Rational Reconstruction Reconsidered*

"*Arnheim*: […] and nothing irrational happens in world history.
Ulrich: But so much that is irrational happens in the world, surely?
Arnheim: In world history, *never*!"[1]

1. Introduction: History of Science versus History of Scientists

Here is a dilemma concerning the history of science. Can the history of scientific thought be reduced to the history of the beliefs, motives and actions of scientists? Or should we think of the history of scientific thought as in some sense independent from the history of scientists?

If the history of scientific thought is just, or is reducible to, the history of scientists, then it seems that philosophy of science has little to do with the history of science. It may (and hopefully does) learn from the history of science, but philosophy in no way does (or should) influence the way history of science is done. By describing the actions and motives of scientists, as well as the relevant sociological, institutional, and cultural background, we get a full picture of the history of science. This approach will have little patience for some of the classic topics of the philosophical history of science that dominated the 60s and 70s of the last century, such as the rationality of theory change.

The other horn of the dilemma would be to claim that the history of scientific thought is (in some sense) independent from the history of scientists. Taking this route often amounts to interpreting science as a system of ideas or thoughts that develops according to its own logic and we do not have to take into consideration the actual actions and motives of individual scientists in order to write the history of science.

The aim of this paper is to carve out an intermediate position between these two. I will argue that the history of scientific thought supervenes on, but is not reducible to, the history of scientists. There is a legitimate level of description for analyzing the history of scientific thought that does not reduce to the individual level of scientists. Yet, every aspect of the history

"Rational Reconstruction Reconsidered" by Bence Nanay,
The Monist, vol, 93, no. 4, pp. 598–617. Copyright © 2010, THE MONIST, Peru, Illinois 61354.

of scientific thought is determined by the actual motives and actions of individual scientists.

Maybe surprisingly, I use Imre Lakatos's controversial concept of the rational reconstruction of the history of science in order to argue for this intermediate position. Lakatos is often taken to be one of the most radical proponents of the second horn of the dilemma I sketched here. I argue that we can use Lakatos's theoretical framework to show why we need to take both the history of scientific thought and the history of scientists seriously if we want to write history of science. The aim of the paper is to examine how these two aspects of the history of science can and should be combined.

2. Internal and External History of Science

I talked about the history of scientific thought and the history of scientists above. I will use a less colorful but more precise formulation in what follows. I will use a distinction between the internal and the external history of science as these terms are used by Lakatos (Lakatos 1970, 1971). It is important to note that Lakatos's way of using these terms is highly idiosyncratic (see Lakatos 1971, 123, n. 1, Hacking 1979, 394, Hacking 1983, 122)[2]: it is very different from what intellectual historians mean by internal and external history (see Shapin 1992 for a good summary of the many ways this distinction is used by intellectual historians). I will use Lakatos's terminology (and not the standard terminology of intellectual history) in what follows.

Lakatos takes the external history of science to be a socio-psychological narrative that describes the beliefs, motives, actions, and other mental states of scientists, together with their institutional background. Internal history, in contrast, is taken to be a description of the history of scientific thoughts and ideas and of "objective scientific growth" (Lakatos 1970, 180).

So it seems that every event in the history of science has two descriptions, an internal and an external one. Take the 'modern synthesis' of evolutionary biology, for example. The external historian would examine the various intellectual influences of specific scientists, for example, of Ronald Fischer, Theodosius Dobzhansky, Ernst Mayr, as well as the interactions between them. The internal historian, in contrast, would talk about the combination of two scientific theories, Darwin's theory of natural selection and population genetics. The specific scientists and their interactions will not play any role in this narrative.

Lakatos has a seemingly extravagant view on the relation between internal and external history, which I will come back to in the next section. For now, I would like to give some examples for writing history of science in a purely external and in a purely internal way. Unsurprisingly, the moral of the story will be that we need to pay attention to both internal and external history in order to write any meaningful narrative of the history of science. The aim of the paper is to analyze how exactly the internal and the external aspects can and should combine.

It is not too difficult to find examples for external history of science that ignore (or pay little attention to) internal history as most historians working in the tradition of 'sociology of science' follow this methodology. A famous and quite extreme example is Bruno Latour and Stephen Woolgar's *Laboratory Life* (1979). Latour, who has no training in molecular biology, spent a couple of years observing the research on growth hormones in a molecular biology lab at the Salk Institute in San Diego. In this description of laboratory life, the authors actively and deliberately ignored the content of the research that was undertaken in this lab. Similar (but maybe less radical) methodology has been used in describing more distant episodes in the history of science (another famous example is Shapin and Schaffer 1985, see also Shapin 1982 for an important theoretical/methodological manifesto for the sociology of science approach).

The sociological approach to the history of science can be, and has been, used in more or less radical fashion. The one I have been focusing on here is the more radical version, the one that denies the relevance of internal history (to which Latour-Woolgar 1979 is a good example). One problem with this radical version of the sociology of science approach is that very often ignoring internal considerations makes it very difficult to describe what is going on in external history. If a theory T implies a claim C, then if we describe a scientist who accepts T, it is easy to explain why she holds C with reference to this piece of internal history. If we cannot use internal history, then the reason why the scientist holds C needs to be explained in terms of the scientist's psychological history. This will look even more difficult if C is not a claim that the scientist holds explicitly, but rather a claim that she takes for granted because she (explicitly) accepts T. In this case, internal history can help us to explain why she takes C for granted, but it is difficult to see how external history alone can do so.

We have seen an example for using external history only in the writing of history of science. But what is the other extreme? Can we do only internal history, ignoring any external considerations?

One example for this purely internalist approach is implied by some evolutionary approaches of scientific change. It has been suggested that like evolved organisms, scientific theories also compete for survival and reproduction: both science and natural selection proceeds by trial and error. As we can talk about natural selection among organisms and lineages, we can also talk about selection among scientific theories. There are various ways of substantiating this evolutionary analogy (some famous examples include Popper 1959/2002, 1963, 1972, 1975, 1978; Toulmin 1967, 1970, 1972; Kuhn 1972, 172; Van Fraassen 1980, 39–40; Hull 1988, 2001; see Bradie 1986 for a typology). But the one that is relevant for our purposes is the meme theoretical reconstruction of the history of science (see Hull 1988, 2001; Aunger 2002; Distin 2005).

According to this proposal, the history of science, like other cultural phenomena, can be explained, at least partially, with the help of the following evolutionary model: Memes are pieces of information and they compete for survival in a quite similar way as genes do; the difference is that they compete for the capacity of our minds. Since the capacity of the human mind is limited, only some of them, the successful ones, manage to get into the minds of numerous people, hence, they survive, whereas the unsuccessful ones die out. A meme can be a tune, the idea of liberalism, or the habit of brushing one's teeth. Those tunes will survive that can get into and stay in many minds. The ones that fail to do so will die out (Dawkins 1989, 1982a, 1982b; Dennett 1995, 2003, 2006). This general explanatory model can be applied to the history of science as well: scientific theories are memes: they spread from the mind of one scientist to the other (David Hull made a very similar proposal in his Hull 1988 and 2001, but see Bechtel 1988, Ghiselin 1988, and Griesemer 1988 for a variety of objections to his account).

Meme theory in general has been severely criticized (Sperber 1996, 2000; Wimsatt 1999; Richerson-Boyd 2005; Sterelny 2006a, 2006b), but even if we assume that there are such things as memes and even if we also assume that we can talk about something like a selection process among them, the meme theoretical reconstruction of the history of science will still look somewhat problematic. If the history of science can be described in terms of the selection pressures of meme selection, then scientists themselves are left out of this process altogether. They are the vehicles of memes at best and the important and explanatory relevant causal relations are to be found among memes, not scientists. If we take the history of science to be the history of memes, then this gives us an example of purely internal history that ignores any external factors. The history of science, interpreted this way, bypasses scientists.[3]

The aim of this section was not to give knock-down arguments against certain branches of the sociology of science or against the meme-theoretical approach to the history of science. I introduced these two examples and pointed out some of their problematic features because they signify the two extremes of writing history of science. It seems that neither ignoring internal history nor ignoring external history leads to an unproblematic historiographic approach. The unsurprising conclusion is that if we want an unproblematic methodology for writing history of science, we need to combine the two. How exactly this could and should be done is the question I now turn to.

3. Rational Reconstruction

Imre Lakatos is often considered to be the champion of internal history and of the ruthless dismissal of external considerations. I aim to show that Lakatos in fact urges an interesting combination of internal and external history.

According to Lakatos, the first step of writing a history of science must be the reconstruction of internal history. This is the step Lakatos calls 'rational reconstruction' (Popper also talks about the rational reconstruction of the history of science; it is unclear who inherited this concept from whom, see e.g., Popper 1972, 179). As Lakatos says, "whatever problem the historian of science wishes to solve, he has first to reconstruct the relevant section of the growth of objective scientific knowledge, that is, the relevant section of 'internal history'" (Lakatos 1971, 106). What he means by rational reconstruction is perhaps more appropriately described as rational construction: there is no guarantee that the rationally reconstructed internal history will correspond to the actual historical facts. Lakatos explicitly acknowledges this: "Internal history is not just a *selection* of methodologically interpreted facts: it may be, on occasions, their *radically improved versions* (Lakatos 1970, 106).[4] In other words, rational reconstruction distorts what we know to be the historical facts and this gives rise to internal history.

Lakatos's writings are full of provocative claims about just how distorted this internal history will look. Here is the most famous (or, rather, infamous) quote:

> One way to indicate discrepancies between history and its rational reconstruction is to relate the internal history *in the text* and indicate *in the footnotes* how actual history 'misbehaved' in the light of its rational reconstruction. (Lakatos 1971, 107)[5]

This quote as well as Lakatos's general seemingly dismissive attitude towards what he calls 'actual history' triggered very strong reactions both from philosophers and from historians. Thomas Kuhn writes that "What Lakatos conceives as history is not history at all but philosophy fabricating examples" (Kuhn 1971, 143). Or, more precisely,

> A *historian* would not include *in his narrative* a factual report which he knows to be false. If he had done so, he would be so sensitive to the offence that he could not conceivably compose a footnote calling attention to it. (Kuhn 1970, 256)

Larry Laudan's reaction is equally strong: according to him, Lakatos's methodology is "consciously and deliberately falsifying the historical record" (Laudan 1977, 170). Gerald Holton is even more negative when he writes about Lakatos's rational reconstruction of Bohr's early work, which he considers to be "an ahistorical parody that makes one's hair stand on end" (Holton 1978, 106, see also Holton 1974, 75). These are not isolated examples, see also Godfrey-Smith 2003, 103–104; Koertge 1976; Richardson 2006; Holton 1978, 105–107; McMullin 1970).

Before dismissing Lakatos as a bad historian of science and dismissing Lakatos's vision of history of science as either ahistorical or crazy, it is important to remember that rational reconstruction for Lakatos is just the first step of writing history of science. It is not the end of the story. Internal history is not the finished product of what historians are supposed to do, but only the first, preparatory stage. Lakatos is so explicit about this that it is striking how many historians and philosophers misinterpreted his account.

For Lakatos, the first step of rational reconstruction needs to be followed by a second phase where "one tries to compare this rational reconstruction with actual history and to criticize both one's rational reconstruction for lack of historicity and the actual history for lack of rationality" (Lakatos 1970, 138, n. 40). Or, as he reiterates, "any rational reconstruction of history needs to be supplemented by an empirical (socio-psychological) 'external theory'" (Lakatos 1971, 91). This is not a passing remark, but the summary of the third of the three main claims he argues for in (Lakatos 1971). In short, the history of science is not identical to its internal history (that is, its rational reconstruction). Lakatos is not a radical internalist. In fact, he is very much against radical internalism. As he says, "since external influences always exist, radical internalism is utopian" (Lakatos 1971, 94). According to him, writing history of science requires attention to both internal and external history as well as to the interaction between the two: "history of science is always richer than its rational reconstruction" (Lakatos 1971, 105).

But if this is true, then how should we interpret his provocative remark about writing internal history in the main text and exile external history to the footnotes? Lakatos, somewhat defensively, describes the infamous passage about footnotes as "a rather unsuccessful joke" (Lakatos 1976/1978, 192) and adds that "of course such parodies may be written, and may even be instructive; but I never said that this is the way in which history actually ought to be written and, indeed, I never wrote history in this way" (ibid). In short, rather than dismissing external history, "Lakatos merely suggests a colourful way of doing something quite orthodox" (Musgrave 1983, 66. See also the similarly charitable interpretation of Brown 1989, 109–111; Hacking 1979, 396, 1983, 125).[6]

So if we follow Lakatos, we need both internal and external history in order to write a history of science (and we should only joke about dismissing the latter into footnotes). But how do these two aspects of history relate to one another? Lakatos finds it important to emphasize a number of times that internal history is primary and external history is secondary (see, e.g., Lakatos 1971, 92, 105). Why? Because "the most important problems of external history are defined by internal history" (Lakatos 1971, 105, see also Lakatos 1976/1978, 191 and Lakatos 1971, 92 for similar formulations). Or, to use his telling metaphor, "the internal skeleton of rational history defines the external problems" (Lakatos 1976/1978, 191).

In other words, internal history is far from being the history of science. It is only the skeleton of history of science and a lot of work needs to be done to get from the skeleton to the full body of history of science (see, e.g., Lakatos 1971, 118, Lakatos 1976/1978, 191–92). Nonetheless, although internal history is merely the skeleton of the history of science, it is necessary to have this skeleton to build on. Internal history is a necessary ingredient of any serious history of science (see Lakatos 1976/1978, 192).

One may wonder why Lakatos takes internal history to be so important. The answer has to do with the fact that the aim here is to explain the history *of science*, a mainly rational social enterprise, which needs to be kept separate from the history of other social phenomena. In other words it is in order to preserve the special status of science that we need to use rational reconstructions. Doing without rational reconstruction would make "scientific change a kind of religious change" (Lakatos 1970, 93). And this "would vindicate, no doubt unintentionally, the basic political *credo* of contemporary religious maniacs" (ibid).

It is time to summarize what we have learned from Lakatos about the ways in which internal and external history could and should be combined. For Lakatos, we need both. If we ignore internal history, then we lose what is special about science. And this is exactly what certain branches of the sociology of science do. Lakatos has little patience for this approach: "The work of those 'externalists' (mostly trendy 'sociologists of science') who claim to do social history of some scientific discipline without having mastered the discipline itself, and its internal history, is worthless" (Lakatos 1971, 128, n. 68).

But Lakatos also holds that external history should not be ignored either. This is especially clear from his criticism of Stephen Toulmin's writings (see, also, Nanay forthcoming).[7] Toulmin insisted that evolution should be more than a mere metaphor when we describe the progress of science (Toulmin 1970, 560–64, see also Toulmin 1972). It is not enough to compare the trial and error method of science to the trial and error method of natural selection. The evolutionary model is indeed explanatory (Toulmin 1967, 470): selection among scientific theories explains some of the features of these theories, most importantly, their survival. Lakatos's main problem with Toulmin's account is that this selectionist explanation bypasses scientists and philosophers, very much like the cunning of Hegelian reason. In other words, he seems to be criticizing Toulmin for ignoring external history—something Lakatos himself is often—as I tried to show, mistakenly—accused of.[8]

In other words, Lakatos holds that both internal and external history are indispensable. And he has an account of how the two are related: internal history is logically prior to external history: historians of science should first give a rationally reconstructed internal history and then give the external history on the basis of that. In order to substantiate this account, we need to examine how exactly external history presupposes internal history and we also need to examine some thorny ontological questions about the entities internal history is about.

4. Why Do We Need Rational Reconstruction?

It is time to depart from the specifics of Lakatos's account of rational reconstruction and examine how the concept of rational reconstruction could be made plausible, regardless of what Lakatos said about it. And, in fact, Lakatos did not give a full account of how rational reconstruction is supposed to work. More precisely, he never tells us why and how internal history sets the problems for external history. As Larry Laudan rightly points out, "Lakatos nowhere establishes the necessity (or the desirability) of making a reconstruction of the past which involves an intentional warping of the historical record" (Laudan 1977, 170). Lakatos gives a lot of extravagant metaphorical claims about the necessity of internal history for writing history of science. Perhaps the most famous of these is the second half of his Kantian paraphrase: "Philosophy of science without history of science is empty; history of science without philosophy of science is blind" (Lakatos 1971, 91). But Laudan's criticism is absolutely correct: Lakatos never give any reason why the rational reconstruction is necessary for writing history of science.

Here is one reason why we could think that rational reconstruction is a necessary feature of writing history of science. Explanation is always contrastive: explaining why x is *F* rather than G is a different explanatory task from explaining why x, *rather than y*, is F (Van Fraassen 1980, 142–43). In other words, explanation is always relative to a contrast class. I am assuming that the same is true for historical explanations.

But then the question is what allows the historian of science to identify the contrast class for her historical explanations. What contrast classes are worth taking into consideration and what are not? My answer is that the contrast class of explanations in the history of science comes from the rational reconstruction of internal history. And this is the sense in which external history presupposes the rational reconstruction of internal history.

Here is an example. Many historians of biology have been trying to explain why Darwin attributed so much importance to Henry Charles Fleeming Jenkin's objection to his theory of natural selection (Gould 1991; Bulmer 2004; Morris 1994; Bowler 1983, 1988, 1990; Bulmer 2003, 141–45; Cookson and Hempstead 2000). Jenkin's paper was published in 1867 and two years later, Darwin wrote that "Fleeming Jenkin has given me much trouble, but has been more real use to me than any other essay or review" (letter to Joseph Hooker, in Darwin-Seward 1903, vol. 2, 379. See also a similar claim in his letter to Alfred Russel Wallace also in 1869). Jenkin's objection about blending inheritance became perhaps the most important criticism of *The Origin of Species* and even of the theory of natural selection in general (before the 'modern synthesis') (see Hull 1973 and Vorzimmer 1963, 1970 for summaries). Stated very simplistically, the objection is that natural selection cannot explain real evolutionary change as, because of the 'blending' nature of inheritance, variations from the average will be watered down to be closer to the average in the next generation. The question is: why did Darwin take Jenkin's objection to be so important?

How should the external historian begin to address this question? The first thing they should do (and many historians in fact do) is to see how good Jenkin's objection really is and what aspect of Darwin's theory it jeopardized. The standard interpretation of Jenkin's objection is that it made it clear that Darwin was using a mistaken theory of inheritance (the 'blending' theory) and as long as we take inheritance to be blending the traits of the two parents, then Jenkin is correct to point out that natural selection will not be able to explain major evolutionary change (see, for example Lewontin 1986). This interpretation of Jenkin's argument is consistent with the textbook narrative of the significance of the 'modern synthesis' that replaced Darwin's original concept of inheritance in the theory of natural selection with a Mendelian one. Thus, Darwin was right to take Jenkin's criticism seriously.

But here is an alternative interpretation. Ernst Mayr argues convincingly that Jenkin's paper is deeply confused and it does not present any good objection (Mayr 1982, 512–14, see also Kitcher 2007, 73–75, who also nicely exposes the racist undertones of Jenkin's paper). Mayr argues that Jenkin failed to grasp what is, according to Mayr, *the* most important element of Darwin's theory, what he labels 'population thinking': the view that in the biological domain individual variation cannot be ignored and subsumed under some fixed and preexisting type. Individual variation is what drives evolution and we can only talk about types that these variations are instantiations of as statistical abstractions (Mayr 1959/1996, see also Sober 1980 and Nanay forthcoming a). Jenkin assumes that variation is always variation within a type. According to Mayr, this is an instance of the 'typological thinking' or essentialism that Darwin was strongly opposed to. Thus, Mayr concludes, Jenkin misunderstood the most important claim of the theory of natural selection. Darwin could have easily refuted him, but he didn't. In other words, Darwin was completely mistaken to take Jenkin's criticism seriously (see esp. Mayr 1982, 514).

We have two different historical explanations for Darwin's assessment of Jenkin's argument. What matters for our purposes is not which one is correct, but in what way they differ. They differ in as much as they give different rational reconstructions of the internal history of the Darwinian revolution. And, as a result, the external historical narrative will also look very different.

If we accept Mayr's explanation and claim that Darwin was wrong to take Jenkin's criticism seriously, then we also need to interpret the related changes Darwin made in later editions of *The Origin of Species* (especially on p. 72 of the 1872 edition of *The Origin of Species*) as insignificant (see, e.g., Vorzimmer 1963, 1970). And if we accept his interpretation, then this sets the agenda for future historical research: the next question the external historian should ask is why was Darwin so mistaken about the vulnerability of his own theory in the face of Jenkin's objection? Mayr's (sketchy) answer is that it is because he was "rather confused on the topic of variation" (Mayr 1982, 514, see also 681–97).

If, in contrast, we accept the standard narrative about Jenkin's criticism, then the changes Darwin made in the later editions will be considered to be very significant indeed. The agenda for future historical research will also look very different: the next step is likely to involve the understanding of the reasons why Darwin was mistaken about inheritance.

Crucially, what constitutes the difference between these two (external) historical explanations is the way in which Darwin's theory is being rationally reconstructed. For Mayr, the crucial element of this theory is population thinking. That is why, according to Mayr, Darwin should have dismissed Jenkin's objection without thinking twice about it. According to the standard interpretation, however, Darwin's mistaken ideas about inheritance are part of the most central Darwinian claims (to be contrasted with the correct ideas of the 'modern synthesis'). Thus, the two historical explanations differ in what they take to be what Lakatos would call the 'hard core' and the 'protective belt' of Darwin's theory. This is a difference in internal history. And, as we have seen, depending on how we rationally reconstruct this internal history, we get different directions for future (external) historical research. Further, depending on how we rationally reconstruct internal history, the course of external history, for example, the change of Darwin's thinking between the earlier and the later editions of *The Origin of Species*, will also look different.

The aim of this brief case study was to show that the rational reconstruction of internal history is presupposed by the writing of external history. Mayr's and the standard historical explanation differ in the assessment of the 'hard core' of Darwin's theory, and, as a result, ask different questions about the influence of Jenkin's criticism on Darwin's thinking. This means that they use different contrast classes when they ask why Darwin took Jenkin's criticism seriously. Mayr asks why Darwin took it seriously, rather than just dismissed it as a piece of old essentialist typological thinking. The standard account asks why Darwin was forced to take Jenkin's criticism seriously, rather than having a theory of inheritance that would not have made his theory to be susceptible to it. And as a result of this difference in the contrast class of the historical explanation, the course of external history of the Jenkin-Darwin episode (as well as the further questions it raises) will also be very different. As Lakatos said, "the most important problems of external history are defined by internal history" (Lakatos 1971, 105).

5. Ontological Worries about Rational Reconstruction

In the last section I departed from Lakatos's concept of rational reconstruction to show that it is a necessary feature of writing history of science. In this section I depart from Lakatos even further: I question some ontological assumptions he made about the nature of (rationally reconstructed) internal history.

External history is about scientists, their beliefs, motives and actions. But what is internal history about? The official formulation is that it is about the growth of knowledge, in fact, the growth of objective knowledge. But what kind of entities does internal history presuppose?

Lakatos makes it very clear that internal history is not about the scientists' beliefs, motives and actions. As he writes, "the internal historian will not need to take any interest whatsoever in the *persons* involved, or in their beliefs about their own activities" (Lakatos 1971, 127, footnote 60, see also Lakatos 1971, 106).

Note that this is an epistemic claim: if we are interested in the history of scientific thought, we need to describe the history of science in such a way that would not need to refer to the beliefs, motives and actions of scientists. But Lakatos goes further and seems to also make an ontological claim:

> The—rationally reconstructed—growth of science takes place essentially in the world of ideas, in Plato's and Popper's third world, in the world of articulated knowledge which is independent of knowing subjects. (Lakatos 1970, 180)

Lakatos makes it clear a number of times that internal history is part of Popper's third world (Lakatos 1976/1978, 191, Lakatos 1971, 179). But what does he (or Popper) mean by the third world? According to Popper, "the third world is man-made and, in a very clear sense, superhuman at the same time. It transcends its makers" (Popper 1968/1972, 159). Further, it seems to have causal powers, independent of the causal powers of the second world (ibid.)

Lakatos is somewhat ambiguous about how demanding an interpretation of the third world he embraces. In the quote above (and elsewhere) he seems to equate the third world with the heaven of Platonic Forms. As he says in one of his last papers, "the third world is the Platonic world of objective spirit" (Lakatos 1976, 128). Elsewhere, however, he takes the third world to be the world of propositions: "the 'first world' is that of matter, the 'second' the world of feelings, beliefs, consciousness, the 'third' the world of objective knowledge, articulated in propositions" (Lakatos 1971, 127, footnote 61). It is also unclear whether he follows Popper in attributing causal powers to the entities of the third world. But what matters for our purposes is that for Lakatos, as for Popper, there is an important ontological difference between scientific thoughts and the thoughts of scientists. According to Lakatos, rational reconstruction gives us 'third world' entities: entities that are "not dependent in the slightest on the scientists' beliefs, personalities or authority" (Lakatos 1971, 106). Much of the discussion about Lakatos's concept of rational reconstruction takes this ontological independence of internal history for granted, whether or not they side with Lakatos (see, e.g., Elkana 1974, 245; Kulka 1977, 331).

Should we follow Lakatos in these ontological claims? Some defenders of Lakatos's concept of rational reconstruction have embraced his claims about the 'third world' (notably Palmer 1993). But I agree with Ian Hacking that these claims about the ontology of internal history are part of the reason why Lakatos's concept of rational reconstruction has often been treated with suspicion. As Hacking says, "Lakatos's internal history is [. . .], in short, to be a history of Hegelian alienated knowledge, the history of anonymous and autonomous research programmes" (Hacking 1983, 122). Hacking's words may seem too strong (especially in the light of Lakatos's criticism of Toulmin for his excessive Hegelian vision of history), but some of Lakatos's early writings seem to show that Hacking was right. Here is what Lakatos says in *Proofs and Refutations*:

> Mathematical activity is human activity. Certain aspects of this activity—as of any human activity—can be studied by psychology, others by history. [. . .] But mathematical activity produces mathematics. Mathematics, this product of human activity, 'alienates itself' from the human activity which has been producing it. It becomes a living growing organism that acquires a certain autonomy from the activity which has produced it. (Lakatos 1963/1964, 146)

Thinking of objective knowledge as 'a living growing organism' sounds Hegelian indeed, which, for some philosophers, may disqualify Lakatos's account of rational reconstruction altogether. The question is whether we can depart from Lakatos's claims about the ontological independence of internal history and give an account of rational reconstruction that does not rely on such hints of Hegelianism.

My claim is that we can and in fact should discard Lakatos's claims about the 'third world' and about the ontological independence of internal history in order to arrive at a plausible interpretation of rational reconstruction. We can think of internal history as being ontologically dependent on the beliefs, thoughts, and actions of scientists, yet, we can also treat it as irreducible to these psychological facts. The proposal is that internal history supervenes on the history of the beliefs, motives and actions of scientists, but it is irreducible to them.

And this is in fact what the analysis of the reasons why we should talk about internal history at all, which I have outlined above, suggests. Recall that we considered two historical explanations of the Darwin/Jenkin episode. The difference between them was a difference in the rational reconstruction of internal history. According to Mayr's interpretation, the hard core of Darwin's theory is population thinking. And this sets the agenda for the explanations of, and future problems for, external history. The question is: what is this entity, Darwin's theory that Mayr makes claims about? Is it an entity ontologically independent from Darwin's beliefs? Or is it just the sum total of Darwin's beliefs? I think that the only plausible answer is that what Mayr (or any internal historian) refers to when talking about Darwin's theory is not ontologically independent from Darwin's beliefs, but neither is it reducible to them. The internal historian cannot take Darwin's theory to be ontologically independent of Darwin's beliefs (say, by taking it to be an 'ideal' theory of evolution) as many aspects of this theory clearly depend on the contingent beliefs Darwin happened to have, say, about the nature of inheritance (which were far from ideal). Nor should the internal historian take Darwin's theory as the sum total of, or as reducible to, Darwin's beliefs. Mayr himself admits that Darwin was not entirely consistent in his treatment of population thinking and that he often fell back to typologist terminology (see, e.g., Mayr 1982, 488). Yet, the internal historian, according to Mayr, should treat population thinking as part of the hard core of Darwin's theory. Darwin's rationally reconstructed theory differs from Darwin's actual beliefs. But Darwin's rationally reconstructed theory is still *Darwin's* theory, a theory attributed to a specific person in the history of science. Thus, close examination of the role internal history plays in the writing of history of science reveals that internal history should not be taken to be ontologically independent from the scientists' beliefs, motives and actions, but they should not be taken to be reducible to these either.

Take another example that was mentioned above: the 'modern synthesis' of evolutionary theory. We have seen that when the internal historian talks about this chapter in the history of biology, she does not examine the beliefs and correspondence of Ronald Fischer, Theodosius Dobzhansky, and Ernst Mayr. The internal historian would talk about the combination of two scientific theories, Darwin's theory of natural selection and population genetics. She will not worry too much about the specific scientists and their interactions. In other words, internal history should be given at a level of description that is not reducible to the level of description of the beliefs of scientists.

But this does not imply that internal history is ontologically independent from the beliefs, motives and actions of scientists. When we describe the shape and size of solid middle-sized physical objects, like my desk, we do not pay too much attention to the level of subatomic particles. Yet, my desk is composed of these subatomic particles. Similarly, the internal history of the 'modern synthesis' supervenes on the history of the beliefs, motives, and actions of specific scientists. If the history of the 'modern synthesis' had been different, the history of the beliefs, motives, and actions of Fischer, Mayr, or others would have had to be also different. Yet, we can describe the history of the 'modern synthesis' as the unification of two scientific ideas without any reference to the beliefs, motives, and actions of any scientists.

Thus, we do not have to posit new entities in a Popperian 'third world' or in a Platonic heaven in order to give a rational reconstruction. Remember that rational reconstruction is supposed to provide a necessary ingredient for any history of science and, as I argued, it is supposed to set the contrast class for any external historical explanations. It is difficult to see how Platonic Forms would be able to do this. My more straightforward (and more parsimonious) suggestion is that internal history is about entities that supervene on (but are irreducible to) the beliefs, motives, decisions, and actions of the actual scientists.

6. Conclusion: Nonreductive Monism about the History of Science

We have arrived at a picture of the relation between internal and external history that could be described as nonreductive monism. Internal history is not ontologically independent from external history. But it cannot be reduced to external history either.

An obvious analogy for this way of thinking about the relation between internal and external history comes from philosophy of mind (Hacking 1983, 124 and Hacking 1979, 393 alludes to the same analogy briefly). According to nonreductive monist theories of the mind, mental states supervene on, but are irreducible to brain states (just one example: Davidson 1970). We could think of the history of science in a similar nonreductive monist manner.

In fact, we can take this analogy further. One could argue that the best way of analyzing mental phenomena is to look at both the levels of beliefs, desires, attention, etc. and the level of patterns of neural activations and especially the way in which the two levels interact. But it is difficult to see how that is to be done if beliefs and desires are ontologically independent from patterns of neural activations. The same could be said of the best way of analyzing the history of science: it should look at both internal and external history. And here we have to agree with Lakatos again, whose dictum is that "both historians and philosophers of science must make the best of the critical interplay between internal and external factors" (Lakatos 1971, 122).

Bence Nanay

University of Antwerp
University of Cambridge

Notes

* I have written much of this material while I was Lakatos Research Fellow at the London School of Economics in 2008. I am grateful for the support of LSE as well as of John Worrall, Miklos Redei, and Alex Bellamy. I am also grateful for comments by Gábor Zemplén and the members of my seminar on this material at the Universty of British Columbia in 2009.

1. Robert Musil, *The Man without Qualities*, New York: Random House, 1995, 185.

2. I am fairly certain that Lakatos's idiosyncratic use of the terminology of internal and external history comes from Heinrich Wölfflin's *Kunstgeschichtliche Grundbegriffe* (1915), where this distinction applies to the history of art.

3. Note that some uses of the selectionist explanation of scientific change that do not help themselves to the concept of meme are also subject to the same criticism. See especially Stephen Toulmin's evolutionary account (Toulmin 1967, 1972 and esp. 1970, 560–64).

4. This may sound a lot like Professor Arnheim's provocative statement in Robert Musil's *The Man without Qualities* that "nothing irrational happens in the history of the world" (chapter 43). As we shall see, Lakatos's claim is significantly less crazy.

5. Lakatos reiterates this idea about relegating actual history to the footnotes three times in one page on Lakatos 1971, 106–107.

6. Lakatos does add that he did write history this way on one occasion, namely, in Lakatos 1963/1964, where his "purpose was to distill a methodological message from the history, rather than to write history itself." This may sound like a good excuse, but in Lakatos 1971 (footnote 64), he confesses that he also used this way of writing history in Lakatos 1970, 138, 140, 146. It is difficult to decide whether Lakatos's dismissal of his earlier footnote comment as a joke was genuine or merely a *post hoc* excuse.

7. The most detailed account of Lakatos's problems with Toulmin's evolutionary explanation is in Lakatos: "Toulmin's Wittgensteinian epicycles" (manuscript in the Lakatos archive, file number 8/4). For a shorter summary, see Lakatos 1976, 137–38.

8. See also his letter to Jon Cohen, who reviewed Toulmin's book in the *British Journal for the Philosophy of Science* (which Lakatos edited) in 1972. Lakatos here explicitly agrees with Cohen's criticism of Toulmin's Darwinism (Lakatos to Jon Cohen, October 22, 1972, Lakatos archive, file number 13/166).

References

Aunger, Robert 2002. *The Electric Meme: A New Theory of How We Think and Communicate*, New York: Free Press.

Bechtel, William 1988. "New Insights into the Nature of Science: What Does Hull's Evolutionary Epistemology Teach Us," *Biology and Philosophy* 3: 157–64.

Bowler, Peter J. 1983. *The Eclipse of Darwinism*, Baltimore: Johns Hopkins University Press.

———. 1988. *The Non-Darwinian Revolution: Reinterpreting a Historical Myth*, Baltimore: Johns Hopkins University Press.

———. 1990. *Charles Darwin: The Man and his Influence*, Cambridge University Press.

Bradie, Michael 1986. *Assessing Evolutionary Epistemology, Biology & Philosophy* 1: 401–59.

Brown, James Robert 1989. *The Rational and the Social*, London: Routledge.

Bulmer, Michael 2003. *Francis Galton: Pioneer of Heredity and Biometry*, Baltimore: Johns Hopkins University Press.

———. 2004. "Did Jenkin's Swamping Argument Invalidate Darwin's Theory of Natural Selection?" *British Journal for the History of Science* 37: 281–97.

Cohen, I.B. 1974. "History and the Philosophy of Science," in F. Suppe, ed., *The Structure of Scientific Theories*, Urbana: University of Illinois Press, 308–49.

Cookson, G. and C. A. Hempstead 2000. *A Victorian Scientist and Engineer: Fleeming Jenkin and the Birth of Electrical Engineering*, Aldershot: Ashgate.

Darwin, Charles 1872. *The Origin of Species*, London: John Murray.

Darwin, F and A.C. Seward 1903. *More Letters of Charles Darwin*, 2 vols., London: Murray.

Davidson, Donald 1970. "Mental Events," in L. Foster and J. Swanson, eds., *Experience and Theory*, Amherst, MA: University of Massachusetts Press, 79–101, reprinted in *Donald Davidson, Essays on Actions and Events*, Oxford: Clarendon Press, 1980, 207–25.

Dawkins, Richard 1976/1989. *The Selfish Gene*, Second edition, Oxford University Press, Oxford.

———. 1982a. *The Extended Phenotype*, W.H. Freeman, Oxford.

———. 1982b. "Replicators and Vehicles," reprinted in Brandon, R.N. and Burian, R.M. eds., *Genes, Organisms, Populations: Controversies over the Units of Selection*, Cambridge, MA: The MIT Press, 1984.

———. 1983. "Universal Darwinism," in Bendall, D.S., ed., *Evolution from Molecules to Man*, Cambridge: Cambridge University Press, 403–25.

Dennett, Daniel C. 1995. *Darwin's Dangerous Idea*, New York: Touchstone.

———. 2003. *Freedom Evolves*. New York: Viking, 2003.

———. 2006. *Breaking the Spell: Religion as a Natural Phenomenon*. New York: Viking.

Distin, Kate 2005. *The Selfish Meme*, Cambridge: Cambridge University Press.

Elkana, Y. 1974. "Boltzmann's Scientific Research Programme and its Alternatives," in Y. Elkana, ed. *The Interaction between Science and Philosophy*, New York: Free Press, 242–97.

Ghiselin, Michael T. 1988. "Science as a Bioeconomic System," *Biology and Philosophy* 3: 177–78.

Godfrey-Smith, Peter 2003. *Theory and Reality*. Chicago: University of Chicago Press.

Gould, Stephen Jay 1991. "Fleeming Jenkin Revisited," in *Bully for Brontosaurus*. New York: W.W. Norton, 340–53.

Griesemer, James R. 1988. "Genes, Memes and Demes," *Biology and Philosophy* 3: 179–84.

Hacking, Ian 1979. "Imre Lakatos's Philosophy of Science," *British Journal for the Philosophy of Science* 30: 381–410.

———. 1983. *Representing and Intervening*, Cambridge: Cambridge University Press.
Holton, Gerald J. 1974. "On Being Caught between Dionysians and Apollonians," *Daedalus* 103 (3): 65–81.
———. 1978. *The Scientific Imagination: Case Studies*, Cambridge: Cambridge University Press.
Hull, David L. 1973. *Darwin and His Critics*, Cambridge: Cambridge University Press.
———. 1988. "A Mechanism and its Metaphysics: An Evolutionary Account of the Social and Conceptual Development of Science," *Biology and Philosophy* 3: 123–55.
———. 2001. *Science and Selection*, Cambridge University Press, Cambridge.
Jenkin, Fleeming 1867. "The Origin of Species," *The North British Review* 46: 277–318.
Kitcher, Philip 2007. *Living with Darwin*, New York: Oxford University Press.
Koertge, Noretta 1976. "Rational Reconstruction," in Robert S. Cohen, Paul K. Feyerabend and Marx W. Wartofsky, eds., *Essays in Memory of Imre Lakatos*. Dordrecht: Reidel, 359–69.
Kuhn, Thomas 1970. "Reflections on My Critics," in I. Lakatos and A. Musgrave eds., *Criticism and the Growth of Knowledge*, Cambridge: Cambridge University Press, 237–78.
———. 1971. "Notes on Lakatos," *Boston Studies in the Philosophy of Science* 8: 137–46.
Kulka, Tomas 1977. "Some Problems Concerning Rational Reconstruction: Comments on Elkana and Lakatos," *British Journal for the Philosophy of Science* 28: 325–44.
Lakatos, Imre 1963/1964. "Proofs and Refutations," *British Journal for the Philosophy of Science* 14: 1–25, 120–39, 221–43, 296–342.
———. 1970. "Falsification and the Methodology of Scientific Research Programmes," in I. Lakatos and A. Musgrave, eds., *Criticism and the Growth of Knowledge*, Cambridge: Cambridge University Press, 91–195.
———. 1971. "History of Science and its Rational Reconstruction," *Boston Studies in the Philosophy of Science* 8: 91–136.
———. 1976. "Understanding Toulmin," *Minerva* 14: 126–43.
———. 1976/1978. "A Postscript on History of Science and its Rational Reconstruction," in John Worrall and Gregory Currie, eds., *Imre Lakatos: The Methodology of Scientific Research Programmes*, Cambridge: Cambridge University Press, 1978, 189–92.
Lakatos, Imre and E. Zahar 1976/1978. "Why Did Copernicus's Research Program Supersede Ptolemy's?" in John Worrall and Gregory Currie, eds., *Imre Lakatos: The Methodology of Scientific Research Programmes*, Cambridge: Cambridge University Press, 1978, 168–88.
Latour, Bruno and Steve Woolgar 1979. *Laboratory Life*, Beverly Hills, CA: Sage Publications.
Laudan, Larry 1977. *Progress and Its Problems*, Berkeley: University of California Press.
Lewontin, R.C. 1986. "How Important is Genetics for an Understanding of Evolution?" *American Zoologist* 26: 811–20.
Mayr, Ernst 1959/1994. "Typological versus Population Thinking.," in B.J. Meggers, ed., *Evolution and Anthropology*. Washington: The Anthropological Society of America, 409–12. Reprinted in Elliott Sober, ed., *Conceptual Issues in Evolutionary Biology*, Cambridge, MA: The MIT Press, 1994, 325–28.
———. 1982. *The Growth of Biological Thought*, Cambridge, MA: Harvard University Press.
McMullin, E. 1970. "The History and Philosophy of Science: A Taxonomy," *Minnesota Studies in the Philosophy of Science* 5: 12–67.
Morris, S.W. 1994. "Fleeming Jenkin and The Origin of Species: A Reassessment," *British Journal for the History of Science* 27: 313–43.

Musgrave, Alan 1983. “Facts and Values in Science Studies,” in Home, Roderick Weir, ed., *Science under Scrutiny: The Place of History and Philosophy of Science*, Dordrecht: Reidel, 49–80.
Nanay, Bence forthcoming a. “Population Thinking as Trope Nominalism,” *Synthese*.
———. forthcoming b. “Popper’s Darwinian Analogy,” *Perspectives on Science*.
Palmer, E. 1993. “Lakatos’s ‘Internal History’ as Historiography,” *Perspectives on Science* 1: 603–26.
Popper, Karl R. 1959/2002. *The Logic of Scientific Discovery*, London Routledge.
———. 1963. *Conjectures and Refutations*, London Routledge.
———. 1972. *Objective Knowledge*, Oxford: Clarendon.
———. 1975/1996. “The Rationality of Scientific Revolutions,” in R. Harré, ed., *Problems of Scientific Revolutions*, Oxford: Clarendon, 72–101. Reprinted in Popper, Karl, R.: *The Myth of the Framework*. London: Routledge, 1996.
———. 1978. “Natural Selection and the Emergence of Mind,” *Dialectica* 32: 339–55.
Richardson, Alan 2006. “Rational Reconstruction,” in *The Philosophy of Science: An Encyclopedia*, Vol. 2 of 2., ed. Sahotra Sarkar and Jessica Pfeifer, London: Routledge, 681–85.
Richerson, Peter J. and Robert Boyd 2005. *Not by Genes Alone, How Culture Transformed Human Evolution*, Chicago: University of Chicago Press.
Shapin, Steven 1982. “The History of Science and its Sociological Reconstruction,” *History of Science* 20: 157–211.
———. 1992. “Discipline and Bounding: The History and Sociology of Science as Seen through the Externalism-Internalism Debate,” *History of Science* 30: 333–69.
Shapin, Steven and Simon Schaffer 1985. *Leviathan and the Air Pump*, Princeton: Princeton University Press.
Sober, Elliott 1980. “Evolution, Population Thinking, and Essentialism,” *Philosophy of Science* 47: 350–83.
Sperber, Dan 2000. “An Objection to the Memetic Approach to Culture,” in Robert Aunger, ed., *Darwinizing Culture*. *The Status of Memetics as a Science*. Oxford: Oxford University Press.
Sterelny, Kim 2006a. “The Evolution and Evolvability of Culture,” *Mind & Language* 21: 137–65.
———. 2006b. “Memes Revisited,” *British Journal for the Philosophy of Science* 57: 145–65.
Toulmin, Stephen 1967. “The Evolutionary Development of Natural Science,” *American Scientist* 55: 456–71.
———. 1970. “From Logical Systems to Conceptual Populations,” *Boston Studies in the Philosophy of Science* 8: 552–64.
———. 1972. *Human Understanding*, Oxford: Clarendon.
Van Fraassen, Bas 1980. *The Scientific Image*, Oxford: Oxford University Press.
Vorzimmer, P. 1963. “Charles Darwin and Blending Inheritance,” *Isis* 54: 371–90.
———. 1970. *Charles Darwin, The Years of Controversy: The Origin of Species and Its Critics*, 1859–1882, Philadelphia: Temple University Press.
Wimsatt, W.C. 1999. “Genes, Memes, and Cultural Heredity,” *Biology and Philosophy* 14: 279–310.