

## **Communication, Rationality, and Conceptual Changes in Scientific Theories\***

Frank Zenker and Peter Gärdenfors

Lund University, Department of Philosophy and Cognitive Science

Kungshuset, Lundagård, 22222 Lund, Sweden

frank.zenker@fil.lu.se, peter.gardenfors@lucs.lu.se

*Abstract:* This article outlines how conceptual spaces theory applies to modeling changes of scientific frameworks when these are treated as spatial structures rather than as linguistic entities. The theory is briefly introduced and five types of changes are presented. It is then contrasted with Michael Friedman's neo-Kantian account that seeks to render Kuhn's "paradigm shift" as a communicatively rational historical event of conceptual development in the sciences. Like Friedman, we refer to the transition from Newtonian to relativistic mechanics as an example of "deep conceptual change." But we take the communicative rationality of radical conceptual change to be available prior to the philosophical meta-paradigms that Friedman deems indispensable for this purpose.

*Keywords:* a priori, conceptual spaces, communication, communicative rationality, incommensurability, Michael Friedman, neo-Kantianism, scientific revolution, theory change, Thomas Kuhn

\* To be presented at the 2014 Meeting of the Nordic Network for Philosophy of Science, 27-28 March 2014, Lund University, Sweden. A previous version was presented at the conference Conceptual Spaces at Work, 24-26 May 2012, Lund University, Sweden.

## 1. Introduction

Thomas Kuhn (1970) famously argued that successive paradigms in the sciences are incommensurable. It is therefore a challenge for philosophers of science to explain how adherents of successive paradigms can nevertheless engage in rational communication during periods of “deep conceptual change.” The transition from Newtonian to relativistic mechanics is widely accepted as *the* paradigmatic example. Michael Friedman’s (2001; 2002a,b; 2008; 2010) strategy in addressing this challenge is to divide the scientific discourse into three levels. In order to account for rational communication about competing theory frameworks (paradigms), he argues that meta-philosophical arguments are required.

Common to Kuhn’s and Friedman’s positions is the assumption that scientific theories are *linguistic entities* and that successful communication about them implies a realist view of semantics whenever theories predict natural phenomena. As a theory’s terms—for instance, *mass*—then apply by *reference* to things—for instance, planets—communicative impediments would appear to arise because of the differences in how referents are construed across paradigms.

Our aim in this article is to present an alternative way of meeting Kuhn’s challenge. Our strategy is to see a theory as a constraint on a spatial structure (the framework of a theory). We model such frameworks in terms of conceptual spaces (Gärdenfors, 2000; Gärdenfors and Zenker, 2013). By analyzing different kinds of changes within such spatial structures, one can understand Kuhn’s challenge in a way that eschews incommensurability and that does not require meta-philosophical discussions.

The communicative impediments arising in deep conceptual change may thus be alleviated by focusing on the structural representation of concepts in

conceptual spaces instead of their possible referents. This shifts focus away from referents and towards the structure of concepts. The theory of conceptual spaces, then, is employed as a particular example of a more general answer to Kuhn's challenge. We also believe that our perspective lies closer to what scientists actually use in their thinking and communication. The paper proceeds as follows: we present how conceptual frameworks can be understood as spatial entities (Sect. 2), how the dynamics of such frameworks may be analyzed in the diachronic case (Sect. 3), and how to address communicative challenges that are caused by meaning change (Sect. 4). We then contrast our approach with Michael Friedman's neo-Kantianism (Sect. 5), and argue that an explanation of rational communication between proponents of successive paradigms need not rely on meta-philosophical paradigms (Sect. 6).

## **2. Conceptual Frameworks as Spatial Entities**

The theory of conceptual spaces builds on a cognitivist view of semantics. It contrasts with both extensional and intensional realistic semantics that include the referent of a linguistic expression as a meaning constituent. On our approach, conceptual knowledge is seen as mental representations, modeled in conceptual spaces. Since we do not believe in the existence of a mental language ("Mentalese"), we do not view such representations as parts of a symbolic system with a syntactic or logical structure. Instead, we treat them as spatial structures that can be analyzed into their constitutive dimensions and properties, representing the semantic knowledge of an agent (Gärdenfors, 2000, 2014).

An empirical theory always presupposes a specific conceptual framework that provides the magnitudes, or dimensions, on which the formulation of this theory depends. These magnitudes can be modeled as collections of dimensions with their inter-relations, that is, as conceptual spaces (Gärdenfors, 2000; Gärdenfors & Zenker, 2011; 2013; Zenker & Gärdenfors, 2013; Zenker, 2014). Put schematically, an empirical theory, T, depends on a conceptual framework, F, that is modeled as a conceptual space, S.

Apart from concepts arising in sensory perception, e.g., color or sound, or in basic scientific measurements, magnitudes include those introduced by science, for instance, *mass*, *force*, and *energy*. The topological or metrical structure of such magnitudes is tightly connected to the methods by which they can be measured, and to their relations to other concepts within a scientific theory. *Newtonian mass*, for instance, can be modeled as a dimension that is isomorphic to the non-negative part of the real number line, and Newtonian space as a three dimensional integral (vector) space with Euclidian metric. (We return to this example below.)

The primary role of the dimensions is to represent various “qualities” of objects in different *domains*. The notion of a domain can be given a more precise meaning by using the notions of *separable* and *integral* dimensions. Dimensions are said to be integral if, to describe an object fully, one cannot assign it a value on one dimension without giving it a value on the other. For example, an object cannot be given a hue without also giving it a brightness value. Or the pitch of a sound always goes along with its loudness. Dimensions that are not integral are said to be *separable*, as for example the size and hue dimensions. Within the context of scientific theories, the distinction should rather be defined in terms of *measurement*

*procedures*. If two dimensions (or sets of dimensions) can be measured by independent methods, then they are separable, otherwise they are integral.

It is therefore part of the meaning of “integral dimensions” that they share a metric, which separable dimensions do not. A theory *domain* can now be defined as the set of integral dimensions that are separable from all other dimensions in a theory, for instance Newton’s domain of absolute space.<sup>1</sup>

The choice of domains need not be uniquely determined, for the organization of dimensions into domains depends on which dimensions are considered basic for a theory. In the literature one finds organization principles that are theoretically motivated, e.g., via transformation classes, or motivated by practical considerations, e.g., via measurement procedures. The set of domains thus depends on the choice of basic dimensions (notably the basic ones given by the International System of Units—SI-units), and the choice of useful derived dimensions. Both ontological and epistemological considerations have historically informed which dimensions are considered basic and which are derived, but our notion of what constitutes a domain is an instrumentalist one. This notion is independent of ontological positions, and while it is compatible with calling certain dimensions basic or fundamental that are found to be epistemologically prior, no such distinction is required.<sup>2</sup>

---

<sup>1</sup> More precisely, domain C is separable from D in a theory, if and only if the invariance transformations of the dimensions in C do not involve any dimension from D; and the dimensions of a domain C are integral, if and only if their invariance class does not involve any other dimension (Gärdenfors & Zenker, 2013).

<sup>2</sup> Albert Einstein, for instance, appears to have been well aware of the Kantian tradition, yet opposed to singling out some concepts as privileged or epistemologically prior:

Contrasting with our instrumentalist notion, Kuhn's incommensurability thesis derives from a view on semantics that requires all concepts to have referents. His paradigmatic example in support of the thesis is that changes in meaning of the scientific variables (dimensions) go along with a change of referent. He writes:

“The variables and parameters that in [the statements which Kuhn refers to as] the Einsteinian  $E_i$ 's represented spatial position, time, mass, etc. still occur in the [Newtonian]  $N_i$ 's and they still represent Einsteinian space, time and, mass. But the *physical referents* of these Einsteinian concepts are by no means identical with those of the Newtonian concepts that bear the same name. (Newtonian mass is conserved; Einsteinian is convertible with energy. Only at low relative velocities may the two be measured in the same way, and even then they must not be conceived to be the same.)” (Kuhn, 1970, 101f.; italics added)

And it is vis-à-vis such a semantic realism that Kuhn can claim that the predecessor and the successor theory in a paradigm shift are non-intertranslatable.

Since our position is basically instrumentalist, we do not count a theoretical term such as *mass* as one that has a direct reference in an external world.

---

“One must not ... speak of the ‘ideality of space.’ ‘Ideality’ pertains to all concepts, those referring to space and time no less and no more than all others. Only a complete scientific conceptual system comes to be univocally coordinated with sensory experience. On my view, Kant has influenced the development of our thinking in an unfavorable way, in that he has ascribed a special status to spatio-temporal concepts and their relations in contrast to other concepts.” (Einstein, 1924, 1690f.; cited after Howard, 2010, 347)

Consequently, we can analyze, and then compare, different mass concepts solely in virtue of their distinct dimensional reconstitutions, and their relations to other concepts.<sup>3</sup> *Newtonian mass*, for instance, is separable from everything else in Newtonian theory; it constitutes a separate dimension because an object's mass remains constant, and so is independent of variations in other magnitudes such as the object's position or velocity. *Relativistic mass*, in contrast, is integral with *energy* in special relativity: the mass of a moving object increases as its kinetic energy increases, given the object has non-zero rest mass. The term *mass* thus takes on different meanings in these two theories. Both meanings can nevertheless be compared—something Kuhn was well aware of (d'Agostino, 2013). But, as we will argue, it is misleading to present the meaning difference as giving rise to communicative difficulties of the kind that Kuhn took to support his claim according to which proponents of successive paradigms *cannot* engage in rational discourse—and so they *must* resort to less rational forms of interaction (see below).

Within a cognitive theory of meaning, as we have stated, the meaning of *mass* is analyzed non-realistically through a dimensional (spatial) representation. This view of the semantics of a scientific concept can be extended so that the dynamics of a scientific conceptual framework may be treated in like manner. We do this by analytically separating the types of changes that modify a predecessor framework into a successor framework. The term *framework* we here understand as the conceptual space that a scientific theory builds on.

---

<sup>3</sup> Conversely, allegedly different concepts are in fact not distinct when sharing the same meaning in virtue of their identical dimensional forms and identical relations to other concepts.

Applying this typology yields a systematic analogue to the actual historical dynamics of theory frameworks insofar as a predecessor conceptual space can be modified into a successor space. As we argue below, it becomes plausible that, by understanding scientific concepts as spatial entities, scientists may successfully and rationally communicate about different concepts, their non-referentially conceived meanings, and thus about the dynamics of empirical theories in which these concepts are used. The threat of incommensurability is thereby eschewed. Before considering the arguments for this, we first turn to the types of changes that modify a conceptual framework.

### **3. The Dynamics of Conceptual Frameworks for Theories**

In earlier work, we have provided a range of historical examples of five types of changes to a scientific conceptual framework (Gärdenfors & Zenker, 2011, 2013). Here we restrict ourselves to a brief overview. In order of increasing severity of change, the following are distinguished: addition and deletion of special laws; change of metric; change in importance; change in separability; addition and deletion of dimensions.

Firstly, in our model, the special laws of a theory provide constraints on the distribution of points over a conceptual space. Newton's second axiom, for instance, restricts points to the hyper-surface described by  $F=ma$  in the conceptual space consisting of the domains *time*, *mass*, (physical) *space*, and *force* (Gärdenfors, 2000; Gärdenfors & Zenker, 2013). The law of gravitation further restricts these points to regions where  $F=GMm/r^2$  holds (Zenker, 2009, 41f.). The addition or deletion of



axioms and special laws is the mildest type of change, as it leaves the conceptual framework intact.

Secondly, each domain is endowed with a scale or metric that determines distances. For instance, day and nighttime are standardly modeled by one circular dimension with 24 equally-sized intervals called *hours*. Before the invention of mechanical clocks, however, the two points on this dimension that separate twelve hours of nighttime from twelve hours of daytime were commonly coordinated *locally* to sunrise and sunset. As these points shift, again locally, over the course of one year, their distance changes.<sup>4</sup> This variable temporal metric has meanwhile given way to one of constant clock intervals. The same occurred in the case of space—yet in the inverse direction, namely from constant to variable—where the Euclidian metric assumed by Newton first gave way to a Minkowskian (special relativity) and then to a Schwarzschild metric (general relativity), which may both be viewed to generalize the Euclidian. It is easy to see that a change of metric leads to a change in the symbolic expressions used to calculate with these distances.

Thirdly, the extant dimensions (or domains) provided by a scientific framework may change in importance. Energy, for instance, remained of little importance to Newtonian mechanics (Gärdenfors & Zenker, 2013). In contrast, forces became ever more important in the development of 19<sup>th</sup> century fluid dynamics, amongst others through a separation into short and long range forces (Petersen & Zenker, 2014).

---

<sup>4</sup> Consequently, the expression “days are longer in summer than in winter” was literally true although nights and days were always twelve hours long. But the length of an hour varied from day to day, ever more markedly so over the contrasting seasons at locations further away from the equator.

Fourthly, more severe yet are changes in the separability of dimensions, for instance the integration of mass with energy in relativistic physics, discussed above, or the formation of integral 4D space-time.

Fifthly, the most radical type of change is the addition to or the deletion of a dimension from a scientific framework as in the addition of the dimension *charge* in electrodynamics, or the abandonment of a dimension known as *caloric* in early accounts of thermo-dynamical phenomena.

The first of the five types addresses *intra*-framework changes; the other four types describe *inter*-framework changes. Categorizing changes of theories into these five types provides a finer grain than Kuhn's distinction between normal and revolutionary change. This gives us a richer toolbox to analyze the changes of a theory or its associated framework over time (for a case study see Gärdenfors & Zenker, 2013). What is commonly referred to as *scientific revolution*, we categorize primarily under the last type: the replacement of dimensions.

#### **4. Meaning Change and Communicative Challenges**

This typology of scientific changes in hand, one can now address cases where a predecessor conceptual framework is replaced by a successor framework, while the latter framework retains terms already used by the former. In short, one can address the kind of meaning change said to be typical of a “scientific revolution.” In contrast to proponents of realistic semantics, our typology of changes allows to avoid the commitment that such frameworks (and the theories expressed relative to them) are non-intertranslatable because their shared terms diverge referentially. Since external referents are not central to our cognitivist account, the communicative challenges

that proponents of pre- and post-revolutionary theories might face do consequently *not* arise from referential divergence. Such divergence, therefore, does not pose a threat to communicative rationality. To follow Kuhn's main example, even if the meaning of *mass* is different in Newtonian mechanics and Einstein's relativity theory, it does not follow that the two meanings cannot be compared, since the geometric and topological properties of the Newtonian and Einsteinian frameworks still have a considerable overlap. Nor does it follow, provided such overlap but without referential stability, that proponents of successive paradigms must engage in broadly non-rational forms of communication.

Put differently, even if communicative difficulties are expectable in periods of deep conceptual change<sup>5</sup>, these impediments may be alleviated by focusing on the structural representation of concepts in conceptual spaces instead of their possible referents. Historically, it seems that scientists' communicative difficulties do not pertain to the role of the central concepts in scientific reasoning, but rather to unrelated factors such as their ontological commitments or research traditions.

We believe that an understanding of conceptual frameworks as dimensionally structured spaces also suffices to explain cases of successful and rational communication between scientists. Our view nevertheless remains at odds with Michael Friedman's account who postulates various explanatory levels to

---

<sup>5</sup> Rudolph Carnap, Kuhn's editor for *Structure* (see Reisch, 1991), for instance, states:

"I have found that most scientists and philosophers are willing to discuss a new assertion, if it is formulated in the customary conceptual framework; but it seems very difficult to most of them even to consider and discuss new concepts." (Schilpp, 1963, 77)

render cases of severe conceptual change as communicatively rational. We now turn to a comparison with his perspective.

### **5. Michael Friedman's Neo-Kantianism**

Having outlined conceptual spaces as a non-linguistic model of conceptual knowledge, we now turn to a presentation of the central tenets of Michael Friedman's influential Neo-Kantian account. It presents a modern version of Kantian epistemology that intends to render scientific revolutions as "communicatively rational" instances of scientific development. Friedman borrows this term from Jürgen Habermas's (1984, 1987) theory of communicative action where it broadly signifies the "non-coercively uniting, consensus creating power of argumentative speech" (Habermas, cited after Friedman, 2001, 54).

Communicative rationality here contrasts with instrumental rationality. The latter governs the strategic choice among extant means, options, or tools, in order to satisfy, or optimize on, goals, purposes, or criteria, broadly conceived. Applied to the scientific case, for instance, instrumental rationality pertains to choosing among a range of available candidates that theory which is simpler, more empirically adequate, more fruitful, has greater scope, etc. A theory thus selected standardly constitutes an instrumentally rational choice with respect to the relevant goal or criterion, *only if* it is not dominated by another theory (maximization of choice utility).

Communicative rationality, on the other hand, pertains—or, as the case may be, fails to pertain—not to the selection among *extant* means or tools, but to the communicative process that makes "the serious consideration of [a theory that only

arises with] the *new* paradigm a rational and responsible option” (Friedman, 2002a, 190, italics added; see his 2010, 714). So communicative rationality refers to the rational quality of the interaction—itsself antecedent to theory choice—that renders the new paradigm (e.g., relativistic mechanics) acceptable as a strategically rational option, and more particularly as an *additional option* for those scientists (e.g., the “Newtonians”) who had formerly adhered strictly to the old paradigm only.

An immediate consequence of Friedman’s view is that coercive or otherwise broadly non-rational forms of communication and persuasion need not be postulated. For whenever such a process can in fact be communicatively rational in the sense of instantiating a form of non-coercive rational persuasion, then it makes little sense to suggest, as Kuhn (1970, ch. IX) did,<sup>6</sup> that the conversion of “old paradigmers” be a matter of *merely assent-directed* persuasion, thereby leveling the distinction between rational consensus and mere consensus.<sup>7</sup>

This being the purpose for which Friedman employs his Neo-Kantian account, we now turn to his notion of the relativized a priori (Sect. 5.1) and the role of philosophical meta-paradigms in deep conceptual change (Sect. 5.2).

---

<sup>6</sup> Addressing the “genetic aspect of the parallel between political and scientific development” that “should no longer be open to doubt,” Kuhn writes:

“As in political revolutions, so in paradigm choice—there is no standard higher than the assent of the relevant community. To discover how scientific revolutions are effected, we shall therefore have to examine not only the impact of nature and of logic, but also the techniques of persuasive argumentation effective within the quite special groups that constitute the community of scientists.” (Kuhn, 1970, 94)

<sup>7</sup> Hence, securing the communicative rationality of paradigm shifts is also an attempt at criticizing the relativistic inclinations of the strong program in the sociology of scientific knowledge (“SSK program”) (Rehg, 2009, 14–80; see Friedman, 1998).

### 5.1 *The relativized a priori*

Immanuel Kant had declared Newtonian time and space as *a priori forms* of human thinking, particularly an absolute space with Euclidian metric, and Newton's notion of absolute time. Because these forms make measurements in the empirical world possible, they are treated as *constitutive* of empirical knowledge. The a priori forms can also be expressed linguistically as *principles*. They then become, for instance, “the Newtonian Laws of Motion in the context of the *Principia*, the light principle and the principle of relativity in the context of [Einsteinian] special relativity, [or] the light principle and the principle of equivalence in the context of general relativity ...” (Friedman, 2010, 713).<sup>8</sup> The a priori forms and their corresponding principles are taken to enable the kind of knowledge that empirical theories express by means of special laws such as Newton's law of gravitation (or later Einstein's field equations)—knowledge that serves in the prediction of natural phenomena. The a priori principles of Newton's *absolute space* and *absolute time*, as Friedman stresses, changed in the transition from Newtonian to relativistic physics at the turn of the 20<sup>th</sup> century, enabled by prior historical developments in mathematics and geometry and culminating in Einstein's theory of relativity. Following Hans

---

<sup>8</sup> Newton's laws of motions standardly comprise *inertia*,  $F=ma$  and, on many accounts, also *actio=reactio*. For general and special relativity, the *light principle* dictates that the speed of light be constant for all observers irrespective of their motion relative to the light's origin. In special relativity, the *principle of relativity* says that the laws of physics be the same for all objects moving in inertial reference frames (i.e. those of constant velocity). In general relativity, the *principle of equivalence* expresses that “bodies affected only by gravitation follow geodesics in a variably curved space-time geometry” (Friedman, 2002a, 187).

Reichenbach (1920), who called them “axioms of coordination” (see Padovani, 2011), Friedman treats such principles as *methodological* a priori propositions. Serving the same function as they did for Kant, the principles remain constitutive of experience but are not forever fixed, and so can be altered as science develops. Relativized to historical eras, then, a principle comes to reflect the systematic stage of scientific theorizing reached at a historical moment, rather than an inbuilt or otherwise pre-determined constraint on the cognitive apparatus of theory-users.

Once relativized, Friedman argues, a Neo-Kantian version of the a priori may be retained, while Kant’s original one must be abandoned. This Neo-Kantian version is assigned with a significant task in the explanation of radical conceptual change in the transition from Newtonian to Einsteinian physics.<sup>9</sup> According to him, the explicit acknowledgement of philosophical meta-paradigms—including Kant’s own, in which a priori forms of thinking have their modern historical origin—is necessary and, in combination with other elements of Friedman’s account, also sufficient to successfully address the “Kuhnian challenge.” The challenge, according to Friedman, is to make understandable how a rational form of mutual inter-paradigmatic communication between proponents of rivaling scientific paradigms, old and new, can be achieved when these paradigms are *incommensurable* in the

---

<sup>9</sup> “[W]hat I call the dynamics of reason is not intended to be a general theory of scientific change at all—rather, it is a particular historical narrative accompanied by a particular philosophical gloss. The point, on my view, is that the transition from Newton to Einstein is the most important challenge to the Enlightenment ideal of scientific rationality championed by Kant . . . , and I am attempting to respond to this challenge, accordingly, by reexamining this particular transition in considerably more historical detail.” (Friedman, 2010, 714)

sense of relying on non-intertranslatable conceptual frameworks, as explained above.

Friedman's response to Kuhn's challenge particularly arises from the need "to explain, prospectively, how the new framework becomes rational, a 'live' option" (Kindi, 2011, 337) for those who adhere to the old paradigm. This notion of prospective rationality, holds Friedman, is distinct from the rationality afforded by a standard account of *inter-theory reduction* (see Batterman, 2012). Here, the set of models,  $M$ , that are provided by a theory of the old paradigm,  $T$ , are (approximately) reduced to limiting cases of particular models from the set  $M^*$  provided by a theory  $T^*$  that belongs to the new paradigm. The models of  $T$  are thus (approximately) likened to those of  $T^*$ .<sup>10</sup> For Friedman, however, a convergence of mathematical structures through inter-theory reduction can at most demonstrate that a paradigm shift is *retrospectively rational*. That is, the change from the old to the new theory is but unilaterally rational, namely from the point of view of the new paradigm. From the point of view of the old paradigm, however, an analogue to inter-theory reduction, one that would similarly render the transition to the new paradigm as a rational instance of conceptual development, is unavailable. Moreover, ontological or referential differences, as we have seen, provide Kuhn with the primary reason to attest a divergence of meaning between (the terms shared by) incommensurable

---

<sup>10</sup> In relativistic contexts, for instance, *momentum* is expressed as  $p=m_0v/\sqrt{1-(v/c)^2}$ , where  $v$  is the velocity and  $c$  the speed of light. The special relativity form converges to the Newtonian  $p=mv$  as  $v$  goes to zero. So the relativistic form reduces the Newtonian one, because  $\sqrt{1-(v/c)^2}$  approaches 1 as  $(v/c)^2$  approaches 0.



conceptual frameworks.<sup>11</sup> And such referential meaning divergence remains unaddressed by limiting case reduction. Friedman essentially accepts referential divergence as capturing “a centrally important aspect of what he [Kuhn] has called the non-intertranslatability or ‘incommensurability’ of pre-revolutionary and post-revolutionary theories” (Friedman, 2002a, 186, n.14). In particular, that such referential divergences arise in paradigm shifts would seem to make it more understandable that proponents of the old and the new paradigm face communicative difficulties—for they may find themselves “living in different worlds,” worlds populated by different things—and so would render it more understandable that they allegedly have to *sway* one another into the adoption of one or the other paradigm.

## 5.2 *The role of philosophical meta-frameworks*

Friedman’s form of Neo-Kantianism takes issue with the holistic picture, standardly ascribed—as the Duhem-Quine thesis—to Pierre Duhem and Willard Van Orman Quine (see Brenner, 1990), according to which our “web of belief” faces the tribunal of experience *in toto* (Duhem), so that each of its elements is as prone to modification as any other (Quine). Friedman instead proposes “an alternative picture

---

<sup>11</sup> Friedman places Kuhn within Cassirer’s *genetic conception* that sees scientific knowledge to “progress from naively realistic ‘substantialistic’ conceptions, focusing on underlying substances, causes, and mechanisms subsisting behind the observable phenomena, to increasingly abstract purely ‘functional’ conceptions” (2008, 244). But both Kuhn’s notion of paradigm shift in *Structure* and his subsequent replacement for paradigm, the *structured lexicon*, also draw on diverging ontologies, or referent-shifts. Kuhn’s position, Friedman argues, therefore also reflects the Meyersonian substantialistic view, although it is directly opposed to Cassirer’s.

of a thoroughly dynamical yet nonetheless differentiated system of knowledge that can be analyzed ... into three main components or levels” (Friedman, 2002a, 189). The first level houses the component that faces experience directly, namely the “empirical laws of nature, such as the Newtonian law of gravitation or Einstein’s equations for the gravitational field” (ibid.). As we have seen, in our diction such laws or equations correspond to specific constraints on the distribution of points over a conceptual space.

The second level comprises the “constitutively a priori principles that define the fundamental spatio-temporal framework within which alone the rigorous formulation of first or base level principles is possible” (ibid.). From our perspective, second level principles correspond to (parts of) the conceptual space itself. For instance, principles expressing Newton’s integral 3D space with its Euclidian metric, as well as Newton’s separate 1D time or Einstein’s integral 4D space-time with its non-Euclidian (Schwarzschild) metric explicitly count as a priori on the second level for Friedman. In contrast, Newton’s separate 1D mass and the integral 3D force—being magnitudes that to successfully apply in predicting natural phenomena presupposes having assigned values on the space and time dimensions—may appear to count as being a priori at most implicitly. But Friedman goes on to claim that “[t]hese relativized a priori principles [at level two] constitute what Kuhn calls paradigms: relatively stable sets of rules of the game, as it were, that make possible the problem-solving activities of normal science [all of which involve the adoption, maintenance, or rejection of laws at level one]...” (ibid.), thus indicating that principles governing mass and force might count as a priori, too. After all, in order to predict a gravitational phenomenon such as a planetary orbit, for instance,

one of the Newtonian “rules” of normal science prescribes that one enrich models that already feature values on the time and the space dimensions by stipulating suitable values for the dimensions of mass and force.<sup>12</sup> Friedman’s second level would then comprise the entire conceptual framework or, in our diction, the conceptual space spanned by the dimensions with their metrics as these combine into domains.<sup>13</sup> This excludes from level two only the constraints on the distribution of points otherwise known as empirical laws, and located at his level one.

The second level provides Friedman with the necessary background so that the adoption, maintenance, and rejection of laws at level one—in the course of Kuhn’s “normal science”—can be understood as an empirical matter, that is, as a matter of methodologically-hardened<sup>14</sup> corroboration or falsification. On the other

---

<sup>12</sup> In structuralist terms, the process is that of enriching a partial potential model of Newtonian mechanics to a full model of the theory. The structuralist approach is compared to the conceptual spaces approach in Gärdenfors & Zenker (2011) and Zenker & Gärdenfors (2013) where it is shown, among others, that the structuralist’s three kinds of models—potential, partial potential, and full model—can be provided with spatial analogues. In each case, respectively, these models are understood as ever more restrictive constraints on the distribution of points over the space.

<sup>13</sup> Friedman perhaps comes closest to such distinctions when he reports the early Carnap as having offered “a generalization of Kant’s conception of spatial intuition according to which only the infinitesimally Euclidean character of physical space is a priori determined by the form of our intuition ..., whereas the choice of specifically ‘metrical form’ (whether Euclidean or non-Euclidean) is ‘optional [wahlfrei]’” (2002b, 24). Compare our second type of change (Sect. 3).

<sup>14</sup> In criticism of Karl Popper’s falsificationism, Imre Lakatos (1978) introduced the notion of a *methodologically hardened fact*. Unlike *naive falsificationists*, who treat empirical data (“facts”) as indubitable—so that facts always “win” in case of a conflict between theory and observation—and also unlike *methodological*

hand, the adoption, maintenance, or rejection of level-two-conceptual frameworks or paradigms—in the course of Kuhn’s “revolutionary science”—cannot be handled as empirical processes in this sense. Friedman writes that,

“no straightforward process of empirical testing [of the paradigm-constitutive a priori principles at level two], in periods of deep conceptual revolution, is then possible. Here our third level, that of philosophical meta-paradigms or meta-frameworks, plays an indispensable role, by serving as a source of guidance and orientation in motivating and sustaining the transition from one paradigm or conceptual framework to another. Such philosophical meta-frameworks contribute to the rationality of revolutionary change, more specifically, by providing a basis for mutual communication (and thus for communicative rationality in Habermas’s sense) between otherwise incommensurable (and therefore non-intertranslatable) scientific paradigms.”  
(Friedman, 2002a, 189)

The role that Friedman envisions for philosophical meta-paradigms or meta-frameworks is that of “guiding the articulation of the new space of possibilities [delivered by the successor paradigm] and making the serious consideration of the *falsificationists*, who treat as indubitable the observational theory that yields these facts, *sophisticated methodological falsificationists* “harden” their facts by presupposing a hierarchy of observational theories. So it is only the acceptance of an order of auxiliary theories that enables the theory to “lose” against experience (*falsification*) or to confirm it (*corroboration*). Conversely, when the predicting theory shall be maintained for some reason, then the facts may be “softened” by doubting the observational theory on which they rely (see Zenker, 2009, ch. 4).

new paradigm [by those committed to the old one] a rational and responsible option” (Friedman, 2002a, 190). Like the second level, also the third functions in the spirit of a Kantian program, a program that seeks to explicate *conditions of possibility*. Among these meta-paradigms, for instance, Friedman counts the “new approach to the understanding of nature self-consciously crafted by Descartes and Galileo against the backdrop of medieval Scholasticism” (Friedman, 2001, 23), and of course Einstein’s “recognition of a new item, as it were, in the space of intellectual possibilities: namely the possibility of a relativized conception of time and simultaneity” (ibid.). A meta-paradigm shall deliver “new conceptions of what a coherent rational understanding of nature might amount to” (ibid.). It is, in brief, “a source of new ideas, alternative programs, and expanded possibilities that is not itself scientific in the same sense—that does not, as do the sciences themselves, operate within a generally agreed upon framework of taken for granted rules” (ibid.)

To summarize, the meta-paradigms at level three make possible the adoption, maintenance, and revision of a priori principles at level two; and the level two a priori principles make possible the adoption, maintenance, and revision of empirical laws at level one.

Friedman’s position is thus that philosophical discourse is necessary for communicatively rational scientific progress in the face of radical conceptual change. The third level, in particular, is seen as *indispensable* in this process. So a consequence of Friedman’s view seems to be this: when scientists each prefer (radically) different conceptual frameworks, but nevertheless manage to engage in rational communication about them, then they must at least to some extent be well-versed in meta-philosophical discourse. We want to show that an alternative

response to Kuhn's challenge is possible. One reason to look for an alternative response is that there are socio-empirical reasons to doubt Friedman's account: By and large, scientists are not well-versed in philosophical discourse, and it is not clear either that they need to be. If so, then the scientists' interaction is for the most part not communicatively rational in the sense that Friedman adopts from Habermas, or successful communication about diverging conceptual frameworks can be facilitated by something other than meta-philosophical discourse. Instead, we suggest that the theory of conceptual spaces explains the communicative success, without committing to the third level of Friedman's three-tiered neo-Kantian view. Our aim in the next section, then, will be to argue that it is not necessary to deploy the entirety of philosophical discourse in order to explain that radical changes to a conceptual framework can be communicatively rational. If we are right, then the theory of conceptual spaces provides an alternative response to the Kuhnian challenge.

## **6. Communication about Theoretical Frameworks as Conceptual Spaces**

It is crucial for our argument to recognize that, like Kuhn's (1970; 1987; 2000), also Friedman's account remains grounded in a tradition that understands conceptual frameworks as *linguistic entities*, while ours models these entities as (abstract) spaces. As we have seen, one may understand genuinely empirical laws as symbolic expressions of constraints on the distribution of points over the space spanned by its dimensions. This was perhaps clearest for cases such as Newton's law of gravitation that predicts these points to lie on the hyper-surface described by  $F=GMm/r^2$  (Gärdenfors & Zenker, 2013). Recall further that the law restricts these points to a

subspace of that described by Newton's second law,  $F=ma$ , which is normally treated as an axiom of the theory. In brief, on our account a theory consists of a conceptual space (the framework) together with constraints on empirically possible points in the space (that can be expressed as laws).

This spatial view extends also to propositions such as Einstein's light principle that, for Friedman, has a priori status (see DiSalle, 2002, 196f.). To illustrate, the principle says that the speed of light,  $c$ , is constant in all inertial reference frames, where  $c$  can be dimensionally analyzed as  $[LT^{-1}]$ , i.e., a length  $[L]$  over a time interval  $[T]$ . Interpreted in terms of conceptual spaces, the light principle restricts  $c$  to a single value, given by the quotient of the dimensions  $[L]$  and  $[T]$ , of 299,792,458 meters/second. In relativistic treatments of 4D space-time, moreover, 3D space and 1D time are integral dimensions and so form a domain. This contrasts with the Newtonian theory where 3D space and 1D time constitute separate domains, gravity is a force acting-at-a-distance, and the speed of light an *additive* magnitude (rather than bound to  $c$ ). And it similarly contrasts with a Cartesian view where, being unrestricted,  $c$  marks infinity—giving rise to an interpretation of light signals as propagating instantaneously, and so coordinating distant events to local ones.

When a conceptual framework is analyzed as a spatial structure, the empirical laws at Friedman's first level are demoted in status because, as we have seen, they are merely the linguistic means to express constraints on a conceptual space. And the same holds for the linguistic expressions known as a priori principles or axioms at the second level. After all, the information that linguistic approaches take to be expressed by such principles may partly be read off directly from the

structure of the space, and so is generated by the combination of dimensions and their metric.<sup>15</sup> And it can partly be read off by interpreting such principles to constrain the space, so that predictions warranted by a theory's empirical content fall within the space's hyper-planes. Consequently, a new conceptual framework,  $F^*$ , with its embedded theory  $T^*$ , that *arise together* as a new paradigm in the course of Kuhn's revolutionary science, can be described by applying the five change operations (Section 3) to the predecessor framework,  $F$ —both  $F$  and  $F^*$  being reconstituted as conceptual spaces. And the set of models comprising the empirical content of the new theory,  $T^*$ , that is formulated against the background of  $F^*$  may similarly be understood as a distinct set of constraints.

The linguistic approach that is used by both Kuhn and Friedman identifies such changes in terms of a replacement of sentences, rather than describing the changes to a priori principles, axioms, and empirical laws of  $T$  into those of  $T^*$  that take place in the transition from the conceptual framework  $F$  to the successor  $F^*$  as being *determined* by a structural modification of  $F$ . Furthermore, there is no evidence in Friedman's writings of a principled distinction between a conceptual framework, on the one hand, and an empirical theory that is expressed relative to that framework, on the other. (In our model, theories are defined by *constraints* on the conceptual space of a theory framework.) For instance, Friedman writes that

---

<sup>15</sup> To give an elementary example, the fact that relations on dimension such as "longer than" and "warmer than" are *transitive* follows immediately from their one-dimensional structure, being isomorphic with the real line. Thus the transitivity need not be formulated as an axiom, as is standard in most linguistic or logical accounts of theories, but is inherent in the dimensional structure of the underlying conceptual space.



“[t]he key move in general relativity ... is to *replace* the law of inertia—which, from the space-time perspective inaugurated by special relativity, depicts the trajectories of force-free bodies as geodesics in a flat space-time geometry—with the principle of equivalence, according to which bodies affected only by gravitation follow geodesics in a variably curved space-time geometry.” (Friedman, 2002, 187, *italics added*)

We take such passages as evidence that both the empirical theory and the conceptual framework that it presupposes are primarily understood as linguistic entities, so that an analysis of their dynamics will need to rely on change operations such as those described in the AGM tradition (Alchourrón, Gärdenfors & Makinson, 1985; Gärdenfors, 1988).<sup>16</sup> Here, for instance, the replacement or revision of proposition *p* by *q* is modeled as the retraction of *p* followed by the addition of *q*; the various propositions of a theory, moreover, are ordered according to levels of entrenchment (Zenker, 2009).

---

<sup>16</sup> We take Friedman’s employment of the terms “succession,” “transformations” and “extension” to reflect the same understanding:

“I have argued, on the one hand, that the transition from Newton to Einstein centrally involves a succession of relativized constitutively a priori principles ..., and the existence of such diverse constitutively a priori principles, on my view, captures the essence of Kuhnian incommensurability. But I have also suggested, on the other hand, that the detailed historical route from earlier to later constitutive principles exhibits the latter as natural transformations of the former—arising as a sequence of ‘minimal extensions’ of our Kantian-Newtonian starting point in a succession of new mathematical, empirical, and philosophical situations.” (Friedman, 2010, 713)

We next turn to the consequences of adopting the spatial perspective on theories and their frameworks for scientific communication. As we have seen, a model of  $T$  is a constrained hyperspace of  $F$  which itself is a spatial object constituted by dimensions, their metrics, and the way these dimensions form domains. Such spatial objects, we submit, form the basis for communication between scientists during periods of what Friedman calls “deep conceptual change.” These spatial objects allow for identifying some or another set of dimensions and constraints as being methodological a priori—making it expectable that scientists display diverging preferences as to which dimensions in fact are fundamental in this sense—because none of these dimensions count as inherently privileged. For instance, in electrodynamics, particle theorists tend to view *electrical current* as a fundamental dimension, while field theorists take *electrical charge* as fundamental (Gärdenfors & Zenker, 2013). And just like Joseph Sneed’s (1971) distinction between *T-theoretical* and *T-non theoretical terms*, on which structuralists rely (Gärdenfors & Zenker, 2011), also the Neo-Kantian distinction between *a priori principles* and *empirical laws* remains perfectly possible in the context of conceptual spaces.

Our critical objection, however, is that neither such distinctions nor intimate knowledge of the philosophical traditions from which they arise appear to be *indispensable* for communication to remain rational in the relevant sense. On a cognitive account of meaning and a representation of theories with the aid of conceptual spaces, rather, communicative rationality appears to be available prior to Friedman’s third level. So, one *may* assign a priori status to some but not other dimensions of an empirical theory’s conceptual framework. But the same appears

not to be necessary in order to explain that rational communication between proponents of different paradigms, about these differences, is possible during periods of deep conceptual change.

Let us spell out our position a bit more. We do not claim that, when successfully communicating, scientists consciously entertain the theory of conceptual spaces as a shared background tool. And we neither claim that, if scientists were to entertain—consciously or not—the theory of conceptual spaces, they would never disagree about issues of theory choice for, as it were, the need to persuade, and be persuaded, never arises. Rather, our claim is that the theory of conceptual spaces can explain scientists' communicative successes—particularly those related to issues arising inter-paradigmatically—without postulating meta-paradigmatic philosophical abilities on their behalf. Our view moreover entails that persuasive attempts at bringing an opponent to adopt a scientific framework she dis-prefers need not, *ab initio*, be outside the realm of rationally reconstructable discourse. This is in contrast to the Kuhnian view we criticize.

At the same time, our account offers little in the way of a positive characterization, or general guidance on, how scientists *should* rationally persuade one another. Analysts studying actual scientific discourse may nevertheless find it worth considering that inter- and intra-paradigmatic communication failure and success alike are not readily explainable by citing the discourse participants' differentially pronounced meta-philosophical abilities. Applying the theory of conceptual spaces also for the purpose of describing scientific discourse will make it easier to provide good accounts of the communicative phenomena.

## 7. Conclusion

Philosophy of science has put itself in a deadlock by representing theories as linguistic entities (symbolically expressed systems of laws) and by assuming semantic realism. This leads to problems understanding how scientists can communicate rationally in times of radical theory change. The deadlock witnessed, among others, Kuhn's claim that paradigms are incommensurable. This claim ultimately accepts, as an *explanandum*, what is in fact a reconstructively incurred artifact owed to Kuhn's own theoretical perspective, which is a linguistic one. As a remedy, Friedman introduces meta-philosophical discussions as a way of accounting for how scientists can communicate in a rational way. However, scientists would generally not see the need for such discussions to resolve the conflict between two competing theories, even if the frameworks of the theories are different.

In contrast, if theory frameworks are modeled as conceptual spaces and theories as constraints on such spaces, the communicative challenge will become smaller. Frameworks have to be compared in scientific discussions, but this can be handled by comparing their geometric or topological properties, instead of relying on meta-philosophical considerations. And once theory frameworks are thus comparable, the incommensurability aspect of conceptual change collapses into referential divergence, or ontological change. In brief, a discussion conducted in terms of meta-philosophical principles won't add much towards guiding "the articulation of the new space of possibilities" and making "the serious consideration of the new paradigm a rational and responsible option" (Friedman, 2002, 190).

We have provided reasons to doubt that addressing radical meaning changes in adequate ways requires, or makes indispensable, the meta-philosophical principles

at the third level of Friedman’s account. Instead, we have argued, if mutual inter-paradigmatic communication between scientists shall be saved—so that, also from the point of view of the old paradigm, the new conceptual framework can be seen as rationally superior to the former (Kindi, 2011, 337)—then a reconstruction of the successive conceptual frameworks as two conceptual spaces already allows for such mutual inter-paradigmatic communication. So Friedman’s three-tiered method of handling Kuhn’s challenge can be replaced by a cognitively and communicatively more economical account.

In conclusion, we find that Friedman’s neo-Kantian perspective has some advantages over a sentence-based account (such as logical positivism) or Kuhn’s position. Nevertheless, his third level has little role to play in actual scientific discourse; for most purposes that are of concern to scientists, it is sufficient to rely on a comparison between the spatial frameworks of the theories under scrutiny.

### **Acknowledgements**

We thank two anonymous reviewers for critical comments that improved an earlier version of this text, as well as the audience at the conference “Conceptual Spaces at Work,” held 26-26 May, 2012 at Lund University. Both authors acknowledge funding from the Swedish Research Council.

### **References**

- Alchourrón, C., Gärdenfors, P., and Makinson, D. (1985). On the Logic of Theory Change: Partial Meet Contraction and Revision Functions. *Journal of Symbolic Logic*, 50: 510–530.

- Batterman, R. (2012). Intertheory Relations in Physics. In Zalta, E. N. (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2012 Edition). <http://plato.stanford.edu/archives/fall2012/entries/physics-interrelate>. (Accessed 23 JAN 2014)
- Brenner, A. A. (1990). Holism a century ago. The elaboration of Duhem's Thesis. *Synthese*, 83: 325–335.
- D'Agostino, F. (2013). Verballed? Incommensurability 50 years on. *Synthese* (online first). DOI 10.1007/s11229-013-0288-y
- DiSalle, R. (2002). Reconsidering Kant, Friedman, Logical Positivism, and the Exact Sciences. *Philosophy of Science*, 69:192–211.
- Einstein, A. (1924). Review of Elsbach 1924. *Deutsche Literaturzeitung*, 45:1685–1692.
- Friedman, M. (1998). On the sociology of scientific knowledge and its philosophical agenda. *Studies in History and Philosophy of Science*, 29(2): 239–271.
- Friedman, M. (2001). *Dynamics of Reason*. Stanford, Ca.: CSLI Publications.
- Friedman, M. (2002a). Kant, Kuhn, and the rationality of science. *Philosophy of Science*, 69:171–190.
- Friedman, M. (2002b). Kuhn and Logical Empiricism. In Nickles, T. (Ed.), *Thomas Kuhn* (19–44). Cambridge, UK: Cambridge University Press.
- Friedman, M. (2008). Ernst Cassirer and Thomas Kuhn: The neo-Kantian tradition in history and philosophy of science. *Philosophical Forum*, 39:239–252.
- Friedman, M. (2010). Synthetic History Reconsidered. In Domski, M., and Dickson, M. (Eds.), *Discourse on a New Method. Reinvigorating the Marriage of*

- History and Philosophy of Science (571–813)*. Chicago and La Salle, Ill.: Open Court.
- Gärdenfors, P. (1988). *Knowledge in Flux*. Boston, Mass.: The MIT Press.
- Gärdenfors, P. (2000). *Conceptual Spaces: The Geometry of Thought*. Cambridge, Mass.: MIT Press.
- Gärdenfors, P. (2014). *The Geometry of Meaning: Semantics Based on Conceptual Spaces*. Cambridge, Mass.: MIT Press.
- Gärdenfors, P. and Zenker, F. (2011). Using conceptual spaces to model the dynamics of empirical theories. In Olsson, E.J., and Enqvist, S. (Eds), *Philosophy of Science Meets Belief Revision Theory (137–153)*. Berlin: Springer.
- Gärdenfors, P. and Zenker, F. (2013). Theory change as dimensional change: Conceptual spaces applied to the dynamics of empirical theories. *Synthese*, 190(6): 1039–1058.
- Habermas, J. (1984) [1981]. *Theory of Communicative Action Volume One: Reason and the Rationalization of Society* (Translated by T. A. McCarthy). Boston, Mass.: Beacon Press.
- Habermas, J. (1987) [1981]. *Theory of Communicative Action Volume Two: Liveworld and System: A Critique of Functionalist Reason* (Translated by T. A. McCarthy). Boston, Mass.: Beacon Press.
- Howard, D. (2010). ‘Let me briefly indicate why I do not find this standpoint natural.’ Einstein, General Relativity, and the Contingent A Priori. In Domski, M., and Dickson, M. (Eds.), *Discourse on a New Method*.

- Reinvigorating the Marriage of History and Philosophy of Science* (333–355). Chicago and La Salle, Ill.: Open Court.
- Kindi, V. (2011). The challenge of scientific revolutions: Van Fraassen's and Friedman's responses. *International Studies in the Philosophy of Science*, 25(4): 327–349.
- Kuhn, T. (1970) [1962]. *The Structure of Scientific Revolutions*, 2<sup>nd</sup> Ed. Chicago, Ill.: University of Chicago Press.
- Kuhn, T. (1987). What are Scientific Revolutions? In Krüger, L., Daston, L., and Heidelberger, M. (Eds.), *The Probabilistic Revolution Volume One* (7–22). Cambridge, Mass.: MIT Press.
- Kuhn, T. (2000). *The Road since Structure. Philosophical Essays, 1970–1993* (with an Autobiographical Interview, edited by Conant, J., and Haughland, J.). Chicago, Ill.: University of Chicago Press.
- Lakatos, I. (1978). *The Methodology of Scientific Research Programs*. Cambridge, UK: Cambridge University Press.
- Padovani, F. (2011). Relativizing the relativized a priori: Reichenbach's axioms of coordination divided. *Synthese*, 181(1):41–62.
- Petersen, G., and Zenker, F. (2014). From Euler to Navier-Stokes: The changing conceptual framework of 19<sup>th</sup> century fluid dynamics (submitted manuscript).
- Rehg, W. (2009). *Cogent Science in Context. The Science Wars, Argumentation Theory, and Habermas*. Cambridge, Mass.: The MIT Press.
- Reichenbach, H. (1920). *Relativitätstheorie und Erkenntnis apriori*. Berlin: Springer. Translated by Reichenbach, M. (1965). *The Theory of Relativity and A*



*Priori Knowledge*. Berkeley and Los Angeles, Ca.: University of California Press.

Reisch, G. A. (1991). Did Kuhn kill logical empiricism? *Philosophy of Science*, 58 (2): 264–277.

Schilpp, P. A. (Ed.) 1963. *The Philosophy of Rudolf Carnap*. La Salle, Ill.: Open Court.

Sneed, J. D. (1971). *The Logical Structure of Mathematical Physics*. Dordrecht: Reidel.

Zenker, F. (2009). *Ceteris Paribus in Conservative Belief Revision*. Berlin: Peter Lang.

Zenker, F. (2014). From Features via Frames to Spaces. Modeling Scientific Conceptual Change without Incommensurability or Apriority. In: Gamerschlag, T., Gerland, D., Osswald, R., and Petersen, W. (Eds.), *Frames and Concept Types: Applications in Language and Philosophy* (69–89). Dordrecht: Kluwer/Reidel.

Zenker, F., and Gärdenfors, P. (2013). Modeling diachronic changes in structuralism and in conceptual spaces. *Erkenntnis* (online first). DOI 10.1007/s10670-013-9582-9.

**Word count: 8620 (incl. a 750 word reference section)**