1. Introduction

It is generally believed that Einstein’s special theory of relativity has changed our understanding of space and time and of the motion of objects in two important ways. First, we no longer think, as adherents of the classical electromagnetic world-picture at the turn of the twentieth century did, that there is a privileged class of inertial frames, the ‘ether rest frames’; and, second, it is widely believed that relativistic phenomena, such as length contraction, do not require a detailed explanation in terms of electromagnetic forces or quantum mechanical interactions—that is, a dynamical explanation—but that they are purely kinematic effects: they are artifacts of the relative state of motion of the frames used for representation and a consequence of the geometry of spacetime. This orthodox view has recently been challenged by Harvey Brown and Oliver Pooley (Brown 2005; Brown and Pooley 2006), who agree with the first part of the orthodoxy but not with the second. According to what Brown takes to be the orthodox view, the structure of spacetime—that is, the fact that spacetime is Minkowskian—explains the fact that the laws of our theories are Lorentz invariant and thereby accounts for the universal behavior of rods and clocks. But Brown argues that this view has the arrow of explanation backward and that relativistic phenomena ultimately require a dynamical explanation. According to Brown, it is a brute fact, which itself permits of no further explanation, that the laws are Lorentz invariant and it is this fact which explains that Minkowski spacetime is the proper arena in which to represent non-gravitational physical phenomena.

Brown’s view has recently been forcefully challenged by Michel Janssen (2009) and in this paper I investigate their dispute. I argue that there is less of a disagreement between the two positions than appears initially and, hence, that Brown’s view presents less of a departure from orthodoxy than it may seem. In section 2 I provide a first summary of Brown’s view and Janssen’s response. In section 3 I argue that the appearance of a rather stark disagreement is at least partially due to the fact that both Brown and Janssen...
try to frame their respective views in terms of Einstein’s well-known distinction between principle- and constructive theories. I appeal to an earlier, related distinction due to H. A. Lorentz involving a notion of general principle broader than Einstein’s to argue that Einstein’s distinction presents a false dichotomy. Once we locate Brown’s and Janssen’s respective positions within Lorentz’s framework, I argue in section 4, much of the disagreement between their respective views disappears and suggest as one potential remaining source of disagreement the role of modal and nomic constraints in scientific explanation. I focus exclusively on special relativity in this paper.

2. Physical Relativity: First Pass
One of Brown’s targets is the view that spacetime substantivalism plays an important role in explaining relativistic effects. The substantivalist takes spacetime to be an entity in its own right. Once we know that objects ‘live’ in Minkowski spacetime, and satisfy the constraints of Minkowski geometry, there is a simple well-known geometric construction that allows us to derive length contraction and time dilation. Thus, length contraction and time dilation appear to be purely geometric effects, which are a straightforward consequence of the structure of Minkowski spacetime. But Brown argues that merely appealing to the structure of spacetime does not answer the question as to why objects obey the constraints of Minkowski geometry. He makes this point especially forcefully in connection of what he takes to be a closely related issue, that of geometric explanations of inertial effects: In positing that force-free objects simply follow the geodesics of Minkowski spacetime, determined by its affine structure, the substantivalist needs to assume that objects have “space-time feelers” (Brown 2005, 24) that are somehow able to sense “the ruts and grooves” of spacetime and this assumption for Brown is the “mystery of mysteries” (143). Spacetime structure, thus, can play no explanatory role and merely serves “as codification of certain key aspects of the behaviour of particles (and/or fields).” (Brown 2005, 25)

By contrast, what can explain the behavior of rods and clocks are the “the details of the bodies’ microphysical constitution” (Brown and Pooley 2006, 76): the explanation of Lorentz contraction “is ultimately to be sought in terms of the dynamics of the microstructure of the contracting rod.” (Brown 2005, 133) A moving rod contracts, that is,
“because of how it is made up and not because of the structure of its spatio-temporal environment.” (Brown 2005, 8, italics in original) Thus, Brown argues for the priority of the dynamical laws over spacetime structure. Another distinction that plays a central role in the account is Einstein’s well-known distinction between principle and constructive theories. Put in terms of this distinction Brown’s thesis becomes the claim that the behavior of rods and clocks calls for a constructive explanation, which provides a dynamical model of the objects’ detailed microscopic structure.

Against Brown, Michel Janssen defends the view that the behavior of rods and clocks ultimately is explained by the geometric structure of spacetime. Janssen agrees with Brown in his rejection of spacetime substantivalism but argues that even within a relationalist framework, which takes spatiotemporal relations to have no existence independent of physical objects instantiating them, the explanatory arrow is from the structure of Minkowski spacetime to the Lorentz invariance of the dynamical laws. As Janssen puts it: “I argue that the space–time symmetries are the explanans and that the Lorentz invariance of the various laws is the explanandum; Brown argues that it is the other way around.” (26) Central to his argument against Brown are case studies of several phenomena for which pre-relativistic explanations in terms of detailed micro-dynamical models were eventually replaced by relativistic explanations showing that the phenomena in question followed simply from the general assumption of Lorentz invariance. Instead of needing to appeal to distinct detailed dynamical models for different relativistic phenomena “the statement that space–time is Minkowskian explains all in one fell swoop” (2009, 49).

Thus, on first sight there seems to be a rather stark disagreement between Brown’s and Janssen’s views. Brown, as Janssen and others read him, appears to believe that only an account of the particular forces pushing and pulling the microscopic constituents of a rod can explain length contraction; Janssen by contrast, maintains that the structure of Minkowski spacetime explains the phenomenon and that explanatory advantage of Einstein’s theory over Lorentz’s consists precisely in the fact that the special theory of relativity teaches us that certain phenomena are independent on the detailed assumptions about the dynamics governing a particular system. Brown argues that the behavior of rods and clocks ultimately calls for a dynamical explanation; while Janssen argues that the
explanation is kinematic. And while both agree that the correct or best explanation of length contraction is constructive, for Janssen the correct constructive explanation appeals to the geometry of spacetime rather than to the micro-dynamics of rods.

Curiously, however, both Brown and Janssen appeal to Wolfgang Pauli in support of their respective views. Brown approvingly quotes (three times in his book!) Pauli’s claim that “the contraction of a measuring rod is not an elementary but a very complicated process” (Pauli 1921, 15). Janssen cites Pauli saying that “it constituted a definite progress that Lorentz’s law of the variability of mass could be derived from the theory of relativity without making any specific assumptions on the electron shape or charge distribution. Also nothing need be assumed about the nature of the mass.” (cited in Janssen 2006, 38-9) While the first quote appears to suggest that a proper understanding of length contraction needs to invoke the detailed and complicated dynamics governing the rod, the second quote seems to imply that it is precisely such a detailed model that relativity theory has shown to be superfluous in accounting for so-called ‘relativistic phenomena.’ Yet, contrary to first impressions, there is no tension between the two quotes and the fact that both Brown and Janssen cite Pauli’s review article on relativity in support of their respective views can serve as a first clue that their disagreement might be less than initial appearances suggest. Since the appearance of a stark disagreement between Brown and Janssen is to a large part due to their appeal to Einstein’s principle- vs. constructive theory distinction, I want to begin my reconstruction with a discussion of this distinction.

3. Principles and Mechanisms

Albert Einstein, in a letter to the London Times, famously introduced a distinction between what he called “constructive theories” and “principle-theories.” (Einstein, 1919 [1954], 228) According to Einstein, we can distinguish constructive theories that treat phenomena as complex and build them up out of “the materials of a relatively simple formal scheme” from principle-theories that rely on empirically discovered, general characteristics of natural processes. As it is usually understood, Einstein’s distinction is a distinction between theories that describe the directly observable macroscopic behavior of a system with the help of phenomenological principles, which are elevated to the status of postulates, and theories that describe a system’s behavior by proposing a model of its
detailed microscopic constitution. This interpretation fits well with the examples Einstein gives of the two kind of theories: the kinetic theory of gases, on the one hand, which derives the behavior of gases from a microscopic model; and thermodynamics, one the other had, which is based on phenomenological principles such as the principle that there can be no perpetual motion.

Einstein says that both kinds of theory have advantages—constructive theories are complete, adaptable, and clear while the advantage of principle theories are “logical perfection and secureness of foundations”—yet he is unequivocal in his preference for constructive theories: “When we say that we have succeeded in understanding a group of natural processes, we invariably mean that a constructive theory has been found which covers the process in question.” (Ibid.) That is, according to Einstein, we simply do not understand a natural process unless we are in possession of a constructive theory of that process. Brown and Pooley follow Einstein in their assessment of the relative merits of the two kinds of theory. They, too, believe not only that constructive theories are explanatorily superior to principle-theories but that “principle theories fail to be explanatory.” (Brown and Pooley 74) Thus, principle theories are “explanatorily inferior,” according to Brown and Pooley in quite a dramatic fashion: they simply “fail to provide any sort of explanation” (Ibid.)

In his letter to the Times Einstein maintains that special relativity is a principle-theory, based on the relativity postulate and the light postulate. The behavior of rods and clocks is derived in the theory by showing that in any world in which the relativity postulate and the light postulate holds all objects, independently of their microphysical constitution, must necessarily exhibit the phenomena of length construction and time dilation. Now, at first sight it might seem rather surprising that Einstein, in a letter devoted to outlining his theory of relativity, would deny that this theory can yield understanding. Brown proposes the following interpretation of this. In 1905 no satisfactory constructive theory that could account for the stability of matter was available and a principle-theory approach that simply postulated stable macroscopic measuring rods and clocks was the only option. But this approach was merely preliminary for Einstein, who, according to Brown, believed that a full understanding of these phenomena had to await the development of a satisfactory constructive micro-theory of matter. Einstein himself, of
course, never was happy with the micro-theory that was eventually to be developed—that
is, with quantum mechanics. But according to Brown it is precisely that theory, perhaps in
the form of a quantum field theory, that provides us with the ultimate explanation of length
contraction “in terms of the dynamics of the microstructure of the contracting rod.”
(Brown and Pooley 2006, 77)

That is, Brown maintains that Einstein’s 1905 principle-theory approach played an
important historical role—it provided logically secure foundations for the treatment of a
class of phenomena for which no promising constructive approach was available at the
time—but a genuine understanding of the behavior of rods and clocks had to await the
development of a constructive theory that provided a microscopic dynamical model for
rods and clocks. The principle-theory approach to relativity shows that rods and clocks
must behave in quite peculiar ways for the two postulates of the theory to be true together.
But this, Brown and Pooley insist, does not constitute an explanation of the behavior of
rods and clocks. Rather, explaining length contraction “involves solving the dynamics
governing the structure of the complex material body that undergoes contraction.” (Brown
and Pooley 2006, 79) According to Brown, then, it is a mistake to think of special relativity
only, or even primarily, as a principle theory. Rather, the proper way to understand
relativity theory post-1905 is as a constructive theory.

As we have seen, Brown and Janssen distinguish a third kind of approach to
relativity theory, aside from Einstein’s 1905 principle-theory and the constructive
approach that Brown favors: a geometric approach explaining the behavior of rods and
clocks by appealing to the structure of Minkowski spacetime. According to Janssen, the
geometric interpretation of the theory also provides a constructive explanation of the
relativistic phenomena with Minkowski spacetime providing the constructive model. A
geometric constructive explanation of length contraction proceeds “by showing that two
observers who are in relative motion to one another and therefore use different sets of
space-time axes disagree about which cross-sections of the ‘world-tube’ of a physical
system give the length of the system.” (Balashov and Janssen 2003, 331)

In response, Brown and Pooley maintain that there are two distinct spacetime
interpretations of relativity, neither of which, however, provides an adequate explanation
of the behavior of lengths and rods. First, one might posit Minkowski spacetime as an
ontologically autonomous element and maintain that the behavior of rods and clocks is constrained to reflect the geometric structure of the spacetime they inhabit. Brown concedes that this approach is genuinely constructive, since it posits Minkowski spacetime as reality ‘behind the phenomena’, yet he maintains, as we have seen above, that it fails to explain the behavior of rods and clocks it has to posit a mysterious influence of spacetime on objects: “How,” he asks, “are rods and clocks supposed to know which space-time they are immersed in?” (Brown 2005, 143) Alternatively one might try appeal to the constraints of Minkowski spacetime within a relationalist framework, as proposed by Janssen. Brown and Pooley reject this account on the grounds that it is not truly constructive, since geometric features in question are phenomenological and do not directly “concern the details of the bodies’ microphysical constitution.” (Brown and Pooley 2006, 76) Hence, the geometric features have a status akin to the postulates of a principle theory and, thus, are not explanatory: “They are about aspects of their (fairly) directly observable macroscopic behaviour. And this reflection prompts an obvious question: why do these objects obey the constraints of Minkowski geometry?” (Ibid.)

According to Brown, then, following Einstein, scientific theories fall into one of two categories: purely phenomenological principle theories, on the one hand, and constructive theories that provide a detailed microscopic model of the reality behind the phenomena, on the other. Since only theories of the latter kind can be explanatory, relativity theory would have to be a constructive theory, if it is to be explanatory. But as presented so far, Brown’s account is deeply puzzling. Has the lesson of the theory of relativity not precisely been that a wide class of phenomena that might have appeared to depend on the micro-structure of objects is independent of the details of their constitution? This, of course, is the point that Janssen stresses forcefully in his recent case studies (2009). But while I agree with Janssen on this point, it seems to me that his criticism of Brown is misdirected. To be sure, given the way Brown attempts to locate his interpretation of the theory of relativity within the framework provided by Einstein’s principle-constructive distinction, it is difficult not to read him as advocating that a proper explanation of relativistic phenomena would involve solving the detailed microscopic equations of motion. But, I want to suggest that part of the fault lies with Einstein’s distinction, construed as exhaustive, and that there is a better
framework available to capture Brown’s view—a view which, I want to claim, is compatible with the lesson Janssen wishes to draw from his case studies.

Even though the distinction between principle- and constructive theories is usually traced to Einstein, a related distinction had almost two decades earlier been drawn by Lorentz (see Frisch 2005). (Lorentz 1900) distinguishes theories that begin by postulating “general principles” or “general laws” from theories that postulate a “mechanism of the phenomena” from theories. General principles, according to him, have the advantage of being versatile and applying to a wide variety of phenomena, since they abstract from and are independent of “the inner constitution of bodies.” Mechanism theories, by contrast, provide us “with flawed, yet lively representations” of the “connections between and the nature of things.” On first sight this distinction might strike one as being equivalent to Einstein’s later distinction between phenomenological and constructive micro-theories. Similar to Einstein, Lorentz takes principles to express “generalized experiences” and Lorentz’s mechanisms seem to be the same thing as Einstein’s constructive models. But Lorentz’s notion of principle is broader than Einstein’s (at least as generally interpreted), for as an example of a general principle Lorentz cites not only the purely phenomenological second law of thermodynamics but also the principle of energy conservation. Even if it were correct that we arrived at this principle as the result of generalizing our experiences, the principle clearly is not purely phenomenological; rather it has become to function as a general, albeit perhaps defeasible, constraint on all physical theorizing. Thus, there is an important category of physical principles that Lorentz recognizes, but Einstein’s distinction between purely phenomenological principles and detailed constructive models of the phenomena leaves out: general constraints that guide theory construction and theory choice and that we take every physical theory, or perhaps every theory within a certain domain ought to satisfy. Indeed, it seems helpful to distinguish between the two kinds of principle theories Lorentz allows, and make the following tripartite distinction, which is implicit in Lorentz’s discussion:

(i) mechanism- or constructive theories, such as the kinetic theory of gases;

(ii) purely phenomenological principles, such as the second law of thermodynamics;
(iii) general principles or constraints on all (or at least multiple) levels, such as the principle of energy conservation.

Once we recognize the existence of such general constraints as an integral part physical theorizing, Einstein’s and Brown’s suggestion that only a detailed constructive account of a phenomenon can offer an explanation of that phenomenon becomes suspect, while Lorentz’s discussion seems closer to the mark. Lorentz preferred theories of mechanism, just as Einstein preferred constructive theories, but while Einstein maintained that only constructive theories can yield understanding, Lorentz stressed that principle-theories, too, can yield “insight” and prefaced his discussion of principle- and mechanism-theories with a commitment to an explanatory pluralism according to which “there are multiple ways by which we try to understand natural phenomena [...] Individual characteristics and inclinations determine the choice for each scientist.” That is, Lorentz unlike Einstein believed that both kinds of theories can yield understanding.

4. The ‘Big Principle’

What, then, is the theory of special relativity, in terms of Lorentz’s framework? It is a general principle theory with the principle of Lorentz-invariance as its core—a principle which is not merely phenomenological but which, like the principle of energy conservation, functions as a general constraint on all levels of theorizing, from phenomenological macro-theories to constructive micro-theories of matter. On my reading, then, there is a deep disanalogy between the principle of Lorentz-invariance and thermodynamic principles. The latter are purely phenomenological, while the former is not. The theory is not committed to accepting measurement rods and clocks as not further analyzable fundamentals into the theory, contrary to what Einstein’s 1905 paper might be taken to suggest. Rods and clocks do have a micro-structure, and providing a model of this micro-structure can be explanatory, for example of the stability of macroscopic objects. Yet this is not part of relativity itself, and relativity is no more a constructive theory than it is purely

---

1 I take it that am here assuming that the notions of explanation, understanding, and of gaining insight are closely related: to gain insight into a phenomenon is to increase one’s understanding of the phenomenon and explanations provide understanding or insight.
phenomenological. In particular, details of the micro-theory are not needed to account for the relativistic behavior of stable macroscopic objects, which is explained solely by appealing to the principle of Lorentz invariance.

Brown himself, in fact, seems to have recognized the limited value of Einstein's distinction and struggles in his book to locate his answer to the question "what is special relativity?" within Einstein's framework, despite his initial commitment to a constructive account. Brown's ultimate view, it seems to me is exactly the view I sketched here: the content of special relativity is given by the 'big principle' of Lorentz invariance (see Brown 2005, 146-7). But his problem then is how to think of this principle, which neither has the character of a purely phenomenological principle nor does it itself provide a detailed constructive model. Thus Brown concludes that “at its most fundamental, SR is a theory that lies somewhere in between a pure principle theory (like thermodynamics, or Einstein's 1905 version of SR) and a fully constructive theory (like statistical mechanics).” (Brown 2005, 147) In light of his initial insistence on the superiority of constructive accounts and his earlier discussion of explanation, this verdict is rather surprising, especially since it commits him to the claim that at its most fundamental relativity lies somewhere in between an explanatory theory and one that simply fails to explain!

Finally, thinking of relativity theory as a general principle theory, allows us to make sense of the two quotes from Pauli above. Since the theory is not a phenomenological theory, it does not take rods and clocks as fundamental and can recognize that the contraction of a rod is a very complicated process. But by the same token, the theory is not constructive and its value consists precisely in the fact that shows that paradigmatically relativistic effects are derivable without making any specific assumptions about the structure of matter.

Where does this leave the debate between Brown and Janssen? In taking Brown to argue that properly interpreted relativistic explanations of time dilation and length contraction need to invoke the details of the micro-dynamics, Janssen appear to have been misled by Brown's attempt to fit his interpretation into Einstein's principle-constructive framework. In fact both agree that the ‘big principle’ of Lorentz invariance is all that is needed in order to account for these phenomena. According to Brown special relativity provides, in the principle of Lorentz invariance, a universal constraint on the nature of the
non-gravitational interactions (Brown 2005, 147). According to Janssen, relativity Lorentz invariance as a constraint that “transcends the individual laws.” (Janssen 2009, 28) Janssen calls a universal constraint that is independent of the details of a theory’s dynamics “kinematical”, while Brown calls the constraint “dynamical”—but this seems to be a disagreement about terminology rather than about substance.

Thus, Brown and Janssen agree on the overarching lesson we can draw from Janssen’s case studies—namely that the general constraint of Lorentz-invariance is a common source of all instances of relativistic behavior. Yet Janssen wants to go beyond that and maintains that the structure of Minkowski spacetime provides a common origin explanation of universal Lorentz invariance. (For Brown, by contrast, the appropriate geometric structure in which to represent the motion of physical objects is Minkowski space-time, precisely because the dynamical laws are Lorentz-invariance.) Yet I share Brown’s puzzlement about this aspect of Janssen’s view (see, e.g., Brown and Pooley 2006, 81), since Janssen explicitly agrees with Brown that the role of Minkowski spacetime is to “encode” the Lorentz invariance of the laws. It is difficult to see how Minkowski spacetime can merely serve to encode the Lorentz-invariance of the laws and yet at the same time also provide an explanation of Lorentz-invariance. Indeed, it would seem that whatever information or claim is encoded is explanatorily prior to the structure in which it is encoded.

Perhaps, however, Janssen’s insistence that Lorentz-invariance is to be explained by something beyond that principle points to what be perhaps the most interesting potential disagreement between his and Brown’s views, even though neither stresses this point explicitly, and this is a disagreement concerning the modal status of Lorentz invariance. According to Brown and Pooley, “the Lorentz covariance of all the fundamental laws of physics is an unexplained brute fact.” But, they continue: “This, in and of itself, does not count against the interpretations: all explanation must stop somewhere.” (Brown and Pooley 2006, 80) Janssen, however, criticizes Brown’s proposal precisely for this feature and argues that the Lorentz-invariance of all of our laws ought not to come out as a “cosmic coincidence” (Janssen 2009, 48). Lorentz-invariance, on Janssen’s account, characterizes the “generic” or “default” behavior of physical systems and he stresses that he wants to endorse what he takes to be the normative connotations of this characterization.
Now, taken at face value what Brown and Janssen say is puzzling. For if Lorentz-invariance is indeed a law—that is—is nomologically necessary, how can it also be a ‘brute fact’ or a ‘coincidence’? One way to make sense of Janssen’s worry is to introduce the notion of a meta-law or meta-constraint, which bears the same relation to the laws as the laws bear to particular matters of fact. Lorentz-invariance, then, might be thought of as such a ‘super law’ in Eugene Wigner’s sense, which “provides a structure or coherence to the laws of nature just as the laws provide a structure and coherence to the set of events.” (Symmetry and Conversation Laws 1964, p. 16-7)

We may have finally reached a point of genuine disagreement: Is universal Lorentz-invariance a meta-law that can explain why all particular dynamical laws are Lorentz-invariant, or is Lorentz-invariance merely a property of all physical laws that that cannot be further explained and has no higher ‘grade’ of necessity than that of nomological necessity? I think what one’s answer to this question is will depend strongly on one’s views on the metaphysics of modality. I take it that someone who believes in nomological necessity and in grades of nomological necessity as fundamental ingredients in the inventory of the world, would find an appeal to Lorentz-invariance as meta-nomological constraint explanatory. It is less clear to me how someone with strong Humean intuitions would view the situation. To be sure, a Humean might be able to mimic all the appeals to the explanatory force of laws, by invoking a version of the Mill-Ramsey-Lewis view of laws. But since MRL laws, and by extension any putative meta-law, are merely summaries of the mosaic of particular matters of fact, it is not clear how one might recover the distinction between different grades of necessity. In fact, it is unclear whether any putative meta-law such as the principle of Lorentz-invariance could even be an axiom of the best system, given that, by hypothesis, it is implied by all the dynamical laws. If a system includes all dynamical laws, adding Lorentz-invariance leads to a less simple system without gain in deductive strength. A system without the dynamical laws and only Lorentz-invariance is much simpler but is much less strong than a system including the dynamics.

Unfortunately Brown’s and Janssen’s remarks are ambiguous on this issue. On the one hand, Brown calls Lorentz-invariance a brute fact, but on the other hand he calls it a “super law.” And Janssen searches for an account of Lorentz-invariance that has normative force, but finds the notion of nomic necessity “questionable.” (Janssen 2009, 28 fn. 7)
Perhaps here, too, there is ultimately less of a disagreement than it may seem and both would like to endorse a broadly Humean conception. But if we adopt a Humean best system account of laws together with spacetime relationalism, then it becomes even more difficult to see how, as Janssen maintains, the geometry of Minkowski spacetime could be explanatorily prior to the Lorentz invariance of the laws. Both the choice of geometrical structure in which to represent the motion of physical objects and the choice of dynamical laws, on this view, is determined by which choice yields the overall best system. The choice of metric, thus, is not prior to the determination of the laws but rather in conjunction with the laws. If the simplest laws are Lorentz-invariant, then the simplest choice of metric is the Minkowski metric. Or, as Brown and Pooley put it: “the appropriate structure is Minkowski geometry precisely because the laws of physics, including those to be appealed to in the dynamical explanation of length contraction, are Lorentz covariant.” (2006, 77)

References


