

## HOPOS and the Paradigm Shifts of Philosophy

[paper for panel on “HOPOS: Methods and Contexts” – HOPOS 2014, Ghent, 2-6 July, 2014]

**Menachem Fisch**

In at least one important respect, the nature of modern philosophy of science changed markedly after Newton. If Bacon and Descartes could be described as radical calling to found science anew on solid philosophical footing, after Newton, philosophy of science was rendered by and large an interpretative undertaking: attempting, not to properly establish science, but to make philosophical sense of science as it was in fact practiced.

To understand a move in post-Newtonian philosophy of science *historically*, requires, therefore, to view it as an attempt to make philosophical sense of some aspect of the science of its day, in critical dialogue with the opposition; a two-tier effort:

To explain first why the thinker in question might have found the available philosophical options wrong or wanting, and subsequently to explain why he considered his own suggestions better suited to meet those challenges.

Put differently, good HOPOS work should purport not only to describe or render a past philosopher’s move explicit, but to expose its author’s reasons for making it. Or in Pittsburghese: it should aim at contextualizing the philosophies it studies within the historical spaces of reasons in which they were initially formed and later defended – which brings me to the question I really wish to address.

Just as major developments in science after Newton constituted framework transitions, so did major developments in philosophy of science and mathematics. And framework transitions, needless to say, pose an exceptionally difficult challenge for this kind of explanation. Paradoxically, the fact that HOPOS has remained largely indifferent to the problem of accounting for reasoned paradigm shifts in the *philosophy* of science owes, I believe, to the most significant recent attempt to account for the rationality of such shifts in *science* itself, namely, Michael Friedman’s *Dynamics of Reason*, which more than suggests that the problem is exclusive to science, and should not concern historians of other disciplines, especially philosophy, where, he implies, thinking out of the box is a matter of course. I would like to first briefly explain the problem, take critical stock of

Michael's solution, reopen it for the history of philosophy, and then outline a solution better tailored, I believe, to the concerns of HOPOS.

### **The Problem**

To act *rationally*, as I shall use the term, is to act *for a reason*. To have reason to act is to deem something sufficiently lacking to justify changing it; to have reason not to act, is to deem it preferable to the alternatives. Hence, as Popper insisted, the crude rational force of faults and failings that renders problem-seeking and solving definitive of rationality.

However, a deeper element of critical *self*-scrutiny is inherently involved in our very capacity to reason. Christine Korsgaard makes the point vividly

our capacity to turn our attention on our own mental activities is also a capacity to distance ourselves from them, and to call them into question. I perceive, and I find myself with a powerful impulse to believe. ... [But] Is this perception really a *reason* to believe? I desire and I find myself with a powerful impulse to act. ... Is this desire really a *reason* to act? The reflective mind cannot settle for perception and desire.... It needs a *reason*. Otherwise... it cannot commit itself or go forward. (SN 93)

Korsgaard thus locates the defining feature of rational agency in the *intrapersonal* realm of self-deliberating which sensations and impulses merit endorsement as reasons to believe and act. But thinkers like Sellars and Brandom insist on giving Korsgaard's neo-Kantian picture a sharp neo-Hegelian twist, as it were, locating rationality's defining feature in the *interpersonal* domain, where people engage in the game of giving and asking for reasons and keep deontic score of each other's commitments and entitlements.

Korsgaard is right to insist that self-constitution is exclusively *intrasubjective*. No one but oneself can endorse one's reasons to act, and be committed to one's norms. No one but oneself can oppose certain of one's desires and impulses and identify with others. To assume agency is to stand in critical self-judgment, wary of the persistent pull of our sensations and impulses, just as she describes. Only then, are we in a position to rationally exercise our agency in action, which, as noted, is to adopt a similarly critical stance toward our world and picture of it. And in subsequently engaging in the game of giving and asking for reasons, a third and higher form of reflective wariness is added, by subjecting our first-order Korsgaardian choices to the critical reflection of others.

Combined, the three components of human agency – endorsing reasons, applying them in

action, and exposing them to the critical appraisal of others – display our rationality as thoroughly steeped in self-criticism, just as critical rationalists insist. It also explains why Brandom and others deem the third component - participation in the intersubjective dialogue of reasons - as the very epitome of human rationality. For in doing so, we *apply our rationality to itself*; as it were, enlisting others to reflect critically on our own attempts at critical reflection. It is here, claims Brandom, that the “I” becomes “we”, as the self awakens to its social discursive embeddedness. But none of this is simple.

To criticize, i.e. to deem something wanting and in need of repair, is to make a *value* judgment. For to act for a reason is to take stock of what *is* the case and act according to what one believes it *ought* to be – all of which necessarily presupposes reference to a system of relevant values or criteria of propriety. To criticize is to pass normative judgment, for which a normative framework must already be in place. Criticism, in other words, is inherently framework-dependent in the strong sense of the term – and, yes, one can hear the Popperians cringing.

At its intrasubjective Korsgaardian level of endorsing reasons and applying them in action, the normative dependency of rationality is understandable and unproblematic. Deeming a desire or sensation worthy or unworthy of endorsement as a reason, and a state of affairs in need or not in need of correction, are clear cases of norm *application*. Conducting oneself rationally at this level is to hold oneself *accountable to one’s existing* norms and standards. But rationality requires of us, not only to apply our standards *in* critical appraising, but to subject them *to* critical appraisal. But how is that possible? How can the norms to which one is committed be effectively subjected to one’s own normative appraisal if it is by means of those very norms that one appraises?

Sellars and Brandom imply that this is what exposing our reasons and actions to others is supposed to achieve. But how? If one is incapable of critically appraising one’s own standards of appraisal, how could one be convinced to do so by others? Standards of propriety to which we are genuinely committed seem by definition to be immune to normative criticism. But that seems absurd! John McDowell makes the point forcefully with respect to ethics:

Like any thinking, ethical thinking is under a standing obligation to reflect about and criticize the standards by which, at any time, it takes itself to be governed. ... But the essential thing is that one can

reflect only from the midst of the way of thinking one is reflecting about.... The thought is that this application of one's ethical outlook would stand up to the outlook's own reflective scrutiny. (1994, 81)

But even McDowell, has failed to show how, from the midst of the way of thinking one is reflecting about, the normative failings of one's norms can in principle be disclosed by subjecting them to their "own reflective scrutiny."

Carnap, Kuhn, Rorty and Davidson argue more or less explicitly that they cannot. Others, notably Popper and some of his school, deny there being a problem by unfoundedly pronouncing framework dependency a "myth." But the vast majority, Sellars, Brandom and Korsgaard included, avoid the question by simply not raising it - and understandably, for the problem of normative self-criticism is real and notoriously difficult. The notable exception is Friedman.

### **Enter Friedman**

Faced by the revolutionary history of each of the modern sciences, philosophers of science could not evade the question of the rationality of framework transitions as easily as other philosophers - certainly not since Kuhn's vivid portrayal of science's framework-dependency; a picture that Friedman, his most prominent latter-day follower and critic, puts it: fiercely challenged "the ultimate rationality of the scientific enterprise". If in paradigm shifts the very criteria of scientific propriety are replaced, there seems, he writes, "no sense left in which such transitions can still be viewed ... as based on good reasons". But Friedman's alternative account is also problematic.

First, by opting for Habermas's notion of communicative rationality, Friedman limits himself in advance to cases of practitioners rationally engaging adherents to different *lebenswelten* from their own. The problem hence becomes that of rationally replacing one's framework by an already functioning alternative. The problem of being rationally motivated to initially *construct* such an alternative, is wholly ignored.

Second, Friedman purports narrowly to present a special case for the rationality of *scientific* framework transitions, which he argues are less radical than supposed. In science, he argues contestably, (a) "earlier frameworks are exhibited as limiting

cases...of later ones", and (b) "the concepts and principles of later paradigms...evolve continuously, by a series of natural transformations, from those of earlier ones" (63).

Thirdly, Friedman also insists that the new ideas that render the new unthinkable from the perspective of the old, all originate outside the scientific community – among philosophers, mathematicians and others whose thinking is not constrained by the reigning scientific paradigm.

Each of the three components of Friedman's solution has been forcefully contested. He himself seems to have had second thoughts, for in the bulky restatement of his position in 2010, Habermas is dropped, as is all talk of rationality, and the theory is further narrowed to apply exclusively to the history of space-time physics.

But my concern today is less with the viability of Friedman's thesis for science than with its worrying implications for HOPOS. In making his special case for science, and philosophers have no problem thinking out of the box, Friedman creates the impression that either the problem of normative self-criticism does not arise outside science, or that non-scientific framework transitions cannot be rational. But if there's any truth in what I've said so far, then since there is no such thing as reasoned action that is *not* framework dependent in the same strong sense premised by *Dynamics of Reason*, it follows that the problem of the rationality of framework transitions is as pressing and as resilient with regard to philosophy as it is with regard to science! And if so, we HOPOI are faced with the same concerns regarding the philosophy of science as those raised by Friedman with respect to science proper.

### **Creatively Ambivalent**

To repeat: the problem of rational framework transitions boils down, *pace* Friedman, to that of accounting for the possibility of normative *self*-criticism; a problem claimed or implied to be unsurmountable by virtually all who take framework dependency seriously. But I believe it can be solved, though not by the kind of intrasubjective self-distancing described by McDowell. Let me quickly outline the proposed solution and how I believe it applies to the historiography of both scientific and philosophical framework transitions.

Since normative self-criticism cannot be achieved by mere *self*-reflection, the key to its possible achievement must certainly lie in exposure to the normative critique of others, as Michael Walzer's work on social criticism, and that of Sellars and Brandom briefly mentioned, imply. But we need to tread carefully. For a framework transition to be rational, it must be normatively faulted *while still in play*, and only then abandoned for an arguably better alternative. Wittgensteinian 'persuasion', Rortian irony and Kuhnian gestalts and conversions, won't do because they reverse the process: frameworks are first replaced and only then rationalized in retrospect. But how is criticism leveled by others supposed to transcend the seemingly impenetrable glass ceiling of our own self-reflection? For this we need to take a closer look at how prudent criticism works.

Unlike doubting, criticism aims, not at *wondering*, but at *proving* to its addressees that something is seriously enough amiss to require their attention. Criticism thus purports not only to be understood, but to be *endorsed* by its addressees as *self*-criticism. Criticism must, therefore, be leveled from its addressees' perspective, doing its best to emulate *their* worldview and values, and mount its case on premises *they* hold true.

And the same applies to normative criticism. The problem is that when criticizing a person's norms it is impossible to frame such an argument from her perspective. For there is never available a set of premises she is liable to consider true, that can be shown, convincingly *for her*, to entail a denunciation of her very norms. But if such criticism can never convince, how can it ever have the desired transformative effect of rationally moving people to rethink their norms?

The key lies in realizing that prudent normative critics are to some extent aware of this. They know (or at least sense) that leveling their arguments squarely from within their addressee's worldview is impossible. Yet they also know that to make its mark their criticism must rest, as far as possible, on premises their addressee *can* recognize as his own. What we normally do, therefore – and this is the crucial point – is to frame our arguments to some extent *untruthfully*, arguing from a perspective as close as possible to that of our addressee, yet sufficiently different from it to be able to mount a valid argument. Arguing from the left, for example, critics will surreptitiously premise certain liberal norms to make their case, while those arguing from the right, they will tend to smuggle in just enough conservative value to make their argument stick.

Prudent critics challenge their addressee's norms by premising a part-portrayal of his *normative framework*, which, though largely true, differs noticeably from his own self-image. Since leveled against heartfelt norms, the argument itself will be flatly dismissed, but the picture it premises of its addressee's normative identity *may well register*.

Such deontic portrayals need not be deemed *true* to attract our attention. It is enough if we consider our critics *sincere* to realize how differently they see us from how we see ourselves. Since the two pictures - our own normative self-portrayal, and that conveyed by our critic - will diverge with regard to norms criticized, their incongruity, not unlike the case of a disturbing playback device, may succeed in destabilizing our initial commitments, and rendering us *ambivalent* toward the norms in question. And norms to which we become *ambivalent*, lose their normative hold, and will be subjected to the critical scrutiny of our surviving commitments. This is how exposure to the echo-chamber of trusted normative criticism can stimulate the kind of self-discordance needed for truly transformative normative *self*-criticism. Viewed thus, normative self-criticism is *not* a simple internalization of its external, neo-Hegelian counterpart. It is prompted by it, *caused*, if you wish. And yet, I insist, it is rational; at least as rational as the potentially ambivalating, self-correcting effect of a playback device.

This is the basis, I submit, for a less problematic and more broadly applicable account of rational paradigm shifts in and outside science. It depicts such transitions as two-stage processes, in which practitioners of standing are first "ambivalated" by external critics, and later succeed in promulgating their indecision within their communities by means of a particular form of published work.

As we have seen, truly transformative normative self-criticism depends necessarily on the impact of potentially 'ambivalating' external critics. But where do scientists and philosopher's encounter external criticism? Peter Galison's idea of the scientific trading zone is useful in this regard. Trading zones are the 'locales' of professional engagement outside one's home community, where practitioners bid for financial support, "trade" with neighboring disciplines for techniques or instruments, offer professional opinion, engage students, and so forth. Galison's work nicely portrays how, in so trading, disciplines are enriched by the findings of other disciplines. But he nowhere asks if and how practitioners might find themselves *challenged* in such settings? He portrays them, like

real trading zones, as providing ‘safe space[s] of shared phenomenology’, cordoned off from the parties’ ontological commitments by the scientific pidgins and creoles they employ. But I think he takes his metaphor too literally, and misses something crucial.

A highly relevant aspect of 'talking shop' with the uninitiated that Galison misses, is how, when abroad, practitioners are often required to articulate aspects of their worldview so taken for granted back home, that they go there without saying; how, when presenting to others, one is so often required to field the friendly and bemused questions of genuinely curious professionals, on whose thinking one's own framework may have far less of a grip. Here, trading and conversing beyond the confines of their professional milieu, is where practitioners are liable to experience the ambivalating force of external criticism.

But to set a framework transition in motion, the ambivalence acquired outside the community, needs to be effectively disseminated *within* it. The problem is that dithering hesitance is normally not deemed publishable, and is relegated to the personal diary or private correspondence, that don't carry their master's voice very far. But at times, the their doubts and indecision come tacitly to inform their creative efforts to overcome them, typically taking the form of confident, yet uneasily split, hybrid attempts to re-represent their field, in which elements of the old framework are combined with imaginative gropings toward new possibilities. Analyzed *prospectively*, such works can be shown to represent anxious, unstable, yet highly creative departures from heartfelt commitments capable of motivating others to seek cleaner, more radical breaks with the old. Classical examples are Tycho Brahe's hybrid planetary theory, and Friedman's two favorites: the uneasy mix of Aristotelian and anti-Aristotelian elements in Galileo, and Poincare's Kantian, yet non-Kantian geometrical conventionalism. But let me conclude by looking at four less known examples found in the writings of some of the Brits I've worked on: George Peacock's meta-mathematics, John Herschel and WR Hamilton's views of mathematical physics, and the unholy fusion of Kantian and empiricist strands in Whewell's philosophy of science - all of whom unwittingly propagated a keen ambivalence, prompting others to take firmer stands. Let me say a little both about the hybridity of the four positions, and the ambivalating contexts in which they were forged.

In 1833, Hamilton wrote to Whewell insisting on there being not one, but two, unrelated “sciences of dynamics”: “one subjective, *a priori*, metaphysical, deducible from



meditation on our ideas of Power, Space, Time; the other objective, *a posteriori*, physical, discoverable by observation” – a distinction he was careful to distinguish from empiricist, Kantian, and instrumentalist accounts of their connection. Although he regarded the former “higher in dignity”, he did “not consider it as including the other, or as adequate ground... for the expectation of any one appearance.” But rather viewed their “wondrous convergence” as due to their “mysterious union” in “the Divine Mind”.

Three years earlier, Peacock’s seminal *Treatise on Algebra* proposed splitting modern algebra into two separate and awkwardly related pertaining, again, to very different spheres of algebraic activity: “arithmetical algebra”: considered as the science of quantity duly purged of negatives and imaginaries, and “symbolical algebra”: perceived as a wholly formal calculus of unconstrained symbols and operations, applicable to negatives, imaginaries and much more. Here too, no simple subsuming of the two was assumed. For Peacock, symbolical algebra was *not* a mere generalization of arithmetical algebra, just as the latter, not merely an application of the former to numbers.

In his extensive studies of science conducted at the time, Whewell insisted, similarly to Hamilton though less crudely, and with considerably more philosophical finesse, that each of the “inductive sciences” comprised two integrated, mutually cultivated, yet antithetical components, arrived at by means of a two-pronged methodology geared to the attainment of two independent notions of truth. The term ‘antithetical’ is his.

And to the explicit and acknowledged splittings presented in the works of Whewell, Hamilton and Peacock, one should add John Herschel’s meticulously compartmentalized writings on natural philosophy and mathematics, which when taken together, strongly imply a similar, and even ruder split segregating the factual and the formal components of mathematical physics to near-incongruous spheres of intellectual pursuit.

All four dualisms were in important respects strained and inherently unstable, with one side of each of them pertaining to, or at least building on commitments from which its author had initially set forth, and the other representing radically diverging possibilities.

Friedman makes much of intermediary figures like Galileo and Poincare, whom he sees as proof of the smooth “naturalness” in which the new gradually evolves out of the old. I disagree. All these works attest to tormented irresolvable inner struggles capable of

yielding no more than a shaky compromise; representing positions plagued by profound, yet *highly inventive* indecision, that took their authors years, if not decades to articulate and refine, holding on to the old while forced to grope creatively toward new options.

Take Peacock's predicament: was algebra to be understood as a logically connected system of *truths* about the realm number, as he was taught and always believed, or was it better considered a purely formal, symbolical *system of formal rules*, like those of a game, or grammar, capable of achieving high levels of coherence, clarity and completeness, but that were true *of nothing*? The two views of algebra give rise to very different ways of doing mathematics, and incompatible notions of mathematical meaning and verity. And yet, they cannot be adjudicated by testing. For Peacock they were not attempts to faithfully *represent* algebra as it was in fact pursued, nor ways of improving algebra's capacity to attain its objectives, but attempts to state how algebra *ought* to be understood and pursued. They offered different *final ends* for understanding and pursuing algebra, if you wish, not different means of attaining them.

Peacock's *Treatise* wasn't an idle nor an abstractly philosophical exercise. As Joan Richards, Kevin Lambert and others have noted, the view of algebra as generalized arithmetic was the one to which he had been firmly committed from the start. The second, formalist option, was a view he encountered in an unpublished work by Babbage on the "Philosophy of Analysis," which he read and contested around 1820. But he was disturbed sufficiently to spend the best of a decade trying unsuccessfully to ground a richer arithmetical algebra on an unprecedented survey of the history of counting and arithmetic. Convinced of his failure, he ended up dithering between the two incompatible options, painfully straddling the unbridgeable gulf, investing the best of his considerable mathematical powers, to fashion a hybrid combination of the two that was as ingenious as it was unfounded. Peacock had become sufficiently ambivalent toward his former view to seriously consider its rival, to which he became sufficiently devoted to be able to endorse them side-by-side in uneasy, undecidable conjunction.

And much the same applies to the different meta-scientific musings of Herschel, Hamilton and Whewell. Herschel was from the beginning a staunch advocate of Lagrange's thoroughly formalistic approach to the calculus, which he swiftly extended to mathematics in general. Yet, unlike Lagrange, whose mechanics was as formalist as his

calculus, Herschel adopted a thoroughly realistic, explanatory Baconian view of natural science, which rendered their juxtaposition in modern mathematical physics uneasy in the least? What could be the role was mathematics in modern physics, if it was not conceived of as *true* or even elucidating of the real, physical quantities to which it was applied? This was the question I believe plagued Herschel's meta-scientific musings throughout his career. But unlike his friends, Herschel kept his deliberations to himself, making no move to articulate his dilemma, or resolve it. He chose instead to avoid the issue by relegating his writings on the nature of mathematics and on the philosophy of "natural philosophy", to quite separate publications. The fact that he remained faithful to both positions throughout his career, clearly attests to the depth of his commitments; the fact that he made no attempt to even raise the question of combining them, can be seen as proof of the depth of his ambivalence. Here was one of the most respected thinkers of the era, caught between painfully jarring normative options to the point of speechlessness!

Hamilton's quandary also concerned the best of mathematical physics, to which he contributed famously. Hamilton, who initially subscribed to a realist, rather than formalist, view of mathematics, was untroubled by the role of mathematics in modern physics. His concern had to do with the dividing line between what we now call theoretical and experimental physics. He believed that dynamics was developed and perfected in the course of two unrelated efforts. On the one hand, by intense conceptual analysis, which far transcended the science's *mathematical* apparatus. Dynamics' entire conceptual structure, its notions of cause and force, velocity and acceleration, energy and momentum, along with their exact and complex relationships, he believed, were the combined product of intense *contemplation*, which, he believed, at no stage depended on experience. Hamilton's view of this "subjective, metaphysical" of the two sciences has been shown by Richard Yeo and others to have owed to a deep and early commitment to a romantic view of poetic insight and cognition born of Hamilton's meaningful relationship first with Wordsworth and later with Coleridge, which, once he had finally settled on his academic vocation, he carried into science and mathematics. And yet he also believed that calculated experiment and observation yielded quite independently of the theoretician's efforts, just as rich a picture of physical reality of their own.

The two scientific undertakings appeared to Hamilton to have nothing in common; to be

motivated and informed by quite different paradigms, and yielded by different methodologies in the hands of different communities. And yet, in some mysterious way the keen, systematic, inward-looking investigation of our conceptions and ideas perfectly coincided with the equally keen and systematic outward-looking study of empirical fact. Hamilton's quandary owed, again, to an inability to reconcile incompatible commitments regarding the means of *production* and sources of validity of modern physics.

Unlike Peacock and Whewell, Hamilton had little to offer by way of bridging the divide. From any human perspective, he deemed their convergence a coincidence of truly colossal dimensions, which he ultimately ascribed to their mystical union in the mind of God. Hence, there was not one, but two science's of dynamics, uneasily and puzzlingly poised in relation to each other, converging perfectly by virtue of nothing except a Divinely mandated pre-established harmony! Like Herschel's thunderous silence, Hamilton's resort to mysticism bears firm witness to his normative ambivalence.

Finally, Whewell's "antithetical" philosophy of science, I still believe, contrary to some of his more recent readings, was the product of a metascientific journey more complex, more uneven and more profound than those of his three close friends. Like Herschel, the hybrid synthesis he ended up fashioning preserved on its one side an early native commitment to a Baconian form of empirical realism, according to which the very best of 'inductive' science can be proven on occasion empirically true in retrospect. But at the same time, he had become just as fully committed to a form of idealism late of Kant, far more radical than Hamilton's, according to which all conceptual content originates in the mind, deeming all assertions regarding the world without to be acts of reading *in*, rather than in reading out. Whewell's epistemology and philosophy of discovery were on the one hand thoroughly Kantian, while, at the same time, preserving a determinately non-Kantian, empiricist commitment to the idea that the best of science can at times be proven true not of the world we experience, but of the world in-itself!

\*\*\*

My readings of the deeply split, hybridic nature of the philosophies of science and mathematics implied by the work of the four friends, the stories I tell of their development and eventual impact, have been, and will probably remain contested. The aim of this paper is not to convince you that I have got them right, so much as to make

the more general, HOPOS-related point regarding their historiographical significance, which may be summarized briefly thus:

- Not unlike science itself, the history of the philosophy of science also witnessed fundamental framework transitions.
- The problem of accounting for their rationality is hence as pressing for HOPOI, as it is for historians of science.
- Such transitions, can be deemed rational, I have argued, when practitioners of standing, first prompted by the ambivalating impact of external normative critique, produce confident, yet inherently unstable new hybridic accounts of the field, that unwittingly succeed in imparting their ambivalence to their peers and disciples.