_g$ and we might prefer simple, parsimonious, unifying theories to vulgar ones but we still require a convincing account of how the same body of evidence can confer differing degrees of epistemic warrant to empirically equivalent rivals. Though some progress has been made by Friedman and Glymour (among others) in this regard, realists have not yet done the necessary work.

So with the realist backed into a wall, defenseless it seems against the underdetermination of affine geometry, the time has finally come for (S3), having laid low in our discussion of Reichenbach's conventionalism and affine structure, to burst forth with all its explosive impact. The argument from (S3) is simple, familiar and utterly devastating to the skeptic's case for underdetermination: as the underdetermination of theories is only an acute problem at the global level, examples of empirically equivalent partial theories cannot secure the strong underdetermination thesis. Recalling Hoefer and Rosenberg's response to the TN(*v*) example, we need not worry about our choice between Flat Newton and Curved Newton as both empirically inadequate theories (which cannot account for relativistic phenomena such as the deflection of light) were eventually replaced by Einstein's GTR anyway. Analogously, the underdetermination of affine structure among relativistic theories of gravitation (which cannot satisfactorily account for the eccentricities of the quantum world) is only a temporary concern as the progress of theoretical physics toward a unified theory may ultimately prove this epistemic dilemma irrelevant as well. The skeptic can still respond that if GTR and its empirically equivalent rivals are simply embedded in quantum gravity theory (*i.e.*, in unifying gravity theory with quantum mechanics, we add on to our physical knowledge with little disruption to the component gravity/quantum theories), the underdetermination of affine geometry will still be retained in a fresh collection of alternative flat/curved theories of quantum gravity. But this is only wishful thinking, especially as recent quantum gravity programs, such as superstring physics and gauge theories of gravity (see next section), often significantly modify the geometrical structure of spacetime. More generally then, specific examples of empirically equivalent physical theories, excepting the possible existence of rival complete theories of quantum gravity, support (at most) only transient under-determination problems in spacetime physics.

The reader with short patience may justifiably (and daresay angrily) ask at this point why we have left the argument from (S3) so late, for to be sure, we already learned that non-transient scientific underdetermination is a global problem from the general philosophical debate in section 1. Our travels among concrete cases of under-determined partial theories in theoretical physics thus smack of unnecessary arm-waving: we have been climbing up the craggy side of a hill, attempting to dissolve under-determination threats with Cartesian and evidential considerations, only to find a smooth path on the other side – and even worse, a path that our friends Hoefer and Rosenberg told us was there all along. At the beginning, could we not just have said that physics is not yet complete and left things at that? But I think the reader has done well to be patient here. For one thing, if scientific underdetermination (even of the transient kind) is rampant throughout physics, we might have good reason to suspect that a complete system of the world would indeed have empirically equivalent rivals. Nevertheless, we have seen that by invoking (S1), many alleged examples of underdetermined physical theories fade away, leaving only the underdetermination of affine geometry in Newtonian and relativistic physics as a genuine non-skeptical case. On the other hand, our reflections on (S2) have not been in vain either. It is all good and well that the continual progress of theoretical physics may resolve scientific underdetermination problems but this does not exactly help the physicist stuck in an epistemic predicament. The physicist, however, does have the principle of parsimony, Friedman's 'unifying power' and other evidential criteria beyond empirical adequacy at her disposal as methodological guidelines for theory choice. In this new role, superempirical virtues need not be guides to truth. So long as they lead to theories of increased empirical strength, evidential considerations from (S2) may still help the physicist make informed choices between rival empirically equivalent spacetime theories.[4]

We conclude this section as we concluded the previous one: a careful examination of the numerous alleged examples of underdetermined spacetime theories reveals only a transient form of underdetermination or no scientific underdetermination at all. But is transient underdetermination really a cause for concern? In the remainder of this essay, we explore two more examples from relativistic physics – Brans-Dicke theory and gauge theories of gravity – and Stanford's (2001) argument that the historical record shows transient underdetermination to be a recurrent problem (what he coins the 'New Induction') in science. We take a very different line to that common in the underdetermination literature (with the notable exception of Lyre and Eynck, 2001) and deliver a final assault on the antirealists' scientific underdetermination project. For we argue that not only is transient underdetermination (in Stanford's sense) a natural feature of scientific inquiry but the existence of empirically equivalent theories also plays a beneficial role in the scientific process by identifying key research areas and stimulating improvements in experimental testing.

3 | Two More Examples from Physics: Brans-Dicke and Gauge Theories of Gravity

There is a popular misconception among physics laymen, schooled by *A Brief History of Time* and Einstein's short *Relativity* that in the early-twentieth century, Einstein's general theory of relativity replaced Newtonian mechanics as the 'true physics', teaching us that space curves and unveiling some of Nature's other tricks and that was that. This is not terribly inaccurate, for general relativity has now enjoyed an uninterrupted reign of over ninety years, but in such a superficial statement of the history of relativistic physics which takes GTR for granted, there is so much missing. It ignores the ten years following the *anus mirabilis* where Einstein struggled to find his final field equations, just beating out Hilbert in an exhausting November 1915 (see Norton (2000) for a vivid, entertaining account of Einstein's 'final emergence into the light'). It ignores the mistakes, particularly Einstein's (1913) *Entwurf* theory, born in a jungle red in tooth and claw where Abraham (1912), Nordström (1913) and others strove to extend relativity theory to incorporate gravitation – "a stately beast, but it lacks three legs!", Einstein called Abraham's theory (Einstein (1912) letter to Ludwig Hopf in Norton 2000: 142). Lastly, it ignores the many competitors to GTR itself as physicists continue to seek a better understanding of gravitation and a unified theory of everything. In this section, we briefly explore two of these more recent rivals to general relativity, Brans-Dicke theory and gauge theories of gravity, both theories which suggest that Einstein may not have discovered the 'true physics' after all.

Our story begins with the Brans-Dicke scalar tensor theory of gravity. Recall that along with the principle of equivalence, one of the key motivations underlying Einstein's search for a relativistic theory of gravitation was the fulfillment of Mach's principle, the idea that physical systems can be accurately described in a fully relationalist manner with no appeal to absolute spacetime or preferred frames of reference (for Mach, inertial effects were to be explained relative to the fixed stars).[5] Seeking to extend the Machian program to the gravitational constant G (held fixed in both Newtonian theory and GTR), Brans-Dicke theory supplemented a non-flat affine connection with a scalar field whose values depended on the mass distribution of the universe, thus allowing G to fluctuate as a function of spacetime. An adjustable constant ω determined the respective influences of curved space and the scalar field on spacetime processes (note the similarity here to the constant κ in the flat-space metric equation $g'_{\mu\upsilon} = \eta_{\mu\upsilon} + \kappa F_{\mu\upsilon}$) introducing some uncertainty into the theory. But analogous to flat-space relativistic gravity theories, the observational consequences of Brans-Dicke theory still converged to those of GTR with large ω. For example, the deflection of light by the sun predicted by Brans-Dicke was 7% smaller than GTR's prediction when $\omega = 5$ but only 0.5% smaller when $\omega = 100$ (Will 1986: 155). As ω could be set at any value, the choice between Brans-Dicke theory and GTR was, in principle, underdetermined by all available evidence, though a high enough ω would lead one to reasonably wonder why we even needed the additional scalar field (whose effect was now negligible) in the first place.

Towards the late 1960s, Brans-Dicke became increasingly popular, the only gravitation theory to seriously challenge GTR in half a century. Will (1986: 156) recounts a passing joke that went around the Caltech relativity group while he was a student there: "On Monday, Wednesday, and Friday, we believe general relativity; on Tuesday, Thursday, and Saturday, we believe the Brans-Dicke theory (on Sunday, we go to the beach)." Interestingly, the first major success of Brans-Dicke theory came in 1966 and centered on the perihelion advance of Mercury. Dicke argued that if the sun was slightly flattened or oblate, GTR's predicted value for Mercury's perihelion precession would be too high (by approximately 3 arcseconds) yet by adjusting ω, the Brans-Dicke value could still match the revised perihelion estimate exactly (see Will 1986: 96-107 for details). Though not conclusive evidence favoring Brans-Dicke theory over general relativity (as the problem of Mercury's perihelion/solar oblateness sparked a lively debate that to this day remains unsettled), solar oblateness nevertheless pointed to the need for a high degree of experimental precision in comparing the predictions of the rival theories and, more importantly, delivered the shock that GTR might actually be wrong. A nice example of the strengthening of experimental objectives, Mariner 6 and 7 were now programmed to measure the time delay of light to a 2 percent precision rather than the usual 10 percent so as to distinguish between Brans-Dicke and general relativity (Will 1986: 157). Will recalls: "It was no longer adequate simply to detect an effect predicted by general relativity; one had to measure it with a high degree of precision."

Unfortunately for Brans-Dicke theory, the same enhanced experimental techniques that might have elevated the theory above general relativity soon led to its demise. As increasingly accurate experimental results agreed with GTR, feasible values for ω so that Brans-Dicke remained consistent with the evidence soared. A 1979 Viking result agreed with general relativity to 0.1 percent accuracy, implying that ω would need to exceed 500 for Brans-Dicke to hold (Will: 157). With the effect of the scalar field now insignificant, Brans-Dicke theory was abandoned.

Even before GTR's experimental battles with Brans-Dicke theory, however, another rival to general relativity was already emerging in the gauge theoretical approaches to gravity of Utiyama (1956), Kibble (1961) and Sciama (1962) (discussed in Lyre and Eynck, 2001). Inspired by the now-standard model of particle physics which characterizes the strong, electroweak and electromagnetic forces by the respective gauge groups SU(3), SU(2) x U(1), and U(1), gauge theories of gravity attempted to derive gravitational interactions from local symmetry requirements. Invariance of physical laws under a group of gauge transformations (*e.g.*, the Lorentz or Poincare group) led to empirically equivalent alternatives to general relativity theory. Replacing the standard covariant tensor formalism of GTR with a fiber bundle framework where fibers (copies of the symmetry group) are associated with points in a base space, or spacetime, gravitational gauge theories sparked another underdetermination problem in relativistic physics, though one eclipsed by the popularity of Brans-Dicke theory.

Advance forty years and the gauge gravity program has become a serious contender in the search for a quantum gravity theory. At the frontiers of spacetime physics is the flat-space Gauge Theory of Gravity (GTG) of Cambridge physicists Lasenby, Doran and Gull presented in two excellent survey articles by Hestenes (2005, forthcoming). Incorporating Hestenes' 'Geometric Calculus', GTG begins with a flat spacetime manifold M and a real geometric algebra defined on each point on M. Whereas in GTR the metric tensor represents the curved geometry of a semi-Riemannian manifold, the differential geometric structure of GTG is generated by gauge principles in the tangent algebra defined on M, rather than on the flat spacetime manifold itself. GTG requires two gauge principles: the global displacement gauge principle (DGP), asserting that spacetime is globally homogenous, and the local rotation gauge principle (RGP), asserting that spacetime is locally isotropic. Though Hestenes compares DGP to Einstein's General Principle of Relativity and RGP to the Principle of Equivalence, Hestenes still stresses that both DGP and RGP express physical equivalence of spacetime points and not simply the equivalence of observers (Hestenes 2005: 22, 24). Enforcing DGP thus requires the introduction of a physical field which is identified with gravitation (Hestenes 2005: 19).

For Hestenes, GTG is not a replacement but rather an extension of GTR:

> [M]y own view is that GTG is not fundamentally different from [GTR]. I see it, rather, as simplifying and clarifying the structure of [GTR] -- in a sense, as perfecting the foundations and completing the development of [GTR]. As a bonus, GTG brings the full power of geometric calculus to the analysis and solution of problems in the domain of [GTR], and it goes beyond [GTR] when torsion is included. (Hestenes 2005: 43)

Yet while GTG may 'perfect the foundations' of GTR, significant structural differences still exist between the two theories. For one thing, the spacetime of GTR is curved while GTG imposes curved-space geometry in a flat spacetime, a striking new turn in the debate over affine structure. The geometric calculus formalism also automatically includes spinors, allowing GTG to incorporate the Dirac equation and treat gravitational effects (such as spin precession) on electron motion. Linking the classical and quantum realms, the co-existence of Dirac spinors with the tensors of field theory in GTG holds out promise for a unified theory, though only electromagnetism has been integrated into GTG at present. Superstrings still predominate but as Hestenes (2005: 50) speculates: “It remains to be seen if extra dimensions are really needed.”

In the preceding sections of this essay, we have already extensively argued that the underdetermination of non-final theories only establishes a transient epistemic predicament. Even so, the wider picture that has now gradually emerged from our considerations of Brans-Dicke theory and gauge theories of gravity is one where the transient underdetermination of theory choice may actually help facilitate scientific progress. Consider the late 1960s and early-1970s when “The Brans-Dicke theory forced general relativity to confront experiment as never before.” (Will 1986: 14) Here the underdetermination of spacetime theories directly led to improvements in experimental accuracy and, correspondingly, the downfall of Brans-Dicke theory.[6] Similarly, future experiments might measure torsion and gravitational effects on electron motion to a level of precision that enables one to distinguish between GTR and GTG. Lyre and Eynck (2001: 19) also appreciate the heuristic element of the underdetermination of partial theories, what they call ‘practical underdetermination’, in pointing to open questions in our scientific knowledge: “Concrete practical examples of theory underdetermination can be seen as possible guidelines for the question of where and how scientific efforts should focus.” This is especially clear in the continued development of the gauge theoretic approach to gravity. While gauge theories may agree with GTR on observational consequences, at least in the case of spinless matter, the shift from Einstein’s principles and covariant tensors to gauge symmetry principles and the new formalisms of fiber bundles and geometric calculus may ultimately pave the way to a complete theory of quantum gravity.

4 | The New Induction

So far, the skeptic has been having a difficult time. First, attempts to ground the scientific underdetermination project in general philosophical arguments collapsed to Cartesian skepticism or demonstrated that the underdetermination of (non-final) theories is, at best, a transient concern. In concrete examples from spacetime physics, there was more of the same and, to exacerbate the skeptic's position, we have now seen that transient underdetermination even has silver linings by promoting enhancements in experimental physics and identifying fruitful research areas. The scientific underdetermination project against realism, then, seems like smoke and mirrors.

However, the skeptic still has one final argument in Stanford's (2001) 'New Induction':

> I propose the following New Induction over the History of Science: that we have, throughout the history of scientific inquiry and in virtually every scientific field, repeatedly occupied an epistemic position in which we could conceive of only one or a few theories that were well-confirmed by the available evidence, while the subsequent history of inquiry has routinely (if not invariably) revealed further, radically distinct alternatives as well-confirmed by the previous available evidence as those we are inclined to accept on the strength of that evidence. (Stanford: S9)

Echoing Laudan's (1981) classic Pessimistic Meta-Induction, Stanford's New Induction draws on the historical record to support the inference that if available evidence E confirms theory T at time t, then there is good reason to believe that at some time $t' > t$, scientists will discover an alternative theory T' at least roughly equally well-confirmed by E (note that T and T' need not be empirically equivalent as all Stanford's argument requires is that *E* supports both theories). Along with the advance of spacetime physics, Stanford provides a wealth of examples where this pattern is allegedly instantiated, including the progression from early corpuscularian chemistry to phlogiston theory to oxygen chemistry to atomic/contemporary physical chemistry, and the shift from Lamarck's autogenesis to Darwinian evolutionary theory. From these cases, Stanford concludes that as underdetermination, albeit of the transient kind, is a recurring problem in science, our epistemic predicament is really something to worry about.

It is hard to see what all the fuss is about. So long as scientists work towards a complete and adequate system of the world, the recurring transient underdetermination of partial theories relative to both actual and unconsidered alternatives is, at bottom, a natural feature of scientific progress. For science advances through debate and argumentation, through the development of a host of theoretical attempts to account for empirical phenomena at any given time, some which do not even get off the ground and others which are sharpened, refined or ultimately overthrown as our experiences accumulate. Returning to spacetime physics, we have good reason to believe that general relativity is presently underdetermined relative to a theory which scientists cannot yet imagine – for this theory is just a theory of quantum gravity, whose discovery would mark the advance of theoretical physics in unifying our understanding of classical and quantum physical processes. But the underdetermination of GTR and quantum gravity theory need not concern the realist for this underdetermination is consistent with a story where GTR is approximately true and theoretical physics is converging to the truth (*i.e.*, convergent realism). While the realist has not provided good reasons to be a realist (see

Note 4), this is not at stake here. The point that must be emphasized is this: *transient underdetermination is not a problem for science but rather a natural feature of scientific progress.*

5 | A Final Note on Skepticism

It is easy to be a skeptic. If this is not immediately clear from the astonishing popularity of Descarte's *cogito, ergo sum*, it should be from the two and a half centuries that have passed without a satisfactory solution to Hume's Problem of Induction. Scientists, philosophers, novelists and other nonbelievers doubt the existence of the universe, the existence of themselves, the truth of our 'miraculously' successful scientific enterprise, that the sun will rise tomorrow – and all are received with utmost seriousness. To be sure, there are sensible reasons to be a skeptic. As human beings, we are limited to observing the world through our senses so it is all too easy to imagine what strange demons and forces may lie in those dusty corners beyond our immediate experience (hence, Cartesian skepticism). Our experiences also unfold in a continuously evolving present so though we may assume and project, we inherently have no direct awareness of the future (hence, the Problem of Induction). Religious figures and fortune tellers may tell us otherwise, explaining the universe in terms of God's will and prophecy or seeing the future in the intricate patterns of coffee grounds – and they may be right – but this takes us outside the realm of scientific thought. Nonetheless, there are also hasty misguided reasons to be a skeptic. As we have attempted to show in this essay, scientific underdetermination is one of them.

NOTES

[1] Stanford (2001: S7) credits Lawrence Sklar with introducing the notion of *transient* underdetermination.

[2] Here 'dappled' alludes to Cartwright's (1999) *The Dappled World*.

[3] Interestingly, such a position was held by Einstein later in his life (see Norton (2000) for a wonderful account of Einstein's Canon of Mathematical Simplicity).

[4] To avoid confusion, it must be emphasized here that this paper is not a call to realism. The reader may balk at our suggestion that physicists use pragmatic criteria in theory choice given our previous doubts that such criteria are guides to truth. But we are not trying to justify a realism interpretation of science and readily acknowledge that with 'No Miracles' arguments still plagued with circularity, such a justification has yet to be provided. Our restricted focus has rather been to show that scientific underdetermination is *not* a good reason to be an antirealist.

[5] see Sklar (1974) for a nice account of exactly how GTR is not a Machian theory.

[6] Note that Brans-Dicke theory is still, in principle, empirically equivalent to GTR if we allow an arbitrarily large ω but in a broader sense, this creates only a transient underdetermination problem as both relativistic theories are only partial theories, neither able to account for quantum phenomena.

REFERENCES

Curd, Martin and J. A. Cover (1998) *Philosophy of Science*. New York, London: W. W. Norton & Company. [C&C]

Duhem, Pierre (1904-5) Physical Theory and Experiment. In C&C (pp. 257-279).

Earman, John (1993) Underdetermination, Realism, and Reason. *Midwest Studies in Philosophy* 18 pp. 19-38.

Friedman, Michael (1983) *Foundations of Space-Time Theories: Relativistic Physics and Philosophy of Science*. Princeton: Princeton University Press.

Glymour, Clark (1977) The Epistemology of Geometry. *Noûs* 11 (3) pp. 227-251.

Hestenes, David (2005) Gauge Theory Gravity with Geometric Calculus. *Foundations of Physics* 35 (6) pp. 903-970.

Hestenes, David (forthcoming) Spacetime Geometry with Geometric Calculus. *Preceedings of the Seventh International Conference on Clifford Algebra.*

Hoefer, Carl and A. Rosenberg (1994) Empirical Equivalence, Underdetermination, and Systems of the World. *Philosophy of Science* 61 (4) pp. 592-607.

Ladyman, James (2002) *Understanding Philosophy of Science*. Routledge.

Laudan, Larry (1981) A Confutation of Convergent Realism. In C&C (pp. 1114-1135).

Laudan, Larry and J. Leplin (1991) Empirical Equivalence and Underdetermination. *The Journal of Philosophy* 88 (9) pp. 449-472.

Lyre, Holger and Tim O. Eynck (2001) Curve It, Gauge It, or Leave It? Practical Underdetermination in Gravitational Theories. E-Print: PhilSci-Archive, ID 514.

Magnus, P. D. (2005) Underdetermination of Theories. In S. Sarkar and J. Pfeifer (eds.) *The Philosophy of Science: An Encyclopedia* (pp. 839-842). Routledge.

Maxwell, Grover (1962) The Ontological Status of Theoretical Entities. In C&C (pp. 1052-1063).

Norton, John (1989a) Einstein, Nordström and the Early Demise of Scalar, Lorentz Covariant Theories of Gravitation. In D. Howard and J. Stachel (eds.) *Einstein Studies Volume I: Einstein and the History of General Relativity* (pp. 1-74). Boston: Birkhäuser.

Norton, John (1989b) How Einstein Found His Field Equations, 1912-1915. In D. Howard and J. Stachel (eds.) *Einstein Studies Volume I: Einstein and the History of General Relativity* (pp. 101-159). Boston: Birkhäuser.

Norton, John (2000) 'Nature is the Realisation of the Simplest Conceivable Mathematical Ideas': Einstein and the Canon of Mathematical Simplicity. *Studies in History and Philosophy of Modern Physics* 31 (2) pp. 135-170.

Quine, W. V. (1951) Two Dogma of Empiricism. In C&C (pp. 280-301).

Reichenbach, Hans (1958) *The Philosophy of Space and Time*. New York: Dover.

Sklar, Lawrence (1974) *Space, Time, and Spacetime*. Berkeley: University of California Press.

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

Will, Clifford (1986) *Was Einstein Right? Putting General Relativity to the Test*. New York: Basic Books.