

COMPETING EXPLANATIONS

Exclusion and Importance in Historical Accounts

Mark Andrew Day

Submitted for the degree of PhD, November 2002

Department of Philosophy, University of Sheffield

Summary

This thesis is a philosophical examination of the nature of explanatory competition between historical accounts. It is usual for a philosophy of explanation to attempt an analysis of explanation, singular. The focus of this work is on *relations between two or more explanations*. In particular, I investigate explanatory exclusion and explanatory importance.

The methodology of the thesis is built upon a conception of descriptive philosophy. I believe that when attempting to philosophise about a practice such as history, we should pay detailed attention to existing good practice. To this end, I develop my conclusions in conjunction with an examination of eight differing explanations of the French Revolution.

Explanatory exclusion should be analysed in terms of incompatibility between explanations. Explananda, explanans, or relevance claims may be incompatible. Exclusion of the last type requires a commitment to explanatory realism, which holds that explanatory relationships mirror appropriate ontic – paradigmatically, causal – relationships. There are different types of historical explanation, yet all make reference to the causal history of the French Revolution, and therefore are candidates for explanatory exclusion. Causal attribution is necessary for historical explanation, but not sufficient. Causes must also be described in the ‘correct’ manner, and differing conceptions of ‘correctness’ lead to division between ‘types’ of explanation.

Historical explanations may *compete*, even where they do not exclude. One significant dimension of competition concerns attribution of historical importance. In order to allow substantive explanatory competition over this feature, I develop a realist analysis of historical importance. A more important cause is one which made more of a difference to the effect. In explicating this *counterfactual* claim, I defend an account of counterfactual decidability based on the idea of counterfactuals as implied experiments. This account is shown to have advantages over traditional metalinguistic and possible worlds analyses of counterfactuals.

Preface

Over the course of researching and writing this thesis George Botterill has been everything one could hope for from a supervisor. George has consistently been generous with his time and enthusiastic in his encouragement. Whilst I directly acknowledge a few particularly clear suggestions of his, many more ideas in this thesis have been positively influenced through discussion with or comment from George.

I have also been the beneficiary of the supportive and lively Sheffield philosophy department in general. In particular, Robert Stern, Steven Makin and Rosanna Keefe have been happy to discuss drafts of elements of this thesis, and have contributed useful comments.

I am very grateful to both Jacky Tivers and Kim Langford for performing the difficult job of proof reading this thesis. Any remaining errors are, of course, my responsibility. Kim has also been wonderful throughout: in discussing ideas in the early stages, and in helping me get through it towards the end.

Finally, I would like to acknowledge the financial support provided by the Arts and Humanities Research Board in the form of a three year award covering tuition fees and maintenance.

Contents

	Introduction	1
PART I	Chapter 1: Description and Prescription	5
	Chapter 2: Histories of the French Revolution	36
	Chapter 3: Explanation	65
PART II	Chapter 4: Principles of Explanatory Exclusion	98
	Chapter 5: Target and Contrast	125
	Chapter 6: Theory and Concept	151
PART III	Chapter 7: Importance	190
	Chapter 8: The Counterfactual Model	218
	Chapter 9: The Counterfactual Model: Defence	239
	Bibliography	263

Introduction

What is the relationship between multiple historical explanations? In most areas of historical scholarship we are faced with an array of different sorts of explanation: including, but not limited to, Marxist, economic, biographical, cultural, and intellectual. Deciding between these – even knowing whether we have to decide between them – can be extremely tricky. The task of this thesis is to investigate the notion of competition between historical explanations. Its distinctive feature is to focus not primarily on the question of what ‘explanation’ (singular) should be taken to mean, but on the relations between two or more explanations.

In particular, we want to know whether multiple historical explanations cumulatively enrich our understanding of the series of events being explained. Or is it the case that the acceptance of one explanation rules out acceptance of all others which have a similar target? Perhaps the situation is more equivocal, explanations ‘overlapping’ in some sense; in which case, we need to specify the conditions for this circumstance. These are questions concerning explanatory *exclusion*, one dimension of explanatory competition in general.

Even where explanations do not exclude one another (and perhaps also where they do), they can often be judged as *better or worse*. Are judgements of relative explanatory goodness limited to the claim that one explanation is factually more accurate than the other? If not, what further issues are relevant? Further, are such judgements of explanatory goodness anything more than the product of a particular historian’s viewpoint, incapable of objective defence? In particular, what should we make of the virtue of focusing one’s explanation on the historically *important* events?

I believe that the above questions lead to highly practical difficulties for students of history in understanding historical progress, controversy and competition. My aims are both practical and theoretical: my hope is that, subsequent to reading this thesis,

the reader may be better able to comprehend the relationship between historical works with which they are familiar, and better able to comprehend the theoretical issues underpinning explanatory comparison. Philosophers of history have had very little to say directly on explanatory competition, but the issue seems to be one where the application of philosophical thinking can bring beneficial results.

There is only one philosopher of history or the social sciences, that I know of, who has explicitly investigated the nature of relations between explanations in the above sense. Alan Garfinkel outlines the project thus:

‘... the first task we might set for a philosophy of explanation is that it give us some account of these conflicts [between explanations], complementarities, overlaps, and displacements, that it give, as it were, an elementary algebra of explanations. Its purpose would be to tell us when they can be added together and when they must be subtracted from one another.’¹

There are, however, other areas of philosophical debate which tackle issues analogous to those of this thesis. I shall mention three. First, there is the long-standing debate between methodological individualism and methodological socialism, concerning the possibility of reducing sociology to psychology. The debate focuses, primarily, on the relationship between explanations at different ontological ‘levels’: the social, and the psychological. The key question is whether explanations at one level exclude those at another, and if so, what the criteria are for favouring one over the other.

Second, those debating the nature of mental causation raise very similar issues; indeed, in what follows, I shall draw on ideas developed in this philosophical field. The core question is whether causal neuro-physiological explanation excludes a genuinely mental explanation (in particular, one in terms of reasons) of the same act. For the physicalist, a physiological explanation seems to be available, in principle, for the explanation of any given action. Therefore, unless we can find space for mental explanation in addition to the physiological, a developed neuro-science will eventually make genuinely mental explanation redundant: presumably not an outcome we should accept lightly.

¹ Alan Garfinkel, *Forms of Explanation: Rethinking the Questions in Social Theory* (New Haven, London: Yale University Press, 1981), 4

A final analogue to the question of historical explanatory competition can be found in the philosophy of religion (it was, in fact, this form of the problem which first prompted my investigation of the analogous historiographical question). We can conceive of a physicalist argument against the existence of God, which we might call the ‘argument from alternative explanation’. Roughly, this would state that no matter what evidence there might appear to be for the existence of God, an alternative physicalist explanation for those same phenomena will always (in principle) be possible. Furthermore, it would be claimed that this physicalist explanation excludes the theistic one. To assess this argument would be to raise very similar issues to those of this thesis.

Like many recent philosophers of history (and philosophers of science), I am concerned by the traditional detachment of philosophical analysis and descriptive knowledge of, in this case, historical practice. To remedy this, my approach is explicitly based upon a case study comprising eight historical explanations. In chapter one I provide the (meta-)philosophical background to this case study. I show why a strongly descriptive bias to philosophical enquiry is both desirable and necessary. The descriptive approach, however, raises a number of tricky questions, concerning the precise role that a case study should play in a *philosophical* enquiry.

For the descriptive approach to be of real worth, the works selected for the case study should be both varied and historically influential. In addition, the philosophical treatment should not do violence to these works by treating them in too simple or homogenous a manner. Following these guidelines, in chapter two I introduce the eight quite different explanations of the French Revolution. In chapter three I complete the introductory work of Part I, in presentation of an overview of approaches to historical explanation. In a novel categorisation, I argue that theories of explanation should be understood as being erotetic, cognitive, ontic, linguistic or contrastive. My own approach draws on at least three of these elements.

The core topic of Part II is explanatory exclusion. In chapter four I develop three fundamental principles of explanatory exclusion, which provide the basis for the remainder of the study. The material of chapter five extends the understanding and

application of these principles, through examination of the role of explanatory targets in shaping exclusion. A theory of contrastive explanation is best placed to aid understanding of explanatory targets, and how targets affect explanatory competition. In chapter six I take up the question of whether there are different *types* of explanation represented in the case study accounts, and to the extent that there are, how this affects explanatory exclusion and competition. In this chapter, I attempt to remain faithful to the widely varying content, style and theoretical approach of the different explanations; whilst arguing that they are commensurable, and therefore may be partially or fully exclusive.

The material of Part III concerns explanatory competition over and above explanatory exclusion. Chapter seven begins with an overview of the nature of historical explanatory virtues, and their role in explanatory competition. The main focus of Part III, however, is on one particular explanatory virtue: *importance*. In chapter seven, I argue, in general, that we need a realist understanding of that term. The material of chapters eight and nine develops such an understanding in detail. The counterfactual model of causal importance developed would, if correct, entail that at least some, and possibly a great many, judgements of importance in history are objectively decidable. This would be one substantive way for one historical explanation to be better than another, over and above being more factually accurate.

There is no dedicated final chapter for ‘conclusions’. Instead, I have collected the main findings of each of Part II and III at the end of those respective Parts.

My philosophical approach is guided by two goals, which can, without care, lead to mutual tension. I want to provide a treatment of the key terms ‘explanatory exclusion’, ‘competition’, ‘target’, ‘relevance’, ‘explanatory type’, ‘explanatory virtue’, and ‘importance’ which is satisfactory with respect to the practice of those historians I consider. But I also wish to provide substantive philosophical answers to the problem of explanatory competition, rejecting a relativism to which a thoroughly descriptive approach perhaps naturally leads. I hope that my conclusions can thereby be both relevant and useful to historians and students of history.

Part I

Chapter One: Description and Prescription

During the formative years of analytic philosophy of history, the topic appeared, somewhat perversely, to be a rather a-historical one. Taking its methodological lead from the logical positivist approach, the discipline was concerned with the conceptual analysis of concepts which it was thought were of relevance to historical study. Analysis, in this case, meant the task of listing the conditions under which the concept should properly be applied; conditions which are individually necessary, and jointly sufficient. In particular, such an approach was applied to the question of whether objectivity is possible in historical accounts, and whether historical explanation conforms to Hempel's Deductive-Nomological analysis (introduced in chapter three).

Such a methodology is a-historical because it assumes *a* timeless concept of (for example) historical explanation. It is assumed that there is only one proper way to understand the concept: an assumption which encouraged a tendency for the analysis to appear quite unrelated to the practice of actual historical enquiry. Of course, conceptual analysis was never intended to be entirely divorced from historical practice, given that the criteria for applying the concept had to guarantee at least an approximate fit with the intuitive extension of the concept. However, I think it is fair to say that the examination of historians' work came a very poor second to conceptual investigation.

Positivist-inspired philosophy of history certainly exhibited many virtues: of clarity, precision, and sometimes (as in Hempel's work) of providing fascinating suggestions which inspire further inquiry. Yet I would not be alone in tracing the cause of some of the main failings of analytic philosophy of history to the positivist approach. Most disturbingly, analytic philosophy of history has been almost universally ignored,

indeed derided as irrelevant, by working historians. Could this be due to the failure of philosophers of history to make their work connect in any meaningful way with historical practice?

Of course, the positivist inspired approach is waning. In the philosophy of history, we can trace the gradual change back to the 1970s, and in some cases even earlier: William Dray, for example, has always been concerned with the detailed use of historical work in his philosophical discussion. I hope that the following study can be a further step in that journey from an abstract analytic philosophy of history to a philosophy of history which actively engages with historical practice: a '*descriptive turn*'.

The evidence of a descriptive turn is also to be found in philosophy at large, and can there be traced further back (notably to the post-war Oxford school of Ordinary Language philosophy). Indeed, certain philosophical disciplines have already been revolutionised by a fresh descriptive approach to their subject matter. Kuhn and Lakatos, from the 1960s, stressed the importance of descriptive adequacy in philosophy of history's elder sibling, the philosophy of science. This movement became known as the 'historical turn', due to the role of the history of science in providing a descriptive basis for the philosophy of science. Rawls' influential theory of justice (developed from 1971) started with the description of ethical intuition, and proceeded via a method of *reflective equilibrium*. More recently (1990), Edward Craig has argued for the importance of a rather different kind of descriptive approach as a way of advancing the analysis of knowledge: one which emphasises the need to examine the *practical use* of the concept under analysis.

Yet, a descriptive turn is not currently overly popular amongst philosophers of history. The only recent attempt to examine historical explanation using an explicitly descriptive approach remains Raymond Martin's (1989) *The Past Within Us*¹. Furthermore, whilst there have been attempts to pin down, in general, what exactly is

¹ Raymond Martin, *The Past Within Us* (Princeton: Princeton University Press, 1989). It must, however, be said that the use of short case studies in philosophy of history is common practice, for example in the journal *History and Theory*

meant by ‘descriptive, or a posteriori, philosophising’, none has provided us with a clear statement of the problems to be solved and possible solutions; not one that can be straightforwardly be applied to the philosophy of history, in any case. We need to know what is meant by ‘description’ in this context, what precisely needs to be accounted for by the resultant philosophy, and how philosophy’s distinctively normative, or prescriptive, role fits with the descriptive. This is the justification for the inclusion of the following extended methodological (in other words, meta-philosophical) discussion: apart from the necessity of establishing my own methodology, I believe the general questions raised have wide applicability, and yet have been under-researched.

Why describe?

The traditional methodology of the philosophy of science was brought to task by Kuhn in his classic *The Structure of Scientific Revolutions*. Kuhn believed that philosophers of science had not paid sufficient attention to the history of science. To the extent that historical data were used, it was 'mainly to answer questions posed by the unhistorical stereotype drawn from science texts'². However, Kuhn argued that 'History, if viewed as a repository for more than anecdote or chronology, could produce a decisive transformation in the image of science by which we are now possessed'³

In short, a decent philosophy of science had to be true to an accurate, non-stereotypical description of science, discovered through historical study. Kuhn points out that the history of science had commonly been used to *exemplify* theory; anecdotes and examples were selected in order to support the philosophical theory. But Kuhn wanted to incorporate the history of science in more fundamental ways

² Thomas Kuhn, *The Structure of Scientific Revolutions* (Chicago: Chicago University Press, 1962), 1

³ *Ibid.*, 1

also: history could *suggest* philosophical approaches, and could be the *judge* of that philosophy.

Kuhn's methodological revolution, one of the most striking aspects of Kuhn's work, appears to have had its desired effect. Lakatos certainly took the suggestion seriously; his work is peppered with detailed historical case study. In general, the history and philosophy of science now cannot be easily separated: no philosophy of science worthy of the name can get by without the detailed backing of a history of science.

Just as Kuhn argued that the philosophy of science had to be more responsive to a rich and accurate history of science, I want to argue that the philosophy of history must be more responsive to a rich and accurate description of historical practice. In this section I will provide three reasons why I believe this. (An additional, more general, motivation is the previously mentioned desire to ensure that the philosophy of history is relevant to practising historians.) I will then outline the methodological questions we face in making good on the aim of producing a descriptively adequate philosophy of history.

Some knowledge of a practice is required even to speak of that practice: no philosophy of history is possible without knowing what 'history' is, and one must know something of what historians do to be able to understand the concept 'history'. This much is uncontentious: any theory of analysis requires that the philosophical understanding of key terms, such as 'history', be not too far removed from common understanding. What is open to debate is the degree to which idealisation is permitted. My general principle will be to examine historical work on its own terms, before philosophical idealisation is considered.

A descriptive approach entails two primary criteria, one positive and one negative. Each criterion also demonstrates a virtue of that approach. The positive is that *a philosophy should account for features of practice which are both pervasive and philosophically important* (in a sense which will be unpacked in the next section). Examining a practice in detail can lead to new and fruitful philosophical approaches. Kuhn's novel theories were achieved through the attempt to account for pervasive features of scientific practice which had been previously overlooked: the 'puzzle-

solving' activity of normal science. In the philosophy of history, one source of the interest generated by Hayden White's theory of narrative is that it attempts to account for previously overlooked features of historical writing: the rhetoric and linguistic devices employed.

The negative criterion states that *where philosophy conflicts with existing good practice, then that philosophy should be challenged and, except in extremely unusual situations, rejected*. This removes the support for 'a priori' philosophising, making philosophy and methodology more relevant to the practice. The negative criterion entails that practice should be used, in some sense, as the data against which the philosophical theory may be measured. Spelling out how existing good practice may be identified, and in what sense it is the 'data' of a resulting philosophy, will be the major task for the remainder of this chapter.

A third advantage of such a descriptive approach derives from the fact that practice must be well described in order to be able to then properly *criticise* that practice. In particular, it is vital to note that we cannot comment on the rationality or justification of any given practice without knowing the aims of that practice. For, as Laudan pointed out⁴, rationality is, at the very least, concerned with the proper fit between means and ends. I shall have more to say concerning criticism within a descriptive philosophy towards the end of this chapter.

The general form of question to be considered is how descriptive and prescriptive (normative) elements can be integrated into a philosophy. I sometimes frame the question in terms of the integration of descriptive and prescriptive into a *methodology*, a formulation I regard as equivalent. Whilst no philosophy is a complete methodology, in the sense of providing a complete 'rule-book' to guide practice, a philosophy does provide the fundamental methodological basis for practice.

To clarify the connections between the specific topic of this chapter and other 'descriptive' philosophical approaches, it will be helpful to present a typology of

⁴ Larry Laudan, 'Progress or Rationality? The Prospects for Normative Naturalism' (originally 1987) in *Philosophy of Science*, ed. David Papineau (Oxford: Oxford University Press, 1996), 201

3. **Descriptive object:** The practice of the discipline
- Resulting philosophy:** Methodology of that discipline
- Example:** Martin's philosophy of history

Similarly to type (1), practice is taken to provide the extension which philosophical rules must elucidate. It is important to make clear that 'practice', in this sense, should be taken as a shorthand for 'norms or intuitions which motivate that practice'. For what is described cannot literally be practice itself (the series of physical movements, or the output of words on a page); philosophical description is concerned with normative guidance.

- 3a. The practice of a discipline may be used to provide an analysis of concepts, such as 'explanation', in addition to providing a deep methodology. I shall say more in the final section of the relationship between conceptual analysis and methodology; contrary to Martin⁵, I do regard this as a close relationship.

My own approach is a variation of type (3): the use of a description of a practice to guide both the methodology of that practice, and analysis of relevant concepts. A possible difference concerns the fact that the focus of this thesis is upon relations between explanations; in that sense, my approach bears similarity to type (2).

The remainder of this chapter is devoted to a careful unpacking of the three main aspects of a descriptive approach: positively accounting for practice, negatively being constrained by that practice, and criticising that practice internally. We need to tackle such questions as: What is a good philosophical description of historical practice? How does a 'philosophical' description differ from, say, a historiographic or bibliographic description? What sorts of historical practice, and what aspects of that practice, should be philosophically described? What does it mean to allow a description to be the judge of a philosophical theory? Can we make room for 'bad' practice in the context of a descriptive philosophical approach?

⁵ In particular: Martin, *The Past Within Us*, 24-25

These are weighty questions, and deserve to be examined as much for their own interest as to ground the present work. We can make use of a variety of notions in tackling the above concerns: primarily Lakatos' 'Rational Reconstruction'; Goodman and Rawls' 'Reflective Equilibrium'; and Wittgenstein's 'family resemblance'. A subsidiary aim of this chapter will be to examine these notions, and ask of each how precisely they can aid our meta-philosophical questions.

Lakatos' theory of Rational Reconstruction

'Philosophy without history of science is empty; history of science without philosophy of science is blind'⁶

I shall first use a critical description of Lakatos' theory of Rational Reconstruction (henceforth abbreviated RR) to approach the above questions. RR is a meta-philosophical theory, the purpose of which is to make sense of a descriptive approach to philosophy of science (an approach of type (2) in the categorisation of the previous section).

I shall treat Lakatos' question: 'how should philosophy of science stand in relation to the history of science?' as structurally equivalent to my methodological question: 'how should philosophy of history stand in relation to the practice of history?' In what follows, I shall frequently use the term *practice* to designate scientific or historical practice, as discovered through examining the works directly, or through examining the history of that discipline. Lakatos' RR aims to provide an answer both to the question of what a philosophical description amounts to, and to the problem of how philosophies (or methodologies) can be judged on descriptive grounds.

⁶ Imre Lakatos, 'History of Science and its Rational Reconstructions' in *Method and Appraisal in the Physical Sciences*, ed. Colin Howson (London: Cambridge University Press, 1976), 1

What should a philosophy (of any given practice) be concerned with? The aspects of a practice can be readily divided into those that are *internal*, the features which a philosophy should concern itself with, and those which are *external* to philosophical concern.

For example, certain features of historical scholarship simply should not require philosophical attention: no attention need be given to the index layout of the historical study, neither to the preface, the length of the book, or the font type. More contentiously, such features as the ordering of the chapters, the language in which it is written, and metaphors used might also be regarded as outside philosophical attention. Consider another example: the philosophical examination of the historical development of science. In this case, changes in research funding resulting from a change in government, chance discoveries of new species, and resignations of key scientists might all be thought of as philosophically external.

On the other hand, internal aspects include evidential relations between observation and theory, explanatory relations between explanans and explanandum, and the relation between value judgements and factual judgements: these are the aspects of scientific or historical practice which are of proper philosophical interest. A philosophy which does not describe a practice using such internal concepts is judged not to be incorrect, but to have missed the point of what it is describing. To describe the history of science without using the ‘internal’ language of evidential support, theory change, falsification and verification, would be to do just this. Is there a general criterion which could distinguish between internal and external aspects? An approximate answer would be that internal features relate to the *rationality* of the practice. This makes sense, since it is generally agreed that the unique feature of philosophy as a whole is its focus on rationality or normativity.

Lakatos’ realisation was that the division between what is internal and what is external is dependent on the particular philosophy that is used to make sense of the practice in question. In other words, the demarcation is theory-dependent, where ‘theory’ in this context means ‘philosophical theory’. It is, therefore, the case that

judgements concerning the rationality of the practice in question are also theory-dependent.

The notion of the 'context of discovery', used in the philosophy of science, provides an example of this theory dependency. ('Context of discovery' refers to the actual development of a scientific theory, and should be understood by contrast with 'the context of justification': the extent to which a theory can be rationally justified or refuted.) Whilst all philosophies would aim to tackle the question of justification, whether the context of discovery is also of philosophical interest varies between philosophical theories. The inductivist maintains that where a scientific theory actually comes from is most important: only theories which are warranted by the proper support of pre-existing data should be asserted. The falsificationist gives the opposite answer: how a theory is arrived at is of no philosophical interest whatsoever. What matters is only how that theory is subsequently tested, and therefore justified.

Examples are not hard to find in the philosophy of history. If the philosophy states that explanation is a matter of providing general laws, or even of citing causes, then it is natural to suppose that the *way* these laws or causes are presented will not be of any philosophical interest. From this perspective, there is a clear split between the form of a historical account, which is philosophically irrelevant, and the content of the historical explanation (laws and causes), which is philosophically relevant and hence internal. Hayden White's theory of historical explanation through narrative emplotment (discussed in chapter three) takes the opposite perspective. To White, the form of the narrative – the way it is written – provides the type of explanation. 'Form' is therefore philosophically vital for White, and must be seen as an internal aspect of the historical practice. The way a narrative is constructed, the language used, and the metaphorical and other linguistic devices introduced are all of philosophical interest and importance.

Lakatos' meta-theory of rational reconstruction follows from the distinction between the internal and external elements of a practice. As we have seen, each methodology (philosophy) gives rise to a particular type of description. The philosophy demarcates the internal features of a practice, description of which provides the conceptual

backbone. For example, the core of a falsificationist history of science would consist of a history in terms of a series of scientific theories. It would assess how risky those theories were in relation to the background theory of the time, and it would describe tests of those theories and how the theories fared. A falsificationist philosophy of science therefore produces a certain description of scientific practice.

Such a philosophical description would be a ‘rational reconstruction’. It represents an idealised history of the subject, had the methodology been perfectly carried out, using only what is internal⁷. Despite the whiff of historical anachronism, Lakatos argues that a rational reconstruction is not all that different from any description, for the reason that all description – whether ‘rational reconstruction’ or otherwise – requires the interpretative backing of theory.

‘History without some bias is impossible ... The bias, of course, may be obscured by an eclectic variation of theories or by theoretical confusion: but neither eclecticism nor confusion amounts to an a-theoretical outlook⁸

The purpose of producing rational reconstructions is that these can then be tested, to see ‘how actual history “misbehaved” in the light of its rational reconstruction’⁹. This suggests two requirements. Firstly, that the rational reconstruction (and, therefore, philosophy) be empirically adequate. The relevant – internal – elements of the practice are to be treated literally as data, to be explained and accounted for by the philosophy. The second criterion distinguishes a normative-philosophical study from any other (biographical, sociological) study: philosophies should reconstruct scientific activity so that its rationality is apparent. This is what is meant, then, by a descriptive philosophy: one which describes practice accurately, and which does so in such a way as to make clear its rationality.

⁷ There is an ambiguity in this statement, as there is in Lakatos’ treatment. The weaker reading would demand that a rational reconstruction avoid contradicting internal elements; though it may also include external elements. The stronger reading would demand that a rational reconstruction use only internal elements. The weaker reading seems more plausible, for the reason that some external elements are necessary in a practice, even if no particular external feature is necessary. After all, the scientific theory must be arrived at *somehow*, the explanation must be presented in *some* form; even if it makes no philosophical difference which.

⁸ *Ibid.*, 19

⁹ *Ibid.*, 18

Internality and rationality

Lakatos' RR introduces two ideas which should be retained in some form. The first idea is that a practice can be divided into internal and external aspects, dependent upon the philosophy used. Only the internal need be philosophically accounted for. The second idea is that rational reconstructions should fit practice as far as possible, thereby providing a rational, yet descriptively adequate, account of the practice.

However, Lakatos' insistence that 'internal' and 'rational' can be entirely specified according to the particular philosophy used, leads to serious problems for a descriptive approach. I will argue that substantive modification of Lakatos' version of RR is necessary, if we are to follow the general intuitions which motivate a descriptive philosophical approach. There are four problems with RR, as it stands.

If the demarcation between internal and external is entirely dependent on the philosophy adopted, then there will be no room for a *positive* standard of descriptive adequacy. The descriptive standard will be, to some extent, self-confirming. For whilst, according to Lakatos' RR, it is possible to criticise a philosophy on the basis that it contradicts internal elements of practice, it is not possible to criticise a philosophy on the basis that it fails to account for internal elements. What a philosophy fails to account for is, for Lakatos, necessarily external. I would claim that it is possible to judge a philosophical theory to be inadequate on the basis of its failure to account for features of the practice which should be of genuine philosophical relevance; yet such a demand is, for Lakatos, simply confused. My positive criterion, suggested previously, has no place in Lakatos' RR¹⁰.

Very much related is the problem that positive *derivation* has no place in Lakatos' RR. A genuinely descriptive philosophy, of the sort adopted by Kuhn (and attempted in this thesis), is descriptive not only in justification, but also in derivation. By this I mean that a descriptive philosophy should develop philosophical questions, problems

¹⁰ We might add that whilst the rational reconstruction is overly theory-dependent, Lakatos' concept of 'actual history', against which the reconstruction is tested, seems naively theory-independent. In particular, it jars with Lakatos' claim (quoted previously) that 'history without some bias is impossible'.

and theories from an appreciation of the existing practice. Practice should be allowed to suggest and guide philosophy, as well as to judge it: indeed, this is the most conspicuous feature of the descriptive approaches listed previously (Kuhn, Rawls, White, Craig). In Lakatos' RR, a philosophy is arrived at a priori, is used to fashion a description in its image, and only then is confronted with the evidence.

The third criticism derives from Lakatos' demand that a good philosophy demonstrate the *rationality* of the practice in question. This demand is ambiguous: it could mean that a philosophy should demonstrate the rationality that the practice has, or that a philosophy should demonstrate that the practice has rationality. Due to his insistence that the given philosophy is sufficient to determine what is internal and rational, Lakatos adopts the latter reading.

The problem with this reading can be illustrated by considering Kuhn's philosophy of science. Kuhn's idea of science, in brief, was that theory change had more to do with the scientists' change in their way of seeing the world, than in a rational demonstration that the later theory was better than the first. Precisely because Kuhn's theory did not demonstrate the rationality of the scientific enterprise, Lakatos deemed Kuhn's ideas to be most interesting, but concerned with 'social psychology', not philosophy.¹¹ Yet Kuhn's work does have philosophical relevance: it is, after all, concerned with the core philosophical problems of theory change, and the rational justification of those changes. Simply because Kuhn gives a negative answer to the question of the rationality of scientific change should not rule him out of the game.

Hence Feyerabend's most insistent criticism of Lakatos' theory of Rational Reconstruction: that Lakatos has already decided, a priori, that the scientific enterprise *should* be reconstructed rationally, whilst other activities should not (Feyerabend's favourite examples are 'pseudo-sciences' such as Marxism, black magic, and astrology). In this respect also, the consequences of Lakatos' RR seem profoundly non-descriptive.

¹¹ Imre Lakatos, 'Falsification and the Methodology of Scientific Research Programmes' in *Criticism and the Growth of Knowledge*, eds. Imre Lakatos and Alan Musgrave (London: Cambridge University Press, 1970), 115, 178

In fact, there are episodes even of ‘science proper’ where Lakatos believes that the practice should not be rationally reconstructed; in other words, where the judgements of normative philosophy should over-ride the judgement delivered by an examination of practice. This may happen where that practice is a ‘degenerating tradition’ or ‘a bad new tradition’. Lakatos’ example of a ‘degenerating tradition’ is of certain facets of modern particle physics; his example of ‘a bad new tradition’ being ‘some of the main schools of modern sociology, psychology, and social psychology’¹². Whilst it certainly does appear sensible to allow that philosophies should not always bow to whatever happens to go on in practice, given Lakatos’ theory so far there appears little space for such a move. What is needed is a principled reason why a philosophy should sometimes be judged according to its fit with practice, whilst in other cases, it should be praised for condemning a practice. In other words, we need to have a closer look at those practices which are *worth* describing rationally, and those which should be rejected.

The final criticism is that, contrary to Lakatos, particular philosophies are insufficient to define what is internal and rational in a practice, because *general conceptions* of rationality are also necessary. Lakatos claims that a philosophy should demonstrate that (most) of the existing practice is rational, or justified. Yet, this demand is undermined by the realisation that the terms ‘rational’ and ‘justified’ are vague enough to be, without further specification, of little use.

‘Rationality’ may be taken as an instrumental term, such that it is sufficient to describe a rational action as one which is taken in accordance with aims and background beliefs and theory. In particular, an instrumental reading would not distinguish between proper or improper goals: a scientist who aimed at simply following a productive and stable career (Kuhn’s ‘normal scientist’!) could be perfectly rational. If we were to adopt such a notion of rationality, then a historical account which proceeded by charting a scientist’s/ historian’s reasons for acting could be deemed a ‘rational reconstruction’; as could a theory such as Kuhn’s.

¹² Lakatos, ‘History of Science and its Rational Reconstructions’, 36

Yet this way of looking at a practice is one which Lakatos wanted clearly to contrast with the philosophical standpoint, since it focuses on the apparently ‘external’ elements of a practice. On the other hand, ‘acting rationally’ may be considered a more demanding ascription, involving not just instrumental rationality, but the designation of proper goals or ends (for example, the goal of having one’s beliefs properly justified by the evidence). Further, the choice will not be simply between one notion or the other, but will allow of an answer on a spectrum between extremes.

Similarly, the notion of a ‘justified practice’ is one which may be interpreted in a more or less demanding way. Rational reconstruction amounts to the demand that a practice be presented as ‘justified’; but we can ask: justified according to the practitioner’s beliefs and theory, or according to the standards we hold to be correct? Here, too, a range of meanings of ‘justified’ is conceivable: from the extreme of only using our standards to justify a practice; to the other of allowing any beliefs of the agent to be taken as providing justificatory basis, including those which may appear bizarre or which were produced as a result of a-rational experience, say, childhood events. In this case, there would be no basis for the clear divide between philosophical questioning of a subject and intellectual history which someone such as Lakatos would desire.

It must be said that Lakatos was aware of the dependence of RR upon a particular conception of rationality¹³. The problem, however, is that the choice between general conceptions of rationality can not be justified in terms of RR. In addition, Lakatos’ own general conception weakens the desired descriptive basis for philosophy.

I would like to make space for a notion of positive descriptive adequacy by qualifying Lakatos’ derivation of the internal/ external demarcation as resulting entirely from the philosophy ‘under test’. Lakatos’ RR can make sense of the negative criterion that philosophies should not contradict those elements of a practice which are relevant (internal). But the insistence that the philosophy in question entirely determines both which elements are internal, and how these elements are to be rationally accounted

¹³ See Lakatos, ‘Falsification and the Methodology of Scientific Research Programmes’, 34, where he comments on Kuhn’s ‘poor theory of rationality’

for, is unmotivated. A methodology for a more comprehensively descriptive philosophy requires that we reject this insistence.

I would suggest, instead, that the content of 'internal' and 'rational' is derived from three sources. Firstly, the meaning is *partly* dependent upon the philosophy in question, which brings to the study its own conception of what is essential to a practice. The examples provided in the previous section demonstrate that Lakatos is at least partially correct in highlighting this dependence.

Secondly, the meaning is partly dependent upon wider conceptions of 'rationality', which may be more instrumental, or more substantive. I will subsequently argue that a descriptive approach fits better with a more instrumental understanding. But note that no understanding of 'rationality' is entirely without substance. Even the demand that one act in best accordance with one's own beliefs and desires is a substantive requirement, though it is much more permissive than other possible conceptions of rationality.

Finally, the meaning can be partly derived from the practice itself. What is internal, and what is rational, can be regarded as *immanent* (the antonym of transcendental) features of the practice. If it is possible to defend this claim, we are able to make sense of the possibility of judging philosophies to the extent that they fail to account for internal features: the positive criterion of descriptive adequacy.

The immanent sense of rationality may seem the most difficult to justify. In the next section I attempt to do just this, using the idea of Reflective Equilibrium (RE). Specifically, I aim to use RE to supplement RR in three ways. We need to know which practices should be rationally accountable; what descriptive *development* of philosophy amounts to; and how criticism should proceed from a descriptive approach. I shall begin with the question of how to identify 'good practice'.

‘Good practice’: Reflective Equilibrium

What sort of practice should be used as a descriptive basis for our subsequent philosophising? The question has some bite because (presumably) all would recognise that there exists *bad practice*, which should not be allowed to inform philosophy. Equivalently, we only want to pay heed to ‘good’ or ‘best’ practice. How are the terms within scare quotes to be fleshed out? We can approach the answer by first considering answers which lead nowhere good. Such answers are either *empty*, *prejudiced*, or *tacitly normative*.

First, the empty approach. I hope that it won’t be thought too unfair to illustrate this with reference to Martin’s position in *‘The Past Within Us’* (given that Martin’s *actual* use of history as case study seems to follow roughly the approach I will subsequently recommend):

‘... it would be just as surprising if the ways in which explanations are actually defended in the best historical work are not fairly close to the ways in which they should be defended. The best historical work, after all, is done by the best historians. And the best historians are those who are most competent to write history.’¹⁴

All well and good, but until we give some substance to ‘best’ - or even give some practical idea how to find the ‘best’ historical work or historians - the above statement is vacuous. To give this content we could plausibly turn in one of two directions. We could identify the ‘best’ historians with the ‘elite’ historians - those who have achieved a high standing within the discipline, have been widely read and cited, or can be said to match other similar sociological criteria. Lakatos’ demarcation is (usually) founded on such considerations: ‘if a demarcation criterion [between science and non-science] is inconsistent with the ‘basic’ appraisals of the scientific elite, it should be rejected’¹⁵

¹⁴ Martin, *The Past Within Us*, 28

¹⁵ Lakatos, ‘History of Science and its Rational Reconstructions’, 23. It should be noted that Lakatos goes on to disagree with this statement; but only to soften the idea by dispensing with a rigid understanding of ‘rejection’. The use of scientific *elite* opinions as ‘data’ is retained.

This, then, is the second approach, the *prejudiced* approach. For the concern with this approach is obvious – why should we assume that the best history is done by the historians who have achieved high standing within the historical community? To use Feyerabend's colourful criticism of Lakatos: 'all *he* can say in favour of these [methodological] rules is that the elite of some enterprise he loves *sometimes* sticks to the rules'¹⁶. What is needed, it seems, is a justification of the identity claimed between 'best' history and 'most sociologically successful' history.

Yet in providing such a justification, we reach the third answer – the *tacitly normative*. For, if we can provide reasons why we should pay heed to some historians and not others – where 'reasons' means the following of certain good methodological practices – we thereby render a descriptive approach unnecessary. If we already know the standard by which to judge history as 'good' or 'bad', why not stick to traditional 'a priori' philosophising, and dispense with the charade of deriving (in any sense) philosophy from description of practice?

We need a way of demarcating between works which derives from features which can plausibly be seen to point to the 'best' works, but which does not in so doing render the descriptive approach irrelevant. The method of *reflective equilibrium* meets this aim, since the key virtue of a practice is that it is 'reflective'; a property referring to the derivation of the practice, which therefore does not pre-judge the normative *content* of that practice.

The notion of reflective equilibrium was introduced by Goodman¹⁷ but was more extensively exemplified in Rawls' work¹⁸. Rawls' starting point for political theory is that of 'widely accepted but weak' premises. The purpose of Rawls' well known 'veil of ignorance' is as a means to extract ethical intuitions. (The 'veil of ignorance' disallows knowledge of yourself and your natural abilities, your position in society,

¹⁶ Paul Feyerabend, 'On the Critique of Scientific Reason', in *Method and Appraisal in the Physical Sciences*, ed. Howson, 315. (The 'sometimes' in this statement stems from the qualification pointed out in the previous footnote.)

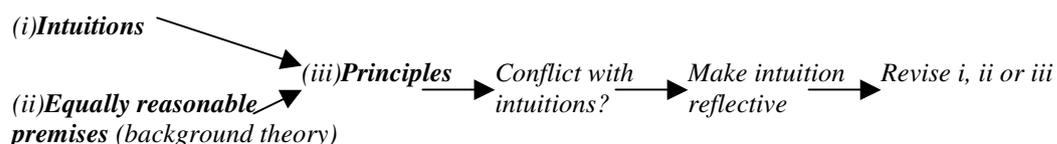
¹⁷ Nelson Goodman, *Fact, Fiction and Forecast* (Cambridge, Mass: Harvard University Press, 1983; originally 1954), 63-64

¹⁸ John Rawls, *A Theory of Justice* (Oxford: Oxford University Press, 1971), 18-19, 41-46

your sex, race, nationality, and individual tastes, in choosing a social ethics.) These intuitions should be logically developed (in conjunction with background theory) to yield a significant set of principles; if this is impossible, ‘equally reasonable’ premises should be selected to allow the principles to be derived. The principles, once developed, should then be tested against intuitive response. If we have no contrary intuitions, or if those intuitions are contradictory or confused, then the principles should be accepted as a good guide to the subject. However, where the intuitions do conflict, we have a more interesting situation: we must seek reflective equilibrium.

In this case, Rawls’ important realisation is that the situation we face is not simply that of a clash between basic intuition and developed theory. Rather, the aim is to arrive at a more considered intuition; an intuition tested through consideration of the competing principles, hence ‘reflective’. If our reflective intuition remains contrary to the theory, then intuition should prevail. However, if the intuition is, on reflection, found to be naïve, then this is clearly an advance. (It should be added that it is ideal at this point to find reasons as to why we initially found the intuition powerful.) Thus, under a methodology of ‘wide reflective equilibrium’ we start with intuition (and background theory); but by the end point it is possible that any of intuition, norms, or background theory be revised.

Diagrammatically:



The concept of reflective equilibrium provides an answer to the question of which historical works should be descriptively respected as a normative guide. Historical works should be subjected to reflection: such that the principles informing these works be taken as starting point, but may be altered through reflective criticism¹⁹.

¹⁹ Whilst the demand to investigate only reflective intuition and practice provides a philosophical answer to the demand for a descriptive starting point, it must be admitted that as a complete *practical* solution, it falls some way short. After all, there are a lot of historical explanations to choose from – even ‘reflective’ ones! In practice, I believe that we have to adopt something of the *prejudiced* approach, as I make clear in chapter two. However, there is still a difference between Lakatos’

I wish to stress two pertinent differences between Rawls' use of reflective equilibrium and the use to which I will seek to put it in providing a method for the philosophy of history (or, indeed, for the philosophy of science). First, isn't it the case that all historical works are already reflective, in a way that everyday moral intuitions are not? To some extent, this depends on what we mean by 'being reflective'; and that is the topic of the following section.

However, surely very little historical work, in contrast to ethical intuition, fails to meet the standards Rawls sets for initial intuition: that we should not include 'those judgements made with hesitation, or in which we have little confidence. Similarly, those given when upset or frightened, or when we stand to gain one way or the other can be left aside'²⁰. This is a difference deriving from the nature of the philosophical subject matter: in the earlier typology: type (3) or (2), rather than type (1). As a consideration of the suitability of RE, though, this consideration only seems to strengthen the belief that historical practice should guide philosophy.

Second, what is the precise nature of the starting point to the process of reflective equilibrium? Should we take one starting point: a basic, broad agreement of historical intuitions? Or allow many starting points, and therefore potentially different paths to reflective equilibrium, which may be inconsistent with each other? In other words, will there be *a* reflective equilibrium?

Whilst it is not feasible to take each individual historical explanation as a separate starting point, I do wish to allow that the direction taken by the process of reflective equilibrium can vary according to differences in historical practice. There *may* not be a single reflective equilibrium; a possibility which I believe a descriptive approach must at least allow.

Gordon Graham also takes the method of reflective equilibrium to be central to his philosophical examination of historical explanation, yet adopts the idea that there should be a single starting point, and hence a single reflective equilibrium. Graham

descriptive starting point and my own: I cannot investigate practice beyond the 'elite' due to lack of time and space; Lakatos would not investigate practice beyond the 'elite' on principle.

²⁰ *Ibid.*, 42

takes more literally Rawls' concern with 'widely accepted but weak' starting premises. As a result, Graham works with a universal, yet very weak, notion of 'historical explanation': simply that historical explanation of an 'event or state of affairs is to give its history'²¹.

Two convictions lead me to question an approach like Graham's. First, I am concerned that the subtleties of actual historical writing, the capturing of which provides one of the major motivations for initially adopting a descriptive approach, will be lost if we are to focus purely on 'widely accepted but weak' intuitions. Second, interpreting 'reflection' as immanent criticism raises the possibility (even, perhaps, the probability) that reflection will not lead to a single end point. I shall expand on this interpretation in the following section. We should be open to the empirical possibility that practice and intuition may not converge. Stich, for example, attempts just such an empirical demonstration²². He argues that the notion of descriptive conceptual analysis which assumes 'that our own case is typical and so [we] can generalise from it to others' is empirically false²³.

What is it to be reflective?

In defending a descriptive approach which relies on reflective equilibrium, it is important to know what 'reflection' amounts to in this context. For Rawls' purposes, it is plausible that he simply demand that intuition should be 'confronted' with wider theory. But for a more complex practice such as history, we need to say more.

²¹ Gordon Graham, *Historical Explanation Reconsidered* (Aberdeen: Aberdeen University Press, 1983), 43

²² Stephen Stich, 'Normativity and Epistemic Intuitions', presented to the University of Sheffield Philosophy Department Seminar on 26/4/02. See also Stephen Stich and Jonathan M. Weinberg, 'Jackson's Empirical Assumptions' in *Philosophy and Phenomenological Research* 62 (3) (2001): 637-643

²³ Stich and Weinberg, 640

My short answer is that ‘reflective history’ means ‘history criticised immanently’. The method of immanent criticism is to import very little to the position under criticism, using instead only the material already present. (I would use the more common phrase ‘internal criticism’, were it not for the fact that this term has already been used in a different sense previously.) Criticising a position in this manner is commonly thought of as a virtue: a more effective criticism of a position is one which demonstrates some inconsistency, a less effective criticism one which simply overrules the position according to some external standard. For example, to criticise a utilitarian theory of morality by attempting to demonstrate that the theory cannot deliver what it promises is more persuasive than simply confronting utilitarianism with an alternative moral theory. Plato’s dialogues are a classic example of immanent criticism; but here I intend to use two different examples to better exemplify the notion.

Hegel’s dialectical method has been described by McCarney as a process of *reflection*: ‘The basic form of such thinking is reflection. ... Intellectual reflection bounces thought, as it were, off its object and back into the self.’²⁴ Hegel’s philosophical criticism always begins from a concrete position, usually one which is intended to represent actual positions taken by past thinkers. This ‘reflective method’ of philosophising about history is well exemplified by Hegel’s criticism of historical method in his *Philosophy of World History*²⁵. In this section Hegel considers six historical methodologies, the last of which, Philosophical History, he adopts as his own. Whether or not Hegel succeeds in implementing a method of immanent criticism, the attempt is instructive. The criticism of each position begins with a description, and challenges each according to the aims and background theory of that position.

One example of these critiques concerns the type of history Hegel calls ‘reflective universal’. This mode of history appears in Hegel’s account as a result of criticism of ‘original’ history, exemplified in the writings of Herodotus and Thucydides. Hegel

²⁴ Joseph McCarney, *Hegel on History* (London: Routledge, 2000), 86

²⁵ Georg Wilhelm Friedrich Hegel, *Introduction to the Philosophy of History*, trans. Leo Rauch (Hackett, 1995; originally 1840), 3-11

claims that ‘original’ history was unduly limited to narration of events linked to the writer’s own personal experience. Hegel’s criticism of ‘reflective universal’ history is that such universal history misses the particularity, vitality, of the events which it describes. It attempts to be universal, yet must ‘make do with summaries and abridgements’²⁶. ‘Reflective universal’ history therefore fails to meet its own standards, and must be developed, as Hegel goes on to do.

Laudan’s philosophy of science demonstrates the method of immanent criticism in a more basic and straightforward way. One of Laudan’s objections to Lakatos’ Rational Reconstruction is that the idea of reconstructing a practice such that it is rational makes no sense without a knowledge of the aims and background beliefs of the practitioners. As we have seen, the weakest sense of ‘acting rationally’ is the instrumental sense, where to act rationally is to act in accordance with one’s aims and relevant beliefs. Stronger senses may add further requirements, but even with an instrumental understanding, knowledge of aims and background beliefs of a practice are necessary to be able to criticise that practice.

Laudan’s methodology amounts to a method of immanent criticism which takes seriously the practitioners’ aims and background theories. Methodological rules are to be framed not categorically – ‘one ought to construct explanations of type x’, but hypothetically – ‘if one’s aim is y, then one ought to construct explanations of type x’.

It has been my intention to delimit the types of practice which are of philosophical importance by equating ‘philosophical importance’ to the property of ‘being reflective’. In the case of a complex practice such as history, I have argued that ‘reflection’ can only mean immanent criticism; criticism using the relevant background of the practice itself. Of course, the nature of ‘immanent criticism’, and ‘reflective development’, cannot be entirely specified in advance of actually attempting the method in practice; but I hope the above examples have at least given a flavour of the sort of method which I envisage.

²⁶ *Ibid.*, 7

As a final note on immanent criticism, I want to return to my original statement of the position; in particular, the claim that in immanent criticism we import ‘very little’ to the position under criticism, relying primarily on elements already present. To see that any form of criticism must import *something* to the position being critiqued, consider two examples. An understanding of logical contradiction, and the presumption that any position which gives rise to self-contradiction should be modified, is perhaps the most basic assumption which must be made. Somewhat more demanding is the requirement that a practice be capable of being justified in some sense. As has been pointed out, even a weak notion of justification is to some extent substantive, and may therefore require the import of ideas foreign to the practice being criticised.

The question of what should be used to criticise a practice is equivalent to the issue, previously raised, of what we are to use in order to understand ‘rationality’ of practice. In answer to that question, I established the need for three different sources: philosophy-dependent, immanent, and general conceptions. The introduction of reflective equilibrium has allowed us to see how the immanent can be given priority, and can function effectively in the context of a descriptive philosophy. However, the bottom line is that the other sources are necessary; a purely immanent approach is not possible. How we are to precisely juggle these different sources cannot, of course, be precisely legislated. It is sufficient to follow the vague requirement that immanent development and criticism be preferred, where possible and where reasonable.

We can now summarise the ways that RE can be used to answer the concerns over RR, raised in the previous section. In the process of reflective equilibrium our starting point is given: that of existing practice or intuition. There is, therefore, no place for the dubious notion of an unbiased description (Lakatos’ ‘actual history’). The major requirement of RE is that philosophical development proceeds from existing practice or intuition; in the case of philosophy of history, this means existing historical practice, historical theory and the concerns which are already present in the work of the historians. This allows for the area of philosophical concern (the ‘internal’) to be partly given by existing practice; and only partly by the particular philosophical approach adopted. It also means, as a matter of course, that the

development of the resulting philosophy is in accordance with a positive criterion of descriptive adequacy.

Perhaps most importantly, RE provides an answer as to when practice should be used to guide philosophy, and when the reverse is true. Practice which is *reflective* should be rationally reconstructed, and used to test the philosophy. Being reflective means being subjected to internal criticism; therefore the proper place of criticism in a descriptive philosophy is also delineated²⁷. Historical practice should be investigated according to its own aims and theoretical practice (which may differ between historians), and should be seen as ‘justified’ to the extent that a coherent account can be given for the methods in use. There need be no demand that such practice conform to further standards not implicit in the practice itself.

Conceptual analysis

In this final section I want to connect the meta-methodology developed above with debate concerning the nature of philosophical analysis. Whilst at the start of the chapter I used the analytic approach of the positivists as the foil to my preferred ‘descriptive approach’, analysis clearly has some place in any philosophy. In addition, the analysis of concepts used in a practice such as history is tightly related to assessment of the methodology of that practice. I will first argue for this latter claim by criticising Martin’s argument to the contrary; I will then draw on a Wittgensteinian theory of philosophical analysis in order to complete the development of my own methodological position.

²⁷ In fact, Feyerabend suggests, a little cryptically, that he would be prepared to accept a theory of rational reconstruction modified in such a way: ‘these judgements [of the rationality of scientific practice] may be retained if we adopt a different theory of rationality, for example Hegel’s’: Feyerabend, ‘On the Critique of Scientific Reason’, 327

As I do, Martin wishes to develop and apply a descriptive approach to the philosophy of history. Yet, unlike me, Martin claims that it is possible to develop such a (meta-) methodology whilst agreeing with the positivist *analysis* of historical concepts. In particular, Martin accepts the positivist Deductive-Nomological (D-N) account of explanation in history, whilst stating that the acceptance of this analysis makes little difference to the kind of descriptive philosophy he (and I) favour²⁸. (The important feature of the D-N analysis for the present discussion is that a D-N explanation provides a set of conditions *sufficient* for the occurrence of the event to be explained.)

Yet Martin's subsequent investigation belies this claim. The negative claim derived from Martin's application of a descriptive meta-methodology is that (contrary to the D-N model) historians do not attempt to provide sufficient conditions for the occurrence of a historical event: for all realise the practical impossibility of such a task. Martin's positive claim is that historians argue for explanations (primarily) by showing one explanation to provide a *more nearly* sufficient set of conditions than another²⁹.

Thus we could say that a historical methodology arrived at by Martin's descriptive approach would demand roughly that 'historians should develop explanations which are comprised of a more complete sufficient set of conditions than are rival explanations of the same event'. The similarity to the D-N account is obvious; since the standard provided by the D-N account is retained, explanations are assessed to the extent that they reach the goal of providing a sufficient explanation. This close connection between historical method and analysis of historical concepts naturally follows from Martin's (non-descriptive) belief in the correctness of the D-N account as an analysis of 'explanation', since any methodology of explanation must take account of what 'explanation' really *is*.

Thus Martin's a priori acceptance of one particular analysis of 'explanation', the D-N analysis, has a major effect on the subsequent attempt to develop a descriptively adequate historical methodology. Martin's approach is insufficiently radical: we cannot afford to treat the process of analysis in a non-descriptive, a-historical manner,

²⁸ Martin, *The Past Within Us*, 127-128

²⁹ *Ibid.*, 40

on pain of rendering the remaining descriptive philosophy inconsequential. If we were to follow Martin's approach, the descriptive basis of philosophy could tweak practical suggestions, but would be barred from the investigation of deeper and more fundamental norms.

The 'standard' method (by which I mean the 'reductive analysis' exemplified in the work of Russell, Carnap, and Hempel) for analysing any given concept has been to list situations which intuitively satisfy the concept, and those which intuitively do not. A set of conditions is then formulated, which are individually necessary, and jointly sufficient for the application of the concept, such that the intuitive decisions are adhered to. The extent to which *all* intuitive examples should be so captured varies according to how disputed our intuitions are, and the premium which matching intuition is given compared to other analytical virtues such as simplicity and coherence.

In relation to this characterisation, one feature suggested by my account of philosophy developed in this chapter would be to place a premium on matching intuition, even where that intuition is divergent. Another is that we should pay more attention to the *development* of conceptual analyses, such that they be tied to practical use, as in the process of reflective equilibrium. The 'standard' model of analysis allows description to be relevant only in a negative sense: as a test which may or may not 'falsify' the analysis. It is in a similar position, therefore, to Lakatos' RR; and I would repeat my injunction that development, as well as justification, should be descriptively based.

My primary focus in what follows is, however, the following presupposition of the 'standard' model: the denial that there can be multiple analyses of a single concept. The idea that the analysis of the concept should be carried out in accordance with the philosophical aims of the analyser, not necessarily equivalent to those of the language user/ practitioner, leads naturally to this opinion.

As I suggested with regard to Graham's interpretation of reflective equilibrium, I think it likely that the analysis of a practice – and related, of a complex term featuring within that practice – will not find a single end point. 'Historical explanation' will be the concept most central to this thesis. Given the variety of ways that terms such as

this are used, the restriction that there should only be one set of conditions which exhaustively apply looks dubious (though, to repeat, whether there is unity or plurality can only be discovered through descriptive examination, not a priori.)

A related point concerning analysis specifically should be mentioned at this point. Necessity and sufficiency should be seen as extremes in our analysis, rather than being the only criteria that can feature in the analysis of a concept. Craig makes this point in envisaging the possibility that belief is a ‘major component’ in the analysis of *knowledge*, but not a necessary condition. Under the ‘standard’ model, Craig complained that

‘the conceptual accountant ... can try to talk about it [belief], just so long as the audience is prepared to listen to such periphra, as of something which very often accompanies knowledge. But when he is asked for the real outcome of the business, the analysis, anything not strictly a necessary condition vanishes without trace.’³⁰

I agree with Craig that such a restriction is implausible and unhelpful.

Though I am content that a descriptive philosophy sanctions the possibility of multiplicity in conceptual analysis, there is a danger in this possibility. What is needed is a means of analysis which allows us to group multiple applications of, say, ‘explanation’; such that, though they are not covered by a single set of criteria, there is a more than ‘accidental’ relationship which obtains between those applications. What we don’t want are multiple applications to be linked simply, by, for example *punning*. Different senses of ‘explanation’ must be analytically connected in a way that different senses of ‘party’ (political party, birthday party, parties to a contract) are not.

I believe the best model we have in the philosophical literature comes from Wittgenstein and his idea of ‘family resemblance’. The essence of the idea is that, for many complex concepts, there is no fixed set of criteria to capture the extension (application) of that concept. Instead, overlapping sets of criteria must be used. Diagrammatically:

³⁰ Edward Craig, *Knowledge and the State of Nature* (Oxford: Clarendon Press, 1990), 14



Applications a, b and c all fall under the description of being examples of concept X. Yet the uses of the concept have no feature common to them all: there is therefore no essence to the concept. The uses are, however, similar; they share enough properties to be considered ‘of the same family’.

Wittgenstein’s core idea is simple, but has radical consequences. I take its chief advantage to be its ability to dispense with the need for a fixed set of necessary and sufficient conditions; yet, through the notion of *similarity*, its ability to show how different uses of the same concept can be linked in a tighter way than just punning. Wittgenstein describes the relation between different uses of the same term as exemplifying ‘a complicated network of similarities overlapping and criss-crossing: sometimes overall similarities, sometimes similarities of detail’³¹.

Concerns raised in the literature over the idea of family resemblance centre on specifying precisely how Wittgenstein’s theory avoids the danger of collapsing into disjunctivism. In the terms of our simple example: why isn’t the analysis of concept X simply the disjunctive ‘X = AB or BC or AC’? Whilst Pelczar³² and Bellaimy³³ differ somewhat in their answers, both agree that the key to understanding the nature of family resemblance must lie not only in an appreciation of similarity, but also in attending to the way in which the concepts are *used*. In particular, Pelczar argues that the distinctiveness of family resemblance concepts lies in the fact that the application of these concepts is semantically open. This means that discretion is permitted in applying family resemblance concept X, such that over time the application of X is widened to encompass different items. X’s extension is widened by language users in ways not determined by the previous use of the concept to other situations which are deemed similar in a respect relevant to previous uses.

³¹ Ludwig Wittgenstein, *Philosophical Investigations* (Oxford: Blackwell, 1953), section 66

³² Michael Pelczar, ‘Wittgenstinian Semantics’ in *Nous* 34 (4) (2000): 483-516

³³ James E. Bellaimy, ‘Family Resemblances and the Problem of the Under-determination of Extension’ in *Philosophical Investigations* 13 (1) (1990): 31-43

Family resemblance allows us to understand a conceptual analysis which leads to related but non-identical final analyses. It allows multiplicity of analysis without that thereby implying simple disjunction. The way that this is achieved relies on *similarity* of application and *use* of concept; the latter notion provides a strong link with my previous remarks concerning the development of a methodology using reflective equilibrium. To repeat: philosophy as (deep) methodology, and philosophy as conceptual analysis, are two sides of the same coin; we should expect similar considerations to apply to each.

Practical Conclusions

I wish to draw together the issues discussed in this chapter by summarising the practical consequences for the subsequent study, and for philosophy of history in general, in terms of four broad points.

The initial intuition was that philosophies of history should be judged according to two criteria. The positive is that pervasive and philosophically important ('internal') features of a practice should be accounted for, by investigating the rationality of those features. The negative is that where the philosophy conflicts with existing good practice, it should be revised. Each of these contributes to making precise the demand that a philosophy 'fit' with historical practice. In this thesis, these criteria will be used in order to discern the nature of explanatory exclusion (chapter four), explananda (chapter five), explanatory theory (chapter six), and historical importance (chapter seven). It is vital to use rich, full examples of historical practice, so as not to miss out interesting or pervasive features.

Secondly, we have learnt from Lakatos' division between internal and external elements of a practice that not all aspects of historical practice need be philosophically accounted for. Where the divide is placed is, in part, dependent upon the philosophy used. However, I argued that what is treated as 'internal' cannot be

entirely theory dependent; indeed, the positive criterion conflicts with such a notion of extreme theory dependence, as we have seen. What is internal is partly given by the pre-existing concerns of the practitioners': their theories, problems and questions may guide philosophical interest. Indeed, in a thorough-going descriptive approach, it may be argued that the latter meaning of 'internal' should be given precedence over the former.

Thirdly, I have argued for the importance of only considering reflective practice; in other words, that practice which has been subjected to immanent or internal criticism. This is how 'good practice' is to be understood. In addition, we can see that criticism of practice should begin from a position which tolerates the underpinning methodology: a case of 'innocent until proved guilty'. Subsequent criticism should import as little to the position being criticised as is reasonable.

Finally, I have pointed out that the fate of *methodologies of history* and of *analyses of concepts used in historical writing* are inter-twined. Analysis, like the development of a methodology for history, must proceed from descriptive and pluralist assumptions. Indeed, the subject matter of this thesis can be put in two ways: as an investigation into the practice of explanatory improvement and competition in historical writing, or as the *analysis* of the concept of 'explanatory competition between historical explanations'.

Chapter Two: Histories of the French Revolution

To supplement the meta-philosophical investigation carried out in chapter one, I shall in this section consider more practical questions. How were the historical works for the case study selected? What exactly will I be looking for in these works, and what role will they play? After answering these questions, the bulk of this chapter consists of an overview of the historical explanations selected. It is unlikely that the reader will have much, if any, knowledge of the historical works in question; some basic introduction to each is essential in order to fully understand the uses to which they will be put.

Selection and role of case studies

My overall task is to examine relations of competition between different historical explanations. It clearly makes sense to choose explanations which have similar targets; for it is such explanations that, in practice, we are eager to compare and rank. The origins of the French Revolution was selected as such a topic for two reasons. Firstly, as one of the most discussed questions in the history of historical writing, an enormous range of historical approaches, theoretical positions and modes of explanation can be found in the literature. In no other historical field is the question 'how should we compare all these proffered explanations?' as obvious, or as pressing. Secondly, there is the simple reason that I already had an acquaintance with the topic from my time as a history student.

So how were the particular histories of the French Revolution selected? I tackled this question philosophically in chapter one, though I also noted there that no selection can be defended entirely objectively. I do not attempt a statistical study of a wide range

of histories; for it is my belief that the importance of the explanations is frequently in their details. Instead I have selected only eight explanations; a number which can hardly be claimed to be fully representative of the thousands of books and articles which have been published in the two-hundred-plus years since the Revolution. Inevitably my selection in part reflects my interests, which books I have been introduced to, which happen to have caught my attention, and so forth.

Excuses notwithstanding, I have attempted to follow two flexible criteria in the selection of historical explanations. The first is that the works selected should be well respected by the contemporary writers' historical peers, or by subsequent historians, as appropriate. In general, all the works have had a broad influence on historical scholarship. The reason for this restriction is to ensure that my conclusions themselves have the broadest possible interest and pertinence.

The second criterion is that the explanations be, to some extent, representative of the array of approaches to the French Revolution adopted by historians. Thus, I have included approaches which emphasise the social, political, geo-political, intellectual and cultural strands of French Revolutionary origins. I have included writers who emphasise the deep causes of the Revolution, and others who attack the idea that the Revolution has such structural underpinnings. The approaches range from the (social) scientific to those which disavow any kind of scientific search for causes or laws. My hope is that, for any influential approach to the origins of the French Revolution which has been adopted, there is an explanation considered here which is a not too distant relation.

One criterion I have *not* tried to consciously employ is that the explanations given well exemplify the philosophical points I wish to make. It is clear that I could have found works which better demonstrated, say, the use of contrast (chapter five), or counterfactuals (chapters eight and nine). But that would have been to mistake the purpose of this exercise purely as one of *exemplification*; a restriction which runs counter to my methodological motivation (and my argument of chapter one). I endeavoured to select works without first ascertaining whether their theoretical background fitted neatly with my own. Instead, where clear exemplification *is* needed, I have not hesitated to turn to examples outside my eight case studies.

One final point: the topic of my case-studies are explanations; not historians, nor books. (Although, of course, you can't find historical explanations without historians to think of them.) For this reason, I make no attempt to deal with the entirety of a particular historical work; I focus only on what is relevant to interpret the particular explanation.

The aim of the following descriptions is simply to introduce the historical works to the reader. The way they are described in this chapter may, therefore, not be equivalent to the more considered and detailed treatment of those explanations in the remainder of the thesis. In this chapter, terms such as *law*, *cause*, *target*, *contrast*, *counterfactual*, *intention*, *important* will be used purely heuristically, and without further analysis. In later chapters the roles will be, at times, reversed: an appreciation of historical practice will be used to shape understanding of the above concepts.

With this in mind, it is natural to focus on the following three features of each explanation, which initially seem of particular relevance in understanding explanatory competition:

- What the historian attempts to explain: the explanandum
- What does the explaining: the explanans
- The theory (explicit to varying degrees) at work behind the given explanation; what *kind* of explanation it is, and (if possible) why an explanation of that kind is given

The terms *explanans* and *explanandum* were first introduced by Hempel. Whilst it could be argued that a strict application of these terms should be restricted to an inferential ('explanation as argument') theory of the sort provided by Hempel, I shall use the terms more generally to refer (respectively) to *those things* (whatever they are) *which do the explaining*, and *that thing* (whatever it is) *which is explained*¹.

¹ A complication arises from the fact that whilst the plural of *explanandum* is *explananda*, there is no satisfactory plural of *explanans*; I shall therefore use the same term for singular and plural.

Georges Lefebvre

Lefebvre was well known as a historian of the details; the author of a number of regional micro-studies, a patient collector of primary source material, a hard working student of the archives. Yet the most distinctive aspect of his explanation is its overarching simplicity. The core of Lefebvre's explanation of the French Revolution is the application of a Marxist law of history: that legal and political power follow economic power.

'These groups [the nobility and clergy] preserved the highest rank in the legal structure of the country, but in reality economic power, personal abilities and confidence in the future had passed largely to the bourgeoisie. Such a discrepancy never lasts forever. The Revolution of 1789 restored the harmony between fact and law.'²

Lefebvre's belief in the primacy of economic-social conditions is indicated not only by his application of the Marxist law, but in his choice of wording. The legal structure is a fiction compared to the 'reality' of economic power. The Marxist idea that intellectual changes are epiphenomena is readily assented to: '[The bourgeoisie] had developed a new ideology which the "philosophers" and "economists" of the time had simply put into definite form'³.

This explanation provides, for Lefebvre, the 'ultimate cause' of the French Revolution. Yet an explanation must do more: 'this deeper cause of the French Revolution does not explain all its distinctive features'⁴. To make sense of 'distinctive', Lefebvre uses contrasting events: political revolutions in England, and nineteenth century 'transformations' on the continent. The differences between the French Revolution and these contrasts provide the two main 'distinctive features': that the Revolution was violent, and that it took the form of a liberation of 'the people' (the Third Estate) *by* the people. The cause of these features is as simply and clearly put as the fundamental cause: it is 'the collapse of the central power', in the form of

² Georges Lefebvre, *The Coming of the French Revolution*, trans. R. R. Palmer (Princeton: Princeton University Press, 1947), 2

³ *Ibid.*, 2

⁴ *Ibid.*, 2

the King's convocation of the *Estates-General* (an ancient, supposedly representative, body divided according to the three orders: clergy, nobility, and commoners).

It is a minor mystery of Lefebvre's work that although the key features of the explanation are given succinctly in the first three pages, the remainder of the book appears to follow a very different course. The work adopts a loose narrative structure; tracing the four acts of revolution, each organised around a particular social class – the aristocracy, bourgeoisie, working class, and peasantry. There is clear continuity in Lefebvre's belief in the primacy of the socio-economic. Yet the introduction and main body can be seen as misaligned; Eisenstein complained that whilst the narrative encourages the belief that all classes had a role to play in bringing about revolution, only the bourgeois' actions are mentioned as part of the 'ultimate cause'⁵. Further, this ultimate cause is not subsequently defended; indeed, it is barely mentioned outside the introduction. There is a suspicion that Lefebvre's Marxist explanation hovers, unsupported, over the detail of the text.

The explanatory style of Lefebvre's narrative is quite standard. It consists of – the introduction notwithstanding – a careful examination of motives, punctuated by 'cross-sectional' analyses of the revolutionary 'actors': the socio-economic groups. Lefebvre's narrative, as with all narratives, can be seen as a detailed explanation in its own right. As such, it would be impossible to attempt to summarise the whole here; more productive is to focus on Lefebvre's narrative at one key point (as defined above) in the narrative: the bourgeois revolution. Lefebvre's thesis is that the bourgeois found their path (in particular, towards social privilege) blocked by the nobility. This prompted the abandonment of the idea of distinction by birth, and the promotion of freedom of science and commerce.

Concerns have been raised over Lefebvre's concept of 'bourgeoisie' as it features in this explanation. Lefebvre is pulled between the recognition of the complicated and diverse nature of what he called the eighteenth-century *bourgeois*, and the explanatory need to provide a social group unified in circumstance and attitude. It is a tension

⁵ Elizabeth Eisenstein, 'Who Intervened in 1788? A Commentary on "The Coming of the French Revolution"' in *American Historical Review* 71(1) (1965): 77-103

exploited by many subsequent historians critical of Lefebvre; Cobban (another of the eight historians considered here) included. Bearing in mind the explanation given, the essential features of the bourgeoisie can be supposed to be their aspiration to nobility, economic behaviour, and belief in enlightenment ideals such as freedom of science and commerce. Yet the applicability of precisely these features has been strongly criticised, both by pointing out the differences within the 'bourgeois' in these respects, and by arguing that these respects are insufficient to distinguish between elite 'bourgeois' and elite nobility. Further, much of the criticism can be (and was) made using Lefebvre's own data and observations!

Lefebvre's work cemented an orthodoxy concerning the origins of the French Revolution which lasted perhaps twenty years (until the 1960s) – a long time in such an active field of historical scholarship. By emphasising four different aspects of revolutionary origins, Lefebvre made his writing amenable to a wide range of historians. In the preface to *The Coming of the French Revolution*, R.R. Palmer claims that '[Lefebvre's] writings have been as little subject to controversy as any on the French Revolution can be, and have been generally praised by all schools of thought on the subject'⁶. Lefebvre's position as a distinguished Chair of French Revolution Studies, and as editor of *Annales historiques de la Revolution Francais* for 30 years gave him a personal influence over wider scholarship. Yet despite complaints such as that 'it is difficult to avoid the conclusion that the orthodox theory of the Revolution has now assumed some of the characteristics of a religious belief'⁷, Lefebvre cannot be accused of instituting such a closed paradigm. Even his opponents (including Cobban) praise him for including in his historical studies an extremely broad range of facts which might be considered relevant, whether or not these facts confirmed his overall theories.

⁶ Lefebvre, vi

⁷ Alfred Cobban, *The Social Interpretation of the French Revolution* (Cambridge: Cambridge University Press, 1964), 10-11

Alfred Cobban

Cobban was one of three prominent historians in the 1950s-60s to first reject the orthodox Marxist approach best exemplified by Lefebvre's writing (the others being Richard Cobb and George Taylor). Their criticism was to come to convince the vast majority of French Revolution historians. However, the reaction to Cobban's own explanation has been less favourable. It is an explanation as thoroughly based on social categories as is Lefebvre's: whilst the Revolution is 'essentially' a political event, 'behind the political regime there is always the social structure, which is in a sense more fundamental and is certainly much more difficult to change.'⁸. Indeed, Lefebvre's social categories are used by Cobban, though Cobban is critical of Lefebvre's understanding of these categories, and is happy to alter the meaning of the terms.

Cobban's work is 'an interpretation'. By this he means that he wishes to answer a 'series of specific historical problems, and ask such questions as: what are the facts of the so-called bourgeois revolution, and in particular who were the bourgeois? What was the feudalism which they are supposed to have overthrown? How was the bourgeois revolution related to the revolt of the peasantry? This raises the problem of the relationship of town and country. Again, what part was played in the revolutionary situation by the lower social elements? This broadly raises the problem of the relation of rich and poor.'⁹

Cobban sees his task as primarily one of a critical *description*. In particular: the description of the nature of social divisions (especially an examination of 'bourgeoisie'), description of the rural economic system ('feudalism'), description of motivations in inciting revolution (for – or against – capitalism), description of the economic effects of the Revolution. In short, Cobban aims to say *what* the Revolution was, not primarily *why* it came about (though, in chapter six, I shall argue that such a distinction is not as significant as it may at first appear to be).

⁸ *Ibid.*, 162

⁹ *Ibid.*, 24

Cobban's theoretical aim, re-iterated at a number of points, is to provide an empirically based interpretation which escapes 'the rigid patterns of system-makers who have deduced their history from their theories'¹⁰. His dislike of Lefebvre's (and others') Marxist approach derives from their reliance on a theory which Cobban perceives as self-confirming, reductivist, and distorting of the historical facts. The charge of self-confirmation is most interesting, and is applicable both to the *facts themselves*, and to the *importance given to those facts*:

'The sociological historian uses his theory as the criterion for the selection of the relevant historical facts, and then on the basis of those selected facts he illustrates and confirms the theory by which they have been selected. Part of the fascination of general sociological theories is that success is built-in.'¹¹

'There is also another way in which sociological laws are self-confirming: by taking one factor in history as basic, all the others can be reduced to conditions in which it operates. Thus if different responses occur in economically identical circumstances, the primacy of the economic factor is maintained by making the factors which produce the aberration into conditions, on which the economic factor acts as a prime cause.'¹²

Clearly, in his own explanation Cobban cannot do without relying on the notions of *cause*, *significance*, and *types*, which he recognises in his chapter one. However, Cobban's ideal theory would involve a reliance on concepts which are multi-dimensional. He would seek to deny the existence of wider historical laws; and would search for a more pluralistic understanding of the strands making up the origins of the French Revolution.

Cobban's explanation reverses the Marxist, orthodox position: 'the revolution was to an important extent *against* and not *for* the rising forces of capitalism.'¹³ Likewise, 'the revolution was a triumph for the conservative, propertied, land-owning classes,

¹⁰ *Ibid.*, 168

¹¹ *Ibid.*, 13

¹² *Ibid.*, 13

¹³ *Ibid.*, 168

large and small.’¹⁴ The Revolution can only be seen as a bourgeois revolution if the bourgeois in question are conservatives: landowners, *rentiers* (those who made a living from investments such as crown bonds), and officials. However, a Marxist idea of bourgeoisie implies that they are the capitalists, financiers and entrepreneurs. Cobban also introduces a conflict between town (industry) and country (land ownership); a conflict which Lefebvre has noticed, but which Cobban thinks ‘more fundamental than he [Lefebvre] allows ...’¹⁵.

Although Cobban makes it clear that he does not want to reduce his interpretation of the coming of the French Revolution to a single *law*, it is capable of summary. Cobban’s fundamental idea is capable of being couched in social-economic Marxist terminology, yet reverses Marxist belief. The Revolution was fundamentally the work of those interested in attempting to maintain the status quo in a changing political and economic world: the work of land owners, conservatives, and those with a vested interest in preventing capitalism.

Jules Michelet

Michelet and Hegel (the following case study) are rather further from the historical mainstream than Lefebvre and Cobban. Michelet writes in an unashamedly personal style, and for unashamedly contemporary reasons: primarily, that he wanted to keep the glorious memory of the Revolution (that between 1789 and 1792) alive, in order to inspire new revolutionaries (such as those of the abortive 1848 revolution). He is perhaps better regarded as a literary figure than he is as an historian, his book being seen by the late nineteenth century as the work of ‘the poet historian’¹⁶. Yet Michelet, for all his stylistic peculiarities, has had a genuine influence on French Revolutionary

¹⁴ *Ibid.*, 170

¹⁵ *Ibid.*, 167

¹⁶ John Hooper, ‘Changing Perceptions of Jules Michelet as Historian: History Between Literature and Science, 1831-1874’ in *Journal of European Studies* 23(3) (1993), 295

historiography since his book's publication in 1853. The origins of the French Revolution are considered at the start of a dramatic and emotive narrative of the French Revolution, itself a part of a wider history of France.

Whilst Michelet's narrative is clearly intended to provide an understanding of the origins of the Revolution, Michelet's focus at the start of his work is on a question which can appear to the modern reader somewhat tangential:

'Is the Law, such as it appeared in the Revolution, conformable, or contrary to the religious law which preceded it? In other words, is the Revolution Christian or Anti-Christian? This question, historically, logically, precedes every other.'¹⁷

The dialectic between Christian ideas and Revolutionary ideas is important to Michelet not only because he wonders whether they might be (or might have been) reconciled. It is also important because it shapes his whole understanding of what the Revolution was and why it came about. During the reign of Louis XIV the people of France lived under the idea of a paternal monarchy: that the King was also the father who would look after his nation. There was no justice, as there was no law: but there was love. This reign of love was the Christian era; since the fundamental Christian principle is the victory of love, or grace, over the law.

For two reasons, this 'rule of grace' became unstable. Firstly, the government was inefficient – due primarily to evil *intendants* (regional governors); here, as elsewhere, Michelet has no qualms about making sweeping moral judgements. This made the Revolution inevitable as early as 1709. Secondly, the 'rule of grace' was only as good as the particular father-king; and, in the case of Louis XV this meant that it was not good at all. The reaction to these problems was firstly, and fundamentally, intellectual: consisting of the work of Montesquieu, Voltaire and Rousseau. 'When these two men [Voltaire and Rousseau] have passed, the Revolution is accomplished in the intellectual world.'¹⁸

¹⁷ Jules Michelet, *History of the French Revolution*, trans. Charles Cocks, ed. Gordon Wright (Chicago: University of Chicago Press, 1967; originally 1853), 17

¹⁸ *Ibid.*, 63

Hence Michelet's insistence that the essence of the Revolution is the rise of reason and justice, explicitly against the Christian principles of grace and love: 'The Revolution is nothing but the tardy reaction of justice against the government of favour and the religion of grace.'¹⁹ Given Michelet's narrative, the answer to his fundamental question of whether the Revolution is Christian or anti-Christian is not surprising: despite their similarity, fundamentally the two are opposed:

'The Revolution continues Christianity, and it contradicts it. It is, at the same time, its heir and its adversary. In sentiment, and in all that is general and human between them, the two principles agree, but in all that constitutes very and special life ... they are adverse and thwart each other.'²⁰

With regard to Michelet's explanation, there is one further striking feature – his emphasis on *unity*. As Hayden White explains²¹, the guiding force of Michelet's narrative is the struggle for unity – a typically Romantic desire. In describing the elections to the Estates General in early 1789, Michelet says that 'when, for the first time, in the course of ages, these words were heard: *All shall assemble to elect, all shall send in their complaints*, there was an immense, profound commotion, like an earthquake'²². Later, he remarks of the elections that 'A movement so vast, so varied, so wholly unprepared, and yet so unanimous, is most wonderful! All took part in it, and (except an insignificant number) they all desired the same thing.'²³

Of course, this was not even remotely true in any strict sense – the electors numbered five million in a population of around 28 million. But in emphasising unity, in believing in it as an ideal, and even in claiming its teleological inevitability, Michelet repeated the revolutionary doctrine itself.

¹⁹ *Ibid.*, 27

²⁰ *Ibid.*, 22

²¹ Hayden White, *Metahistory: the Historical Imagination in Nineteenth-Century Europe* (Baltimore, London: Johns Hopkins University Press, 1973), 150-152

²² Michelet, 83

²³ *Ibid.*, 87

Georg Hegel

Although it might seem odd to regard Hegel as a ‘proper’ historian, there really is no more reason to exclude Hegel than there is, say, Michelet (as Hayden White has argued in *Metahistory*²⁴). In particular, both wrote intellectual history, both used abstract terminology, and some degree of teleological support is provided in both accounts. One difference, however, is Hegel’s formidable theoretical backing for his history, of a different magnitude to that of the other historians considered here. I will give a brief account of the relevant parts of Hegel’s theory, before reconstructing the explanation he provides for the French Revolution. It is worth noting that my purpose is to use Hegel’s philosophy to understand better his history; whereas it is more usual in the philosophical literature to use Hegel’s history to better understand his philosophy. I should also note that I follow here Hegel’s treatment in *The Philosophy of History*, rather than the slightly different account presented in the *Phenomenology of Spirit*.

Hegel’s ‘Philosophical history’ is based on the *presupposition* that ‘Reason rules the world, and that world history has therefore been rational in its course.’²⁵ Yet, Hegel does not want to proceed using a priori ideas, but empirically: he states that the idea of history’s rationality is not only presupposition, but also the outcome of the study of history²⁶. Aren’t these two positions – the idea that we should adopt the presupposition that history is rational and the desire to proceed empirically – inconsistent? A possible solution is to be found in Hegel’s quite modern recognition that no historian can examine the past without some ‘categories’; that there can be no ‘purely empirical’ history. Since categories must be involved, whether consciously adopted or not, Hegel self-consciously uses the category of reason through which to

²⁴ White, *Metahistory*, 427-428

²⁵ Hegel, *Introduction to the Philosophy of History*, 12

²⁶ *Ibid.*, 13

examine history: ‘To him who looks at the world rationally, the world looks rational in return.’²⁷

This interpretation is supported by an interesting passage from Hegel’s *The Philosophy of Right*:

‘An event, or a situation which has arisen, is a *concrete* external actuality which accordingly has an indeterminable number of attendant circumstances. Every individual moment which is shown to have been a *condition, ground, or cause* of some such circumstance and has thereby contributed *its share* to it may be regarded as being *wholly*, or at least *partly, responsible* for it. In the case of a complex event (such as the French Revolution), the formal understanding can therefore choose which of a countless number of circumstances it wishes to make responsible for the event.’²⁸

The position Hegel describes should be seen as similar to Mill’s view of causation: that there are limitless conditions for any actual event, between which some selection must be made. Hegel in fact goes on to criticise the above conception of selection as arbitrary choice; *some* principle of selection must be adopted, and Hegel believes his to be well founded. The first element of Hegelian theory which requires emphasis is, therefore, the assertion that a *philosophically principled standard is required* in order to select between conditions.

Over and above this principle, Hegel emphasises the importance of *rationality* in historical explanation. Roughly, an explanation should consist in demonstrating that the event rationally had to happen. More precisely, a historical explanation of a *particular* event *x* (where *x* is an individual or social action) properly consists of showing that a *type* of event *X* (of which *x* is one example) rationally had to happen. This view has two important consequences. First, explanation, or at least the best explanations, are for Hegel conceptual, not causal. Causal explanations are too contingent: only conceptual explanations provide the degree of ‘satisfaction’ required. Second, the target of a Hegelian explanation will be a broad *type* of event, not, for example, the French Revolution in all its particularity. The type of event is, for Hegel, almost always a certain sort of change in conceptions of *freedom*.

²⁷ *Ibid.*, 14

²⁸ Georg Wilhelm Friedrich Hegel, *Elements of the Philosophy of Right*, trans. H. B. Nisbet, ed. Allen Wood (Cambridge: Cambridge University Press, 1991; originally 1820), Section 115

Hegel believes that the true subject of world-history is the development of the spirit of the age: the Hegelian ‘Spirit’. This means the development of *freedom*; since Hegel claims that

‘Just as the essence of matter is gravity, so the essence of Spirit is its freedom. Everyone will immediately agree that Spirit is endowed with freedom, among other characteristics. Philosophy, however, teaches us that all the characteristics of Spirit subsist only by means of freedom ...’²⁹

Hegel’s history therefore essentially consists of the changing ideas of freedom, of which the idea of freedom underlying the French Revolution comes near the end of the story. History’s true movement and progress can be gauged on these developments alone. World events are explained as resulting from new ideas of freedom, and the contradictions in those ideas explain (dialectically) the practical working-out of those events:

‘It has been said, that the *French Revolution* resulted from Philosophy, and it is not without reason that Philosophy has been called “Weltweisheit” [World Wisdom] for it is not only Truth in and for itself ... but also Truth in its living form as exhibited in the affairs of the world.’³⁰

Hegel’s explanation of the origins of the French Revolution and his explanation of its demise thus stem from the same source: the idea of freedom which motivated the Revolution. This ability to explain both rise and fall should be contrasted with, for example, Michelet’s account; which struggles to explain the failure of revolution. The idea of freedom which lies behind the French Revolution originates in Enlightenment philosophy – in particular, in the work of Kant and of Rousseau. It is an idea which is purely universal; meaning a-historical and unrestricted:

‘According to it [Kantian philosophy] the simple unity of Self-consciousness, the ego, constitutes the absolutely independent Freedom ... Rationality of Will is none other than the

²⁹ Hegel, *Introduction to the Philosophy of History*, 20

³⁰ Georg Wilhelm Friedrich Hegel, *The Philosophy of History*, trans. J. Sibree (New York: Dover Publications, 1956; originally 1840), 446

maintaining of one's self in pure Freedom – willing this and this alone – Right purely for the sake of Right, Duty purely for the sake of Duty.’³¹

Immediately following, Hegel raises a contrastive question: ‘Among the Germans this view assumed no other form than that of tranquil theory; but the French wished to give it practical effect. ... why did the French alone, and not the Germans, set about realizing it?’ (Hegel considers the response that ‘the French are hot-headed’, fortunately regarding this solution as ‘superficial’.) The answer is that Germany had been through the Reformation, whereas France had not:

‘In Germany the entire compass of secular relations had already undergone a change for the better; those pernicious ecclesiastical institutes of celibacy, voluntary pauperism, and laziness, had already been done away with; there was no dead weight of enormous wealth attached to the Church, and no constraint put upon Morality – a constraint which is the source and occasion of vices; there was not that unspeakably hurtful form of iniquity which arises from the interference of spiritual power with secular law, nor that other of the Divine Right of Kings ... The principle of Thought, therefore, had been so far conciliated already’³²

Thus, in France, ‘The new Spirit began to agitate men's minds: oppression drove men to investigation.’³³ Hegel is also concerned to ask why the Revolution was initiated by ‘the people’, why the revolution came ‘from below’³⁴. The reasons that the government did not undertake the change demanded were that (i) the Court, Clergy, Nobility and Parliaments (surely Hegel meant, *parlements*: the French law courts staffed primarily by nobility) were unwilling to surrender their privileges; (ii) that the government could not reconstruct the State on the basis of individual wills; and (iii) that as a Catholic state, adherence to religious ideals in France had to take priority. Social revolution was therefore inevitable.

³¹ *Ibid.*, 443

³² *Ibid.*, 444-445

³³ *Ibid.*, 446

³⁴ *Ibid.*, 446-447

François Furet

It would not be too great an exaggeration to claim that Furet is the most influential contemporary historian of the French Revolution. His interpretations of the *Jacobin* regime (1793-late 1790s) and the role of the Terror are at the centre of his writing, yet he also inquires into the origins of the Revolution. In *Interpreting the French Revolution* Furet uses a reading of Alex de Tocqueville and Augustin Cochin to shape his own interpretation. Whilst Tocqueville emphasised the (primarily) economic continuity between pre-revolutionary and post-revolutionary France, Cochin emphasised the political and cultural discontinuity.

Furet's explananda are given according to a series of questions. Like Tocqueville, Furet sees a great deal of continuity before and after the Revolution; yet, given this fact, he asks why a revolution was 'required' at all:

'... between the origins and the end result ... there is a blank page that Tocqueville never filled. ... why did the process of continuity between the old regime and the new involve a revolution?'³⁵

Turning the Marxist interpretation upside-down, Furet believes the Revolution to have had a negligible socio-economic effect; 'neither capitalism nor the bourgeoisie needed revolutions to appear in and dominate the history of the major European nations in the nineteenth century'. Why was France then 'the country that, through the Revolution, invented democratic culture, and revealed to the world one of the basic forms of historical consciousness of action'³⁶?

A few pages later, Furet questions why the Revolution 'erupted with such torrential violence'³⁷, and why the French adopted cultural assumptions 'which precluded for example ... the English idea of representation'³⁸. Furet understands the Revolution as

³⁵ François Furet, *Interpreting the French Revolution*, trans. Elborg Forster (Cambridge: Cambridge University Press, 1981; originally 1978), 23

³⁶ Both quotes: *Ibid.*, 24

³⁷ *Ibid.*, 28

³⁸ *Ibid.*, 36

independent of socio-economic forces; it was an abnormal, even unnatural, event. We need an explanation as to why only *France* went through this profound, violent change.

It is clear that Furet's belief in the inadequacy of the Marxist theory (in particular, its idea of the French Revolution as a normal, predictable phase in human history) affects his theoretical approach in general. His criticism of the notion that the Revolution was inevitable leads to a suspicion of causal analysis in general:

'A phenomenon like the French Revolution cannot be reduced to a simple cause-and-effect schema. The mere fact that the Revolution had causes does not mean that they are all there is to its history.'³⁹

This interesting theoretical claim has a direct bearing on my interpretation of the case study explanations, and for this reason I shall comment in detail on Furet's theoretical convictions in chapter six. An important component of Furet's views on this topic is that pre-Revolutionary causes are insufficient to explain the development of the Revolution; only the earlier history of the Revolution itself suffices. This is suggested by such comments as that

'One could, for instance, explain the revolt of most of the deputies to the Estates General by the political crisis in the Ancien Regime; but the situation created in its wake by the vacancy of power and the ensuing insurrection added a totally unprecedented dimension to that crisis, with consequences that no one could have foreseen two months earlier.'⁴⁰

Furet does, in fact, provide a conventional causal explanation, though it emphasises the unpredictability of the event:

'the 'revolution' was born of the convergence of very different series of events, since an economic crisis, a complex phenomenon in itself, involving agricultural, 'industrial', meteorological and social factors) took its place alongside the political crisis that had begun in 1787. This convergence of several heterogeneous series, surely a fortuitous situation, was

³⁹ *Ibid.*, 22

⁴⁰ *Ibid.*, 22

to be transformed as early as the spring of 1789 by a retrospective illusion in which it was seen as the inevitable consequence of bad government.’⁴¹

Yet these causal claims are only one element of Furet’s explanation. Much of Furet’s thinking on the origins of the French Revolution is shaped by his understanding of the Revolution as mediating between the two concepts of ‘state’ and ‘society’. Thus, ‘In the dialogue between societies and their States that is part of the underlying texture of history, the Revolution tipped the scales against the State and in favour of society.’⁴²

The answer to the question posed above concerning France’s inability to adopt the English model, or some other peaceful resolution, is in terms of state and society. In brief: Louis XIV (King of France from 1643-1715) had closed pre-existing avenues which allowed society to be represented to state. Those bodies which did subsequently claim powers of representation – particularly the *parlements* – were too conservative to win the support of educated society in general. So the French turned to the *philosophes* and men of letters, with disastrous consequences:

‘The men of letters tended to substitute abstract right for the consideration of facts, principles for the weighting of means, values and goals for power and action. Thus the French, deprived as they were of true liberties, strove for abstract liberty; incapable of collective experience, lacking the means of testing the limits of action, they unwittingly moved toward the *illusion of politics*.’⁴³

Furet sees the Revolution as a political and ideological event, whose socio-economic aspects are negligible. It is to be interpreted primarily as the coming to political power of society, and the replacement of (what was in theory) the absolutist monarchical state. A fortuitous conjunction of economic and political crises brought the Revolution into being; but its development could only be explained by the peculiarly French role of intellectuals in pre-revolutionary political debate, and by the unfolding of the Revolution itself.

⁴¹ *Ibid.*, 24-25

⁴² *Ibid.*, 24

⁴³ *Ibid.*, 37

Simon Schama

Simon Schama is one of the most prominent contemporary historians specialising in any period. *Citizens*, his 900-odd page chronicle of the French Revolution, is eminently readable, informal, and packed with idiosyncratic and surprising events and anecdotes.⁴⁴ Schama deliberately fashions his explanation in the form of a narrative, rather than in the more analytical, abstractive style which was more prominent when *Citizens* was published. This is because Schama believes that only the narrative form can make sense of the fact that ‘the Revolution was a much more haphazard and chaotic event and much more the product of human agency than structural conditioning’⁴⁵.

The precise target for explanation is not given, but the aim can be understood as the traditional narrativist one of bringing *meaning* to the events of the narrative, without compromising the belief in the complexity and haphazardness of the sequence. The belief that the Revolution resulted primarily from a succession of accidents is at the core of Schama’s treatment. Schama disputes the idea that because the event was of ‘epochal significance, then the causes that had generated it had necessarily to be of an equivalent magnitude’⁴⁶. So how is the chaotic made meaningful? There are, to some degree, general themes which organise Schama’s narrative, as we shall see. In addition, Schama builds his compelling narrative around various lively, distinctive *characters*.

Thematically, Schama takes his approach to the French Revolution to have furthered certain ideas which had already been suggested. Three in particular: in a Tocquevillean spirit, continuities (between pre-revolutionary and post-revolutionary France) should be emphasised as much as discontinuities, particularly in socio-economic terms. Secondly, ‘The Revolution’ is better viewed as multiple revolutions,

⁴⁴ Or, in Murray’s words: ‘Schama’s *Citizens* is as brilliant in presentation as it is suspect in scholarship’: Bill Murray, ‘Review Article’ in *Australian Journal of French Studies* 26(1) (1989), 100

⁴⁵ Simon Schama, *Citizens: a Chronicle of the French Revolution* (London: Viking, 1989), xv

⁴⁶ *Ibid.*, xiv

dictated more by local passions and interests than by the interests of whoever ruled in Paris. Finally, individual agency, and in particular revolutionary utterance, was more important than social ‘structure’.⁴⁷

There are three more positive themes which Schama explicitly states will motivate and organise his account. He wishes to emphasise the relationship between the concepts of *patriotism* and *liberty*; the analogy between *citizenship* and the *family* (which was also explicitly followed by the revolutionaries); and, most especially, the violence of the Revolution.

In a narrative of this length – it takes over 400 pages to reach the fall of the Bastille – it is hard to select particular sections without that selection being, to some degree, arbitrary. Yet the following passages demonstrate Schama’s privileging of the theme of *patriotism*, which (as promised) re-occurs throughout:

‘The monarchy collapsed when the price of its financial rescue was measured not in profits or offices but in political concessions. ... henceforth, an alternative conviction was in the ascendant: that patriotic freedom would produce money where reforming absolutism had not.’⁴⁸

‘The bringing together of political patriotism with social unrest – anger with hunger – was (to borrow the revolutionaries’ favourite electrical metaphor) like the meeting of two live wires.’⁴⁹

In his theoretical approach, Schama’s is quite typical of modern historiography. The explanation is broadly causal, yet resists summary into one or two major causes, as Lefebvre (for example) was keen to do. In addition, causal explanation does not exclude interpretive explanation (the explanation of social action in terms of rules which motivate those actions); as can be seen in the exploration of the citizenship/family notion.

⁴⁷ *Ibid.*, xiv

⁴⁸ *Ibid.*, 288

⁴⁹ *Ibid.*, 294

Schama downplays any 'structural' causes, replacing them with an array of micro-causes. He also tends to emphasise more unusual and overlooked causes, such as the culture of patriotism in the old regime. Schama's judgements of significance, implicit in these decisions, is the element of his work which has attracted most criticism. At least two reviewers of *Citizens* (Murray and Spitzer⁵⁰) attack Schama's approach as 'unbalanced': in his excessive emphasis first on the stable and irreproachable aspects of the old regime, and then on the blood lust of the lower classes which led to the Terror.

Despite Kellner's claims⁵¹ that Schama should be seen as a post-modernist historian because he eschews meta-narrative, Schama in fact has certain very clear general themes which are implied in his selection and emphasis. The Revolution was a glitch en route to a modern democratic French state. The old (pre-revolutionary) regime was adequate to the tasks facing it, whilst the Revolution became prone to an evil excess. Though there are themes identifiable in Schama's work, the overall effect of his narrative is to instil a belief in the *multiplicity* of the Revolution: that the causes of the Revolution were as numerous as they were fortuitous, that there was no *one* Revolution, and that we should be wary of such broad types as the 'bourgeois'.

Theda Skocpol

Skocpol is a macro-sociologist, rather than a historian. As such, it is no surprise that the explanation she offers is more scientific than the other seven considered here. Her overall aim is to explain social revolutions in general; but she is also careful to pay detailed attention to the history of each particular revolution she considers (French, Russian and Chinese). Her understanding of her task reflects this awareness of universal and particular:

⁵⁰ Alan B. Spitzer, 'Narrative Problems: the Case of Simon Schama' in *Journal of Modern History* 65(1) (1993): 176-192

⁵¹ Hans Kellner, 'Beautifying the Nightmare: The Aesthetics of Postmodern History' in *Strategies: A Journal of Theory, Culture and Politics* 4 (1991): 289-313

‘Each such revolution, furthermore, has occurred in a particular way in a unique set of social-structural and international circumstances. How, then can a sociologist hope to develop historically valid explanations of social revolution as such?’⁵²

To answer, Skocpol turns to comparative analysis. Her method is to first select cases whose histories are as similar as possible: France, Russia and China primarily, with Prussia, Japan and England as similar cases but where there was no revolution. These cases are then compared according to Mill’s Methods of Agreement and Difference, in order to discern causal influence. The Method of Agreement states that where two cases share the same outcome, the cause is a feature which they share in common. The Method of Difference states that where two cases have different outcomes, the cause is a feature in which they differ. This ‘comparative historical analysis’ is, says Skocpol ‘the mode of multivariate analysis to which one resorts when there are too many variables and not enough cases.’⁵³

Skocpol recognises some of the problems of this kind of approach, including the necessity of a prior understanding of which variables are *relevant*. The methods are only perfectly applicable where the cases stand in a very particular relationship; the Method of Agreement only yields a definite answer where the cases share only one feature, whilst the Method of Difference only yields a definite answer where the cases differ in only one feature. Given that actual comparative cases never reach these ideals, we need to screen candidates for causal factors by an initial judgement of relevance (the *problem of multiple differences* facing the Method of Difference will be relevant elsewhere in the thesis, notably chapter seven). Skocpol recognises that ‘strategic guesses have to be made about what causes are actually likely to be operative – that is, which ones could, or could not actually affect the object of study.’⁵⁴ Similarly, Skocpol recognises that historical theory is, to some extent, necessary before proceeding with comparative historical analysis; in order to provide concepts for use in that analysis.

⁵² Theda Skocpol, *States and Social Revolutions: a Comparative Analysis of France, Russia and China* (Cambridge: Cambridge University Press, 1979), 33

⁵³ *Ibid.*, 36

⁵⁴ *Ibid.*, 39

Indeed, Skocpol is, if anything, too wary of the comparative method. Two examples: firstly, she worries that ‘comparative historical analysis necessarily assumes (like any multivariate logic) that the units being compared are independent of one another’⁵⁵, whereas this is not true of her examples. I do not see the justification for this restriction; in particular, so-called ‘incompatible contrastive explanation’ (see chapter three) makes perfect sense, yet the two contrasts are dependent upon each other in the sense that if one contrast is actualised, the other cannot be. Secondly, she states that comparative historical analysis ‘cannot provide the causal hypotheses to be explored’⁵⁶. Yet, given appropriate terminology and an understanding of what is relevant, this is exactly what comparative methodology can do: for it asks us to look for differences (or similarities) between the cases, and to treat these as causal hypotheses.

Skocpol’s approach to explaining social revolutions is ‘statist’:

‘We shall analyse the causes and processes of social revolutions from a nonvoluntarist, structural perspective, attending to international and world-historical, as well as intranational, structures and processes. And an important theoretical concomitant will be to move states – understood as potentially autonomous organizations located at the interface of class structures and international situations – to the very centre of attention.’⁵⁷

Specifically, she wants to establish the statist view as superior to the Marxist ‘social class’ understanding of revolutions. This conviction she bases on the claim that the statist model picks out the more salient causal factors:

‘Perhaps especially because the factors that they consider are indeed an important part of the story, Marxists have failed to notice a crucial point: Causal variables referring to the strength and structure of old-regime states and the relations of state organisations to class structures may discriminate between cases of successful revolution and cases of failure or non-

⁵⁵ *Ibid.*, 39

⁵⁶ *Ibid.*, 39

⁵⁷ *Ibid.*, 33

occurrence far better than do variables referring to class relations and patterns of economic development alone.’⁵⁸

Turning specifically to the French Revolution, the key fact requiring explanation for Skocpol is why France ‘devolved – despite a half century of vigorous economic expansion – from near-dominance in Europe to the humiliations of martial defeats and royal bankruptcy’⁵⁹. This is a key fact, since it ‘renders comprehensible the specific political crisis that launched the French Revolution’.

The reasons for this military and financial breakdown are three-fold. Firstly, royal absolutism was always more an ambition than a realised goal; institutions such as the *parlements*, and the sale of state offices (venality) left the crown with the political obstacle of a socially dominant class. This was to prove a major hindrance, in particular in attempting to alter tax collection so that the nobility were no longer exempt.

Secondly, French agriculture remained ‘backward’: both in the sense of consisting of highly divided small-holdings, and in maintaining the seigniorial rights of large landowners. These problems were masked from 1730-1770 by a combination of good weather and internal order, but were liable to lead to food crises in times of bad harvest. Furthermore, this agricultural weakness hampered the development of French industry, particularly as compared to the British.

Thirdly, the strategic position of France necessitated wars on land and at sea, wars which in the long run simply could not be adequately funded. These three causes combined to leave the French monarchy without sufficient funds to avert an ‘administrative and military breakdown’. From Skocpol’s comparison with Prussia, Japan and England, she aims to show that these three causes are each necessary, and jointly sufficient.

Skocpol’s explanation privileges structural causes over voluntary agency; indeed, as the explanation is given above, voluntary agency has no causal role whatsoever. Political and strategic factors take complete precedence over social and ideological.

⁵⁸ *Ibid.*, 34-35

⁵⁹ *Ibid.*, 52

Roger Chartier

The title of Chartier's *The Cultural Origins of the French Revolution* makes it clear that any explanation found here will be limited in scope to cultural aspects. The use of 'cultural' origins is consciously intended to replace the investigation of 'intellectual' origins found in, for example, the work of Michelet, Hegel, or Chartier's professed forerunner, Mornet. In making this change, Chartier wishes to place emphasis on the 'collective decisions' which comprise culture, rather than the 'clear thoughts' of individual conceptual innovators⁶⁰.

Chartier claims that the emphasis on cultural origins carries three benefits. Firstly, it should lead us to abandon the question of whether or not the *content* of the Revolution corresponded to the content of the ideas preceding it (those of the Enlightenment). Instead, we should focus on whether and which of the ideas made the Revolution *conceivable*.⁶¹ Secondly, it allows investigation of 'not only clear and well elaborated thoughts but also unmediated and embodied representations; not only voluntary and reasoned engagements but also automatic and obligatory loyalties'⁶². In other words, Chartier wishes to investigate those less conscious aspects of life, which can not necessarily be reduced to clear, underlying ideology. Thirdly, it reverses the relationship between practice and idea. Chartier wants to deny 'that practices can be deduced from the discourses that authorize or justify them'⁶³; rather, 'even the most powerful and most original conceptual innovations are inscribed in the collective decisions that regulate and command intellectual constructions before they achieve expression in clear thoughts.'⁶⁴ This understanding of 'culture' *as practice*, rather than in terms of clearly articulated ideology, shapes Chartier's explanation.

⁶⁰ Roger Chartier, *The Cultural Origins of the French Revolution*, trans. Lydia G. Cochrane (Durham, N.C., London: Duke University Press, 1991), 2

⁶¹ *Ibid.*, 2

⁶² *Ibid.*, 6

⁶³ *Ibid.*, 18

⁶⁴ *Ibid.*, 2

To make the novelty of the approach clear, compare it to Mornet's. His task was to chart the diffusion of Enlightenment ideas throughout public opinion; one way he proposed to do this was by analysing book sales. As Chartier points out, the assumptions in such a task are that ideas are fundamental, whilst public opinion and practice secondary; and that public opinion is a passive receptacle for those ideas.

Chartier's work is similar to Furet's, not least in theory and terminology. Like Furet, Chartier is suspicious of causal explanation:

'History has become more circumspect in the designation of causality, and historians have learned prudence and scepticism from the difficult task of subjecting the brutal emergence of the revolutionary event to rational categories, as well as from their inability to conceive of historical development as necessitated and commanded by one discernible thread.'⁶⁵

Chartier prefers the idea of 'effective history': an approach which 'deals with events in terms of their most unique characteristics, their most acute manifestations.'⁶⁶ Most theories of explanation would struggle to comprehend such a practice as explaining at all: unity, not diversity, is at the heart of most notions of explanation. The one exception would be Hayden White, who considers under the title of *formist* explanation the idea that one may explain by revealing an event in all its particularity and uniqueness⁶⁷. This idea will be considered in the context of White's general theory of explanation, in the following chapter.

Chartier's aim, then, is to present and discuss certain 'conditions of possibility' for the French Revolution. The central cultural change in the old regime is the rise in importance of – indeed, practically the creation of – public opinion. Cultural shifts of this kind brought about the Revolution by making such an event *conceivable*, hence possible. They did not *cause* the Revolution, in the sense of being sufficient to explain the *content* of the Revolutionary ideas.

⁶⁵ *Ibid.*, 1

⁶⁶ *Ibid.*, 5

⁶⁷ White, *Metahistory*, 13-15

As an example of this point (and also of Chartier's privileging of practice over ideas), consider Chartier's treatment of the reading habits of eighteenth century France. The way that the educated eighteenth century French person read fundamentally changed in the latter part of that century, in that it became more critical. The importance of the cultural shift in reading habits was not so much in *what* was read, but the critical way in which it was read. For this made conceivable a critical approach to established authority more generally.

In his conclusion, Chartier returns to his worries over the causal approach; since, despite his earlier theoretical points, he admits that the idea that 'the event and its origin belong to distinctly different, clearly separated, sets of facts connected by a causal relation' has been for him a 'working hypothesis'⁶⁸.

The real novelty of Chartier's explanation is not so much to claim that there *are* cultural origins, nor even to ascribe more importance to them than is commonly recognised, for he clearly does not attempt this. It is, rather, to challenge definitions – *culture* in particular – and to challenge the relationship between originating factors, particularly ideas and practice.

What can be learnt from the case study?

The primary aim of this chapter has been to provide the reader with a passing acquaintance with the historical works which will be referred to in the remainder of the thesis. Yet, it is of interest to note that even a study as rough as that above can produce interesting results for a descriptive philosophy. The advantage of considering actual historical examples, and in attempting to respect the complexity of those examples, is that we can start to appreciate that there are multiple elements of historical explanation which require philosophical examination. The explanations

⁶⁸ Chartier, 197

considered above are (usually) not simple affairs, as might be suggested by the usual stock set of hackneyed examples.

We would be wise to presume that explanatory competition won't be reducible to factual disagreement between historians. The explanations considered in this chapter suggest that an explanation contains multiple philosophically relevant aspects, any of which *may* have a role to play in accounting for explanatory competition. One way that I structure Parts II and III of this thesis is by examining each element in turn, and asking of each whether, and how, that aspect affects explanatory exclusion and competition.

Most basically, each explanation makes use of **facts**; which I take to include facts concerning which **causes** were operative. Clearly factual disagreement will play at least some role in explanatory consistency and exclusion (chapter four).

Historians do not only provide positive claims, but also **negative** or **critical** claims. Whilst sometimes a criticism of another's explanation may be clearly distinguished from one's explanation itself, this is not always the case. Negative or critical claims appear to be not just peripheral to the explanation proper, but essential in determining what precisely the content of that explanation is (also explored in chapter four).

Each explanation has a **target** – that which demands explanation. It is clear that explanations using roughly similar wording ('the French Revolution') can still point to different targets; we need to account for this. We should also take account of the role of presupposition in framing explanatory targets. Finally, we need an account of the relation between sameness or difference of explanatory target, and explanatory consistency or exclusion (chapter five).

Each explanation also has a methodological **theory**, whether or not that theory is explicit. In particular, each historian has ideas concerning what is to be regarded as a good explanation. There may also be explicit constraints over the application of such explanatory terms as 'cause', 'law', or 'structure'. We need to ask what the relation is between such theoretical aspects and the more basic elements of the explanation. We

also need to know to what extent the adoption of different theories leads to explanations which simply cannot be compared (chapter six).

A **key concept** is a recurrent feature of explanations. By this I mean the concept or concepts which organise understanding of the subject matter. This ‘organisation’ can take different forms: in privileging a particular concept, a historian may provide a historical law, show there to be a teleological process, or illustrate that disparate events should be grouped together (also chapter six).

Each explanation inevitably deploys the related tools of **selection** and **weighting** (the claims of one thing being more important than another). The crucial question to ask of these concepts is how they are related to explanatory competition: are judgements of selection and importance a result of personal preference or interest, or can explanations be ranked in a more substantive manner? (chapters seven, eight and nine).

Chapter Three: Explanation

In chapters one and two I have investigated the methodological issues relevant to this thesis, and have introduced the case studies to be used. The third preparatory task is to investigate the concept ‘explanation’. Whilst it will not be an explicit aim of this thesis to provide a thorough analysis of the concept, it is clear that an investigation of ‘competing explanations’ dovetails with an understanding of what an explanation is. In this chapter, I will summarise some of the more influential approaches to explanation, especially those that relate to my specific task of charting competing *historical* explanations. In particular, I will be interested in material relevant to the analysis of explanation of *historical particulars*, such as the French Revolution. Most philosophical analyses of explanation are conducted, explicitly or otherwise, with reference to scientific explanation; much of this work, though by no means all, is relevant to historical explanation.

The categorisation of approaches to explanation adopted below distinguishes four fundamental strands: *erotetic*, *cognitive*, *ontic* and *linguistic*. Most writers fit neatly into one of these groups, though some do combine elements from more than one strand. I certainly do not wish to imply that this latter position is not possible; for the typology is one which is intended to describe typical approaches, rather than to map logically exclusive options. Indeed, the position which I shall recommend is formulated so that the virtues of each strand (as I see them) can be retained. The fifth approach, *contrastive explanation*, is important to later issues, so will also be introduced in some detail. It is, however, an approach to explanation which can only be argued to be supplemental to the more fundamental analyses of explanation represented by the first four types.

The aims of this chapter are two-fold. Firstly, I will introduce approaches to explanation, knowledge of which will be required for the arguments of chapter four

onwards. Certain explanatory theories are referred to a number of times in subsequent chapters, and hence require introduction: in particular, those of Bas Van Fraassen, Carl Hempel, Jaegwon Kim, David Lewis, Hayden White, and Peter Lipton.

Secondly, I will begin to elucidate *my* approach to historical explanation through comparison and criticism of these positions. I present a preliminary sketch of a theory of explanation at the end of this chapter; these ideas will be both applied and elaborated in the investigation of explanatory competition from chapter four onwards.

One of the main desiderata in developing a theory of historical explanation will be to remain faithful to the overall vision of philosophical analysis defended in chapter one. This implies that the term ‘explanation’ be applied *non-reductively*, and *permissively*. We should not *presume* that the concept will be capable of analysis under a single set of necessary and sufficient conditions; and we should investigate different types of explanation on their own terms, at least initially. To this end, I shall identify ‘explanatory reductivists’: those who would attempt to limit the notion of scientific and historical explanation to a single type; for example, causal or nomological. Yet a pluralistic approach does raise its own concern: namely, if there are different types of explanation, what (if anything) is common to them all? This issue will re-occur in this chapter.

Erotetic explanation

Van Fraassen pioneered the approach of explanations as why-questions in his 1980 *The Scientific Image*¹. Explanations have traditionally been viewed as *products*: arguments, statements, or similar. Regarding explanations as, fundamentally, *processes* shifts attention to the actual spoken or written acts, performed in particular conversational contexts. According to the erotetic theory, the correct way to analyse ‘explanation’ is, therefore, in terms of the logic of questions: erotetic logic.

¹ Bas C. Van Fraassen, *The Scientific Image* (Oxford: Clarendon, 1980)

The content of a why-question, Q, is determined by three features: (a) the topic to be explained, (b) a contrast or contrasts which constrain the answer (the nature of which will be examined later in this chapter²), and (c) a relevance relation which also, in a different way, constrains answers. The relevance relation delineates a class of answers which are of the right type, according to the particular context of the question. Whilst it may appear that the key notion *is* this relevance relation, Van Fraassen attempts to minimise its role by refusing to provide any substantive content to the notion: ‘Is this not where the inextricably modal or counterfactual element comes in? But not at all; in my opinion, the word ‘because’ here signifies only that A [the answer] is relevant, in this context, to this question’³.

This emphasis on the pragmatics of explanation is the distinguishing feature of Van Fraassen’s erotetic theory. The conversational context – the implicit desires and motives of the questioner in what precisely they want answered, and in what way – determines which precise question is asked. In particular, interpretation of the explanatory relevance relation requires knowledge of the particular pragmatic context. Knowledge of the explanatory topic and the contrast(s) seems not to be sufficient to produce knowledge of what is relevant. This is so because ‘relevance’ also depends upon what kind of question the questioner had in mind. Van Fraassen’s example is of the question ‘why did you get up at 7 o’clock this morning?’. To this, correct answers could be either ‘no reason, really’, or could be ‘because of the milkman’s clatter’; dependent upon whether the relevance relation restricted answers solely to motives (so no reason can be given), or whether it allowed causal factors in general.

Van Fraassen’s introduction of the pragmatics of explanation has been extremely influential; so much so, that even those philosophers who would attempt to provide a more substantive analysis to ‘explanation’ than Van Fraassen (Wesley Salmon, for example) see pragmatic aspects of explanation at least as one important feature amongst others. No one can now ignore pragmatic and contextual features in a theory

² The introduction of the notion of *contrast* remains one of the most influential aspects of Van Fraassen’s theory of explanation (though he cannot lay sole claim to its introduction: see later in this chapter). However, since I will argue that contrastive explanation is formally unrelated to erotetic explanation, I do not consider contrast at this point.

³ Van Fraassen, 143

of explanation⁴. Once a space is opened for contextual factors to influence what is explained, and the adequacy of that explanation, then it is conceivable to suppose that empirical study might illuminate features of explanation. In this vein, the psychologist Denis Hilton has described the typical ways for types of relevance relation to vary according to different types of question⁵.

Answers to why-questions are to be judged according to four criteria: (i) they should be true; (ii) they should fit with the truth of the explanatory topic and the falsity of the contrast(s); and (iii) they should satisfy the relevance relation. Further, Van Fraassen considers using (iv) probabilistic relevance as a guide to explanatory relevance; such that answers which probabilistically favour the topic over the contrast class are to be preferred. Criteria (i) and (ii) are insufficient to properly capture the adequacy of answers to contrastive why-questions, as I shall argue in the section on contrastive explanation. Criterion (iv) will also be examined in relation to contrastive explanation later in this chapter; and, for somewhat different purposes, in chapter eight. Here, though, I will be less concerned with providing an erotetic theory of answers than I will be in the general conception of ‘explanations as why-questions’.

Whilst the erotetic theory of explanation has brought certain benefits to the philosophy of explanation in general, in introducing the ideas of *context* and

⁴ Of course, these comments do rely on a certain notion of ‘pragmatic’: specifically, that pragmatic features of explaining are those which derive from implicit, shared and context-dependent understanding of the questioner and responder. Achinstein, for one, has argued that this understanding of ‘pragmatic’ is insufficient, since non-pragmatic theories of explanation may also incorporate the idea that knowledge of context is required to fully specify the explanatory target (Peter Achinstein, ‘The Pragmatic Character of Explanation’ (originally 1984) in *Explanation*, ed. David-Hillel Ruben (Oxford: Oxford University Press, 1993): 326-344). My notion of ‘pragmatic’ states only that the request for explanation is underdetermined by explicit demands, hence implicit and contextual clues must be used. However, once the request for explanation has been made definite and complete in this way, the resulting explanation may still be *assessed* entirely conventionally: i.e. non-pragmatically. Achinstein regards the notion that standards of assessment be pragmatically given as the hallmark of a truly pragmatic theory of explanation. I am not averse to this idea (see chapter six for a similar claim), but would still wish to regard the emphasis on context in the specification of explanatory target as an important step in developing a pragmatic theory of explanation.

⁵ In his most recent paper, for example, (Denis Hilton and John McClure, ‘Are Goals or Preconditions Better Explanations? It depends on the question’ in *European Journal of Social Psychology* 28 (1998): 897-911) Hilton finds that where questions are framed in terms of *why an action occurs*, or which request a causal explanation of the action, then goals are rated as better explanations than preconditions. However, if the question requests a causal explanation of an action *which has been obstructed*, then preconditions are seen as better explanations than goals.

pragmatics, contrast, and explanatory process, I would argue that as a complete theory it is insufficient. Two questions are particularly fundamental, and seem tricky to answer using the resources of an erotetic theory.

First: are there to be any restrictions on what counts as a proper ‘relevance relation’⁶? If not, then any proposition may, in the right context, be treated as a correct answer (and hence explanation) of any other. For instance⁷, relations of *astral influence* may be appropriately selected as an erotetic relevance relation; say, in order to explain JFK’s death by citing the positions of the planets. Such an answer meets Van Fraassen’s criteria for answers ((i) to (iv) above): the claims about the planets’ positions may be true, the claims fulfil the desired relation, and the answer, along with astrological background theory, favours the topic over the contrasts in the appropriate way. With no restriction on proper ‘relevance relations’, then the erotetic analysis finds good explanations where it shouldn’t.

But if ‘relevance relation’ is to be restricted, then it seems clear that this can only be done by moving beyond the resources of a purely erotetic theory of explanation. Indeed, Kitcher and Salmon claim that if Van Fraassen wishes to avoid an “anything goes” account of explanation, this ‘ought to commit him to solving most (if not all) of the traditional problems of the theory of explanation’⁸. I agree with their conclusion that whilst Van Fraassen has drawn attention to the pragmatics of explanation, his theory can, and must, be enriched by turning to other notions of explanation.

One answer which may be attractive to the erotetic theorist is that appropriate relevance relations should be restricted according to the maxim that scientific explanation should rely only upon relations allowed by the accepted science of the day: astral relations cannot feature in scientific explanations as they don’t feature in scientific theorising in general⁹.

⁶ This criticism is also at the core of Kitcher and Salmon’s critique of Van Fraassen’s theory (Philip Kitcher and Wesley Salmon, ‘Van Fraassen on Explanation’ (originally 1987) in Wesley Salmon, *Causality and Explanation* (Oxford: Oxford University Press, 1998): 178-189).

⁷ The example is from Kitcher and Salmon, 183-185

⁸ *Ibid.*, 188

⁹ This response is suggested by Van Fraassen, 141, 126

Yet this response is insufficient, and it is instructive to realise why. Whilst astral relation may not be a scientifically acceptable category, and hence is *never* an appropriate relevance relation, others may be scientifically respectable, yet still not always appropriate. To answer ‘why did that crystalline structure grow?’ by citing biological relations between the elements of the crystal would be incorrect, even though biological relations are scientifically acceptable. Biological relation is not an appropriate relevance relation in *this* case, and this cannot be explained using only erotetic criteria.

The second tricky question for the erotetic theorist is whether all answers to why-questions are really explanations; and, conversely, whether all explanations are answers to why-questions. Seemingly not; ‘why should one behave non-egoistically?’ is a why-question which is not a request for explanation (rather, it is a request for justification). And it seems obvious that explanations can be given in response to different kinds of questions (particularly, how-questions and what-questions), or even without any preceding question at all.

To the second counter-example, we might say that so long as an explanation-seeking why-question can be reconstructed – whether or not it was actually uttered – then the erotetic theory can be usefully employed. One reply to the first counter-example would be to limit explanations to answers in response to ‘explanation-seeking why-questions’ (a phrase which has, in fact, become common in the recent literature on explanation). However, this reply is analytically circular: with this move, ‘why-question’ has become a term of art, only specifiable relative to a prior understanding of ‘explanation’; for it presumes a prior ability to distinguish ‘explanation-seeking’ from ‘non-explanation-seeking’ why questions. As an *analysis* of explanation (as opposed to an illumination of certain features of explanation), the erotetic approach seems inadequate¹⁰.

¹⁰ It is disappointing that recent defences of the erotetic theory do not attempt to answer these questions. Matti Sintonen, for example, in his ‘Why Questions and Why Just Why-questions?’ in *Synthese* 120 (1999): 125-135 argues that it is the erotetic view which can best unify various insights into and approaches to explanation, since it is descriptively accurate with regard to scientific practice (126-128). But without, in particular, an answer to the problem of distinguishing explanation-seeking

It may be properly claimed that it is unfair to press the erotetic theory in the above way. Perhaps it is better to regard the erotetic theory as providing an analysis not of the explanatory relation, but of explanatory relata. The erotetic approach treats explanatory relata as questions and answers; and as we have seen, this response has the interesting consequence of focusing attention on the acts of explanation, and hence the context in which those acts are carried out. However, to better understand what constitutes a good relation between question and answer, we must rely on other resources.

Cognitive explanation

A cognitive approach to explanation would analyse explanation as that which brings cognitive benefit(s) of a particular kind. Equivalently, explanation is that which leads to understanding of a particular kind, or kinds. This approach has the benefit of bringing to the fore the fact that explaining is an *act* carried out for epistemic reasons, and that the goodness of the explanatory *product* is measured (over and above the basic truth of the explanans) according to epistemic criteria. It demonstrates why we seek explanation over and above simple knowledge of historical fact; for it brings a different cognitive reward. The key phrases requiring examination are, of course, ‘cognitive benefit’ and ‘understanding’; the wide variety of potential ways to conceive of these terms is reflected in the range of what may be called cognitive models of explanation. To the criticism that to analyse ‘explanation’ in terms of ‘understanding’ is to replace one awkward term with a worse one, I would say that many traditional analyses of ‘explanation’ may fruitfully be seen precisely as attempts to analyse (explanatory) ‘understanding’.

It should be recognised that turning to cognitively based criteria is an obvious, and popular, way to extend the erotetic approach. Sintonen, in asking what kinds of

questions from other questions, it is unclear how precisely the erotetic view *can* unify approaches to explanation.

questions are good requests for explanation, comes close to a cognitive analysis in stating that ‘Scientists value questions which, if answered, would further understanding’¹¹. Van Fraassen concludes his discussion of explanation by saying that ‘scientific explanation ... is an application of science. It is a use of science to satisfy certain of our desires’¹².

I shall mention here some of the more influential attempts to analyse ‘cognitive benefit’, or ‘understanding’: in terms of providing expectability or unification, removing a particular sort of cognitive predicament, and grasping appropriate internal rules. Each associated theory of explanation is familiar in philosophy of explanation literature, though the association which I draw between them is perhaps more unusual.

Hempel’s ‘Deductive-Nomological’ (D-N) analysis of (scientific, social scientific and historical) explanation¹³ is surely the best known of all modern approaches to explanation; it has certainly received the most comment and criticism. To recount the main points briefly: according to Hempel, an explanation should cite one or more empirical facts, and one or more laws, which together form non-redundant premises in a deductive argument whose conclusion is the fact to be explained.

Diagrammatically:

Particular facts: $C_1, C_2 \dots C_x$

General laws: $L_1, L_2 \dots L_x$

[Linking principles: *joining terminology of C_x with that of L_x*]

Conclusion (explanandum): E

¹¹ *Ibid.*, 128

¹² Van Fraassen, 156

¹³ See, for example, Carl G. Hempel, ‘Aspects of Scientific Explanation’, in Carl G. Hempel, *Aspects of Scientific Explanation, and Other Essays in the Philosophy of Science* (New York: Free Press, 1965)

The equally well known modification to the above model replaces the requirement of deductive implication with a requirement that C_x and L_x show E to be *highly likely*; such a model is termed Inductive-Statistical (I-S).

What cognitive benefit does such an explanatory argument bring? It shows that the explanandum was to have been *expected* (and hence that, given sufficient knowledge before the event, it could have been predicted). I believe that the appeal of Hempel's theory (aside from its close fit with a positivist philosophy, if that is considered an appeal) lies fundamentally in the deep connection felt between an explanation 'why', and the demonstration that the explanandum event had to happen.

Yet, as it has been demonstrated time and again in this context, whilst such a connection cannot be denied, this does not imply that the D-N model provides a satisfactory analysis of scientific explanation. Arguments showing that 'expectability' is neither sufficient nor necessary for 'explanation' form the bulk of the many criticisms of the Deductive-Nomological analysis. Expectability alone is insufficient, for the laws relied upon need to be 'of the right sort'.

The explanandum may be logically deduced from true premises of the appropriate sort, and yet not be explained by those premises. The fact that John is a person, and that all (or most) people who take birth control pills are not pregnant, implies the true fact that John is not pregnant. Yet, despite correctly conforming to the D-N schema, the explanans are irrelevant: John is a man! Bromberger's flagpole example¹⁴ demonstrates a special sort of irrelevance: that due to asymmetry. Whilst the height of a flagpole and position of the sun can be used to imply, and explain, the length of the shadow of that flagpole, the length of the shadow plus the position of the sun can imply, but not explain, the height of the flagpole.

Salmon has argued forcefully that neither is expectability necessary; he claims that the explanans do not have to imply that the explanandum had to happen, nor even (as with I-S) that it was highly probable¹⁵. All that is required of explanans is that they

¹⁴ Despite the notoriety of Bromberger's flagpole, apparently it was never introduced in any of Bromberger's published works.

¹⁵ See, for example, Wesley C. Salmon, 'A Third Dogma of Empiricism' in *Basic Problems in Methodology and Linguistics* eds. R. Butts and J. Hintikka (Dordrecht: D. Reidel, 1977)

show the explanandum to be *more likely* than it would have been without the presence of the explanans; in other words, that the explanans made a difference (with respect to the appropriate reference class).

A related approach is one which locates the cognitive benefit of explanation in *unification* of phenomena. Explanation analysed as unification emphasises the explanatory power of being able to reduce the number of independent theories required to deduce the phenomena we wish to explain. Such an approach is ‘global’, since we need to examine our theoretical framework in general in order to discern where unification, and hence explanation, has been achieved. Yet its similarity to Hempel’s approach can be seen in the central role given to deduction in both analyses.

The general link between the reduction of independent assumptions and explanatory power seems correct, but difficulties remain which mirror those of the D-N analysis. Consider, for example, that whilst disparate meteorological phenomena are unified through correlation with less disparate barometer positions, they are not thereby explained by the barometer positions. Of course, we believe that the explanation runs the other way: the meteorological phenomena explain the barometer positions. In general, just as with the D-N analysis, the laws which are intended to unify phenomena must be ‘of the correct kind’; statistical correlation is insufficient. To specify what ‘correct kind’ means seems impossible without resorting to a more fundamental explanatory notion: such as that the correlation instantiate *causal relations*. Such a move introduces an ontic conception, examined in the next section.

A different kind of cognitive approach was attempted by Bromberger¹⁶. Bromberger first locates a particular kind of epistemic predicament which, he claims, is uniquely present where explanation is requested. An explanation can then be defined as the removal of this type of predicament. A P-predicament exists when the recipient cannot conceive of a possible correct answer. For example:

¹⁶ Sylvain Bromberger, ‘An Approach to Explanation’ in *Analytical Philosophy, Second Series*, ed. R. J. Butler (Oxford: Oxford University Press, 1965): 72-105

(1): 'How many French revolted in 1789?'

(2): 'Why did the French revolt in 1789?'

Whilst one who asks (1) can conceive of many possible correct answers (various numbers of French people), the idea is that one who asks (2) cannot. Often, we will find that an explanation involves the removal of a P-predicament, yet Bromberger points out that this is not always the case: sometimes we *can* conceive of a possible answer to an explanation seeking why-question, for example. Therefore, there is a need for a second type of predicament: a B-predicament. This condition is intended to be more objective than the subjective nature of the P-predicament; a B-predicament exists when the recipient cannot conceive of *the correct* answer. Bromberger then defines an explanation as that which *either* removes a P-predicament, *or* removes a B-predicament.

Bromberger's approach, though I have categorised it as cognitive, proves a particularly fruitful extension of the erotetic idea. For, at the least, Bromberger's model allows us to be able to distinguish in a non-circular way 'explanation seeking why-questions' from 'non-explanation seeking why-questions', in terms of types of cognitive predicaments.

It is important to note that Bromberger's definition is not equivalent to more usual (scientific) definitions of explanation. In particular, the temporal nature of explanation usually emphasised is not a feature of Bromberger's model. The answer to

(3) 'What sort of life could there be on gas giants?'

would not usually be regarded as a request for explanation, since it does not require features which were previous to, and somehow led to, that which requires explanation. However, for Bromberger (3) could conceivably be an explanatory inquiry deriving from a P or B-predicament, since I cannot conceive of a correct answer (no more than I can to (2), at any rate).

It must be said that Bromberger's ideas have not been widely adopted. The reliance on the notion of a P-predicament places perhaps too much emphasis on the subjective,

and variable, notion of ‘what one can conceive of’; and leads to apparently odd categorisation of explanations, as (3) demonstrates. The introduction of ‘B-predicament’ is intended to remedy these deficiencies, at least in part. However, Bromberger’s insistence that this latter notion make reference to an objective ‘correct answer’ leaves the theory in the same place as the erotetic theory; the major task of specifying what a truly ‘correct answer’ amounts to is not tackled. In addition, we surely require a notion of ‘proper removal’ of an epistemic predicament in order to define explanation; since we must be able to rule out being hit over the head, brainwashed, or just given counselling to make us stop worrying, as explanatory episodes – even though they may indeed remove explanatory predicaments.

One final type of epistemic or cognitive benefit which I shall mention further demonstrates the breadth of cognitive approaches to explanation. Winch’s conception of a social scientific explanation is that which provides *understanding of internal rules*¹⁷, also known as ‘interpretivism’. A good explanation of a social action is one which describes social relations in a certain way: a way that makes clear the social context within which the action is performed, and which refers to the rules which the participants themselves use to orientate their behaviour. To take a popular example, an interpretivist answer to ‘why did that person exchange a piece of paper for some money?’ might be to describe what ‘cheque’ means in our society, and in particular, what the rules governing how cheques are treated by banks are.¹⁸

These *social rules* should be clearly distinguished from *scientific rules*. A social rule guides our action ‘from the inside’. A scientific rule, or law of nature, is a statistical correlation. An example of Winch’s may illustrate the difference: ‘The difference [between rules and laws] is precisely analogous to that between being able to

¹⁷ Peter Winch, *The Idea of a Social Science* (London: Routledge, 1958)

¹⁸ This interpretation of Winch’s ideas is the most common one, but is not the only one possible. For example, in a short article Sharrock and Anderson argue that Winch’s thesis concerns constraints upon *description* in the social sciences, not explanation (R. J. Anderson and W. W. Sharrock, ‘Understanding Peter Winch’ in *Inquiry* 28 (1985): 119-122). They take Winch to be arguing that to successfully describe an action means to make it intelligible; and this means showing it to be rule-governed. Showing action to be rule-governed, in Winch’s sense, is therefore a precursor to proper explanation, rather than an alternative form of explanation.

formulate statistical laws about the likely occurrences of words in a language and being able to understand what was being *said* by someone who spoke the language'¹⁹.

Criticisms of Winch's thesis have focused, firstly, on whether such an approach to social science is really necessary. Understanding social action in terms of rules formulated from the perspective of the participators themselves, and in their social context, is often clearly fruitful and important. However, is this the *only* way to provide social explanation? Interpretivists object to explanations citing structural factors which are not part of the conceptual scheme of the society whose actions are being explained, but it is unclear how this restriction can be defended. It is claimed that only an explanation in terms of social rules treats humans as purposive agents, rather than as lumps of matter. But even if this is the case, different perspectives on historical individuals and societies are allowable; at any rate, historians certainly *do* take widely different perspectives, as we shall see in particular in chapter six.

Secondly, a common criticism is that an approach such as Winch's is, despite first appearances, capable of subsumption under a scientific approach. In particular, if it can be established that interpretative explanations cite reasons, and that 'reason' is capable of analysis as a sub-species of scientific 'cause' in general, then interpretivism does not represent the radically different (anti-scientific) methodology that it is claimed to be.

The above summary has provided at least a flavour of the variety of cognitive approaches to explanation, though I am certain that the list could be extended. Rather than get enmeshed in the details of any one proposal, I wish to consider certain more general questions: Is it possible to conceive of multiple cognitive benefits which might thereby define what an explanation consists of? And if so, is there a well motivated way to delineate these types? In other words, what makes the above mentioned types of understanding *explanatory* modes of understanding?

Wesley Salmon's recent writing on explanation²⁰ certainly demonstrates a willingness to conceive of more than one type of explanation. His primary models are

¹⁹ *Ibid.*, 115

explanation through unification, and explanation through citation of causal mechanism. The latter type of explanation will be examined in the following section. What is important to note here is that Salmon treats both under a cognitive heading, such that each produces a complementary type of understanding: *understanding* of ‘the general scheme of things’, and *understanding* of mechanisms.

There does seem little a priori reason to reject the notion that there can be more than one cognitive ‘type’ of explanation. Focusing on Salmon’s two cognitive virtues of explanation through unification (the ‘Unificationist Thesis’) and explanation through citation of causes (‘the Causal Thesis’), Barnes attempts to refute the possibility, yet fails²¹. He claims that ‘the Unificationist Thesis and the Causal Thesis are ultimately at odds with each other, and thus cannot be easily conjoined (despite Salmon’s ... argument to the contrary)’²². To show this, Barnes imagines three phenomena to be explained (e1, e2 and e3); which are explained by three causes (C1, C2 and C3). Each cause is separate, and contributes only to its respective phenomena. In this case, the proponent of the Causal Thesis must affirm that there has been a gain in explanatory understanding, whilst the Unificationist must deny it.

Certainly this is true; the question is whether it is a problem. Barnes states that ‘the Unificationist and Causal Theses diverge in this case’²³; however, nothing in the possibility of multiple types of cognitive explanation precludes this. Indeed, it is entirely to be expected that the standards of different types of explanation would disagree over the explanatory power of particular examples, since they demand different properties from explanations. And, once we accept the possibility that there may be different types of explanation, we are at liberty to suppose that certain types may be more appropriate for certain cases. We need only dispense with the idea that there should only be one answer to ‘how good is that particular explanation?’

Salmon’s types of ‘scientific’ understanding are set within a wider array of types of understanding, yet it is unfortunate that Salmon gives us no reason *why* certain types

²⁰ For example, Wesley C. Salmon, ‘Scientific Explanation: Causation and Unification’ in *Critica* 22 (1990): 3-21

²¹ Eric Barnes, ‘Inference to the Loveliest Explanation’ in *Synthese* 103 (1995): 251-277

²² *Ibid.*, 264-265

²³ *Ibid.*, 265

are to be regarded as explanatory in a scientific sense. He does distinguish ‘psychological’ understanding from ‘scientific’ understanding on the basis that only the latter guarantees that the understanding is based upon ‘objective evidence’²⁴, but it is doubtful whether this condition can do much work. Just as with Salmon’s own criticism of Van Fraassen, we can point out that insisting that scientific explanations be true with respect to the best science of the day does not properly limit the types of relation which can be traced between those truths.

More charitably, we could interpret Salmon as being unwilling to impose *a priori* constraints upon types of explanation (or explanatory understanding), because he wishes to rely on *a posteriori* evidence concerning how scientists actually operate. I believe this is close to what Salmon would say; in any case, it is the answer I shall recommend. We should not limit – in a precise way – the sorts of understanding which may ground ‘historical explanation’ before examining historical explanation in practice.

To summarise the findings of this section: analysing explanation in terms of the sorts of cognitive benefits, or understanding, that it brings is a promising approach. I see no reason why we shouldn’t expect the *specifics* of such a programme to be dependent upon an empirical study of the explanatory practices of historians, in line with the philosophical methodology proposed in chapter one. We will need to look at what historians are aiming to do in order to delineate the cognitive virtues of explanation. However, we should also be entitled to expect general comments concerning why certain types of understanding are explanatory, and others not. I will return to this question in chapter six, after having examined the historical accounts in the above way.

Yet we have seen that cognitive approaches taken in isolation are susceptible to counter-examples. These generally take the form of an intuitive demonstration that the ‘explanation’, possessing the cognitive virtue demanded by the particular theory

²⁴ Wesley C. Salmon, ‘The Importance of Scientific Understanding’ in Salmon, *Causality and Explanation*, 90

of explanation, is nonetheless irrelevant, hence non-explanatory. A possible remedy to these counter-examples of this kind is to incorporate ontic notions.

Ontic explanation

An ontic approach takes explanation to be that which cites some underlying mechanism at work behind the phenomena to be explained. It has been argued that *causation* is paradigmatic of such a mechanism; but there may be other types of mechanism, including functions or goals, structural factors ('standing conditions'), relations of supervenience, or dispositions. To what extent does such an approach require substantive metaphysical commitment? At a minimum, it only requires that the appropriate mechanisms can be recognised; no particular metaphysical conception of these mechanisms need be elaborated for these purposes. However, it is necessary that the ontic mechanisms be regarded as more than purely statistical correlations: to an empiricist such as Van Fraassen, this commitment would be unacceptable.

An advantage of an ontic approach is that a clear line can be drawn between providing explanatory information, and providing descriptive information: only the former concerns relational mechanisms. A second benefit is that a minimal sense of relevance and asymmetry come built in to the analysis, given that relationships such as cause-and-effect are asymmetrical. Counter-examples of the sort which plague the D-N analysis (like Bromberger's flagpole) are thereby avoided. A final advantage is that the model may be commendably *a posteriori*, since the precise 'mechanism' can be specified according to the science of the day (this is, for example, Salmon's approach).

We must note that, strictly speaking, 'citing some underlying mechanism' should not always be necessary, even under an ontic approach. This should be replaced by 'citing information concerning underlying mechanism', in order to allow for the common practice of explaining by citing absences. ('Why did Blair get elected, rather than Hague?' 'Because he didn't make so many mistakes!') This need not be a

problem for an ontic approach; in particular, a theory of contrastive explanation (to be considered in the final section) can enable a particularly satisfactory understanding of explanation by absences.

One way the ontic position must be supplemented is with an account of the pragmatics of explanation. This is for the reason that a vast range of mechanism lies behind any given phenomena we wish to explain. For example, the event ‘Blair getting elected’ has a causal history stretching back to the formation of the universe; yet only a tiny fraction of this causal history will be truly *relevant* for any given demand for explanation of why Blair did get elected. Citing ‘the Big Bang’ does no explanatory work, except perhaps in the most peculiar of contexts, even though the Big Bang is in the causal history of Blair’s being elected. We need to understand ‘relevant’ in a pragmatic sense, since the Big Bang is as relevant as any other cause of Blair’s being elected from a purely ontic perspective. I believe the most satisfactory way to make precise this notion of pragmatic constraint as it relates to ontic explanation is through the use of contrastive explanation, discussed below.

One type of ontic model is provided by Aristotle, in his theory of the Four Causes²⁵. It is usually argued that Aristotle’s ‘causes’ are better interpreted as types of explanation, rather than strictly causes in the modern understanding of the term. The four types of explanation given by Aristotle derive directly from an Aristotelian metaphysics. They are explanation in terms of *substance* (or ‘form’: referring to the structure of that bringing the effect about); *matter* (as in ‘bronze is the cause of the statue’); *source* (or ‘motive’: the type most similar to the modern conception of ‘cause’); and *purpose* (or ‘final goal’, to which the item strives towards). It is usually appropriate to cite any of these four types in explaining a phenomenon, with certain exceptions: such as the inability to produce mathematical explanations on the basis of ‘source’. Hence to cite one type of Cause is not to exclude the citation of the others in explanation of the same event. Aristotle claimed that mistakes were made by previous philosophers precisely because they sought only one type of Cause. For

²⁵ Aristotle, *Metaphysics*, trans. John Warrington (London: Dent, 1956), Book A

example, Thales only sought explanation by ‘substance’, Anaxagoras only according to ‘purpose’.

Ruben’s interpretation of the Four Causes²⁶, which I am inclined to accept, strongly emphasises the ontic nature of Aristotle’s theory: Aristotle should be understood as demanding that an explanation should cite some underlying mechanism. Ruben himself provides a list of relations which can properly underpin explanations: causes, laws of co-existence, internal causation (inertia), self-explanation (mereological or reductive explanation), Cambridge dependency, supervenience and disposition²⁷. A similar ontic analysis of explanation is given by Jaegwon Kim²⁸.

The alternative ontic position would oppose the analysis of ‘underlying mechanism’ in terms of an extensive list such as Ruben’s. It would argue that we can restrict ontic explanation purely to causal explanation: such that an explanation cites some of the ‘causal mechanism’ or ‘causal history’ leading to the event to be explained. One of the clearest defences of such a position comes from David Lewis²⁹.

Any debate concerning the relative merits of the Aristotelian (‘florid-metaphysics’) position and Lewis’ (‘causal-reductivist’) will turn primarily on two, not unrelated, issues. Firstly, how broad is to be the definition of ‘cause’? Is it the case, for example, that the term ‘cause’ can encompass standing conditions, inertia, and Aristotle’s ‘substance’ and ‘matter’? Secondly, is it metaphysically possible to consider disposition, function, supervenience and the like to be reducible to causal structure? I am less concerned about these issues than I am about the standing of the ontic model with regard to the other approaches mentioned in this chapter, though I will return to it in chapter six. Suffice to say that causal explanation will be seen to be the paradigmatic sense of ontic explanation in general, though it may not be able to encompass the entire meaning of the latter.

²⁶ David-Hillel Ruben, *Explaining Explanation* (London: Routledge, 1990), chapter III

²⁷ *Ibid.*, 217-230

²⁸ In particular, Jaegwon Kim, ‘Noncausal Connections’ in *Nous* 8 (1974): 41-52

²⁹ David Lewis, ‘Causal Explanation’ (originally 1986) in *Explanation*, ed. Ruben

It seems sensible to combine the cognitive and ontic approaches in some way. Explanation is fundamentally a cognitive activity; it is that activity which produces a certain understanding. However, as we have seen, even if we can specify what ‘understanding’ is to mean in this context, not just any understanding of the correct type is truly explanatory. This is where an ontic criterion proves useful: for the demand to ‘cite an appropriate mechanism’ provides a way of understanding what more is needed for explanation. My conception of historical explanation, developed throughout this thesis but particularly in chapter six, will rely on both cognitive and ontic notions.

Can we, however, dispense with the cognitive element to the analysis, and rely solely on ontic (plus pragmatic) notions? Salmon thinks not, since ‘understanding’ can be deployed in more than one sense, only one of which refers to understanding of causes. Yet certain writers claim that an ontic analysis can dispense with the cognitive element. Such a conviction has been argued for in two different ways: by rebutting alleged examples of non-causal explanations, or, more positively, by demonstrating that all cognitive explanatory virtues reduce to ontic notions.

Lewis attempts to rebut supposed examples of non-causal (scientific) explanation of particulars, on a case by case basis, either through re-interpreting them as cases of causal explanation, or by arguing that they are not really explanations at all. I will not examine Lewis’ arguments in detail, but will mention two grounds for concern. Firstly, in his third case³⁰, Lewis argues that the supposed ‘explanation’ is, in fact a re-description: ‘Why is Walt immune to smallpox?’ ‘Because he possesses antibodies capable of killing the smallpox virus’. Lewis instead regards the possession of antibodies as being the disposition which *is* Walt’s immunity. But even if this is so, it seems clear that we *can* sometimes explain a phenomenon by re-describing it (a claim which will be substantiated in chapter six). If this is the case, then Lewis’ answer may not be sufficient to rebut the counter-example.

Secondly, Lewis claims that ‘understanding’ can be dispensed with, since ‘understanding why an event took place might, I think, just mean possession of explanatory information about it – the more of that you possess, the better you

³⁰ *Ibid.*, 190-192

understand'³¹. This simply doesn't seem true; indeed Lewis' phrase on the following page suggests a counter-example: "Who is Bob Hawke?" No need to write the definitive biography ...'

The second argument for a purely ontic notion of 'explanation' proceeds by attempting to demonstrate that the cognitive virtues underlying cognitive analyses of explanation in fact reduce to the purely ontic virtue of 'knowing the underlying causal mechanism'. If this is the case, then 'understanding' plays no substantive role in the analysis, and the only explanatory good is more knowledge of the causal history (just as Lewis claimed). This is the strategy of Barnes and Rappaport³², in their respective criticisms of Lipton's account of Inference to the Best Explanation. Barnes, for example, aims to reduce Lipton's explanatory virtues of 'precision' and 'coherence with background beliefs' to knowledge of causal mechanism, so that 'Lipton's various criteria of explanatory loveliness rest upon antecedently accepted principles of the world's overall causal structure'³³.

Linguistic explanation

Hayden White's ideas are possibly less well known to analytic philosophers of explanation than those already introduced. However, the influence of those ideas on the philosophy of history of the past thirty years has been immense; and, in any case, much of what White has to say on narrative and explanation has a bearing on the central questions of this thesis.

I call White's philosophy of history 'linguistic' since it focuses purely on the *discourse* of historians. For White, historical explanation is analysed as deriving from the linguistic devices a historian uses. An explicit part of this approach is that the

³¹ *Ibid.*, 195

³² Steven Rappaport, 'Inference to the Best Explanation: is it Really Different from Mill's Methods?' in *Philosophy of Science* 63 (1996): 65-80

³³ Barnes, 274

explanatory concepts of the discourse do not refer to anything in the historical record. There are, in the historical past itself, no narrative structures, no historical laws, no relations of importance or significance. These items are not discovered by the historian in the historical past; rather, they are rhetorical devices created by the historian in their writing. White described and compared these rhetorical devices in some detail; an outline of his four-fold categorisation of explanatory devices follows³⁴.

First, the historian explains the historical events through *emplotting* those events in a narrative form which takes a familiar shape. These forms include *Romance* (featuring a hero who succeeds by transcending the world of experience); *Tragedy* (terrible fates are suffered, yet lessons can be learnt by the spectators of these events); *Comedy* (temporary triumph is achieved through a reconciliation of sorts); and *Satire* (the opposite of Romance: ‘man is ultimately a captive of the world rather than its master’³⁵). Note that the similarity of fictional-narrative forms and historical-narrative forms is, for White, quite intentional.

Second, the historian explains through *formal argument*. This mode of explanation is more akin to theories of scientific explanation, especially Hempel’s D-N theory. However, White wishes to take into account the fact that ‘historians disagree, not only over what are the laws of social causation that they might evoke to explain a given sequence of events, but also over the question of the form that a “scientific” explanation ought to take.’³⁶ For this reason, explanation by ‘scientific’ laws represents only one option: the *mechanistic*. The other three possibilities are *formist* (explanation proceeds by identifying the unique aspects of that explained); *organicist* (explanation proceeds by identifying the common final goal which unifies the events explained); and *contextualist* (explanation proceeds by tracing the threads which connect the event to be explained to its wider context).

³⁴ The major source for these ideas is the Introduction to White’s *Metahistory*. In this work, White examines eight nineteenth-century historians, but it is clear that White believes that a similar notion of historical scholarship is applicable more generally.

³⁵ *Ibid.*, 9

³⁶ *Ibid.*, 12

Third, the historical events are always provided by the historian (often unconsciously) with an *ideological implication*, which serves to explain in a moralistic sense. The four options which White conceives of are *Anarchism*, *Conservatism*, *Radicalism*, and *Liberalism*. In addition to these three basic explanatory devices found within the historian's discourse, there are differing *historiographical styles*. These styles, or 'tropes', are ways of verbally categorising the historical past, prior to deploying the three explanatory strategies. These explicitly rhetorical devices are *metaphor*, *metonymy*, *synecdoche*, and *irony*; and each style tends to be associated with particular explanatory types from those above. (I will say no more concerning the categorisations of ideological implication and historiographical style, since they will be of less relevance to the current study.)

It is possible to question White's over-formalism (that he is overly restrictive of the possible options in each category of explanation), and perhaps his reliance on an out of fashion and over-structural literary theory. Debate has also raged over the extent to which narratives are created by the linguistic practice of the historian, as White claims, or are discovered in the historical past. However, more important in the context of this thesis is to ask in what sense White's philosophy *is* a philosophy of explanation.

The only general answer is that all these types of explanation somehow endow the events of the historical account with a sense of *meaning*, through the use of various rhetorical devices. (White's account could, therefore, be seen partly in terms of a cognitive theory of explanation.) Explanation by ideological implication explains by providing a *moral* meaning to the set of events. Explanation by emplotment explains by providing a *familiar* meaning to the events (it is claimed that the various types of plot are deeply embedded within our shared culture; even that they are culture-transcendent³⁷).

³⁷ Hayden White, 'The Value of Narrativity in the Representation of Reality' in Hayden White, *The Content of the Form: Narrative Discourse and Historical Representation* (Baltimore: John Hopkins University Press, 1987), 1

Explanation by formal argument provides meaning by providing “‘the point of it all” or “‘what it all adds up to”’³⁸. Little more can be said about explanation by formal argument in general, considering the variety in the types of formal argument. In particular, a type such as ‘formist explanation’, which explains not according to unity but particularity, is particularly hard to explicate in terms common to the other types. Even White, at times, struggles to comprehend formist explanations as genuinely explanatory. Commenting on the transition from organicist to formist history, exemplified in the work of Ranke, White says that

‘The transition [from organicist to formist] can be characterised as a modification in which the *impulse to explanation* is sublimated to a desire simply to *describe* the process as it unfolds before the historian’s gaze.’³⁹

Stating *why* precisely the various rhetorical devices should be regarded as explanatory devices represents something of a gap in White’s writing. One of White’s few statements on this topic is to maintain that narrative (and presumably the other rhetorical devices) are the only way to move beyond simple representation. Following Barthes, White states that ‘narrative ceaselessly substitutes meaning for the straightforward copy of the events recounted’⁴⁰. Later in the same essay, he makes it clear that ‘meaning’ should be taken, primarily, as ‘moral meaning’: ‘it seems possible to conclude that every historical narrative has as its latent or manifest purpose the desire to moralize the events of which it treats.’⁴¹

The breadth of White’s scheme may be admired as a virtue, since quite different historical approaches are treated on their own terms, free from the influence of a restrictive, a priori, conception of explanation. Yet this very breadth is liable to invite criticism. In particular, we are surely entitled to ask what these different types of ‘explanations’ have in common; in other words, what makes them (and not other rhetorical devices) explanatory. To the extent that White does provide an answer – in

³⁸ White, *Metahistory*, 11

³⁹ *Ibid.*, 186

⁴⁰ White, ‘The Value of Narrativity’, 2

⁴¹ *Ibid.*, 14

terms of providing meaning – we can ask whether all ways of providing meaning are explanatory.

I hope, primarily in chapter six, to develop an understanding of historical explanation which can account, as does White, for the range of explanations produced by historians. Yet, this understanding proceeds by identifying at least relevant, and substantive, similarities between types of explanation. In addition, in order to properly account for explanatory exclusion and competition, I will argue that we need to abandon White's overt and extreme anti-realism, and his emphasis on the moralising nature of all historical writing.

Contrastive explanation

The idea that explanation is contrastive amounts to the thesis that the explanandum is not represented by a simple proposition 'p', but a complex represented by the phrase 'p rather than q (or q₁, q₂, q₃ ...)'⁴². The terminology: p is the fact, whilst q is the foil, or contrast. The overall complex I call the explanandum, or equivalently the explanatory target. Contrastive explanation has become an influential and well developed body of theory since the idea was first introduced by Dretske⁴³, and then developed by Van Fraassen⁴⁴, Garfinkel⁴⁵, and Lipton⁴⁶

Van Fraassen introduced the notion of contrast in the context of his erotetic theory of explanation, and contrastive explanation and erotetic explanation have been

⁴² Contrastive theory might be applied to other concepts as well as 'explanation'. 'Knowledge' is an obvious candidate: one may know 'that p (rather than q)' – i.e., be able to distinguish p and q – and yet not know 'that p (rather than r)'. This may be useful in understanding sceptical arguments designed to demonstrate our inability to distinguish between, say, the real world and a certain sort of illusion.

⁴³ Fred I. Dretske, 'Contrastive Statements' in *Philosophical Review* 81 (1972): 411-437

⁴⁴ Van Fraassen, chapter 5

⁴⁵ Garfinkel, chapter 1

⁴⁶ Peter Lipton, 'Contrastive Explanation' (originally 1990) in *Explanation*, ed. David-Hillel Ruben (Oxford: Oxford University Press, 1993): 207-227

commonly treated as a package. However, I would argue that it is preferable to consider contrastive explanation as a thesis concerning *what precisely it is that gets explained* (the explanandum, or target). It is, therefore, a thesis which can be used to supplement any of the above four approaches. The erotetic approach would speak of contrastive why-questions; the causal (ontic) approach would refer to causal information required to explain a contrastive phenomenon. In analysing ‘contrastive explanation’, reference to why-questions is redundant, as Markwick points out⁴⁷. I do commonly refer to ‘contrastive questions’ and ‘contrastive answers’ in what follows, but this should not be taken to mean that those phrases are analytically basic. A contrastive question is one which has as its object a contrastive target; a contrastive answer is one which answers a contrastive question.

In what follows, I will firstly discuss contrastive *questions*: their form, and constraints upon good questions. Next, I will discuss contrastive *answers*: in particular, criteria of adequacy governing replies to contrastive questions.

A contrastive question is of the form ‘why p, rather than q (or q₁, q₂, q₃ ...)?’. Such a question presupposes (i) that the fact is true, whilst the foils are not, and (ii) that at least one of the set had to happen. The earlier view (Garfinkel, Van Fraassen) maintained the presupposition that one *and only one* of the set had to happen; I will follow Lipton in dispensing with this exclusivity requirement, for the following reason.

Lipton was the first to point out that there are different types of contrastive question – compatible and incompatible⁴⁸. In the case of *compatible* contrasts, it is possible that both fact and foil could have been realised; though, as it turned out, only the fact was. *Incompatible* contrasts, however, preclude each other: there is no way that both fact and foil could have been realised (though it could have been the case that what we call the foil was realised whilst what we call the fact was not). This distinction is particularly relevant to historical questions, for there are examples of each. ‘Why was

⁴⁷ P. Markwick, ‘Interrogatives and Contrasts in Explanation Theory’ in *Philosophical Studies* 96 (1999), 188-9

⁴⁸ Peter Lipton, ‘Contrastive Explanation’, 211

there a revolution in France in 1789, rather than in England?’ is compatible; ‘why was there a revolution in France in 1789, rather than 1787?’ is incompatible⁴⁹.

Contrastive explanation provides a framework for understanding the role of implicature and the pragmatics of explanation. The contrast or contrasts are frequently not explicitly specified; this is because in most situations, the contrast is very obvious. The usual contrast to ‘why did you hand your essay in a day late?’ is unlikely to be ‘rather than handing in something else’, though this may be implied in certain contexts. More likely is the contrast ‘rather than handing in the essay on time’. The contrast is implied by the pragmatic context, and is then used to shape the relevance of potential answers. In discussion of the ontic conception, I mentioned that it is necessary to provide a pragmatic understanding of explanatory relevance; I will argue in chapter five that the purpose of contrastive explanation is to focus the explanatory information which is relevant in a particular explanatory context.

Whilst the ultimate guide to what contrast is intended can only be understood relevant to the particular explanatory situation, there are certain features which make contrasts more or less *useful*. A good contrastive question is typically one in which the foils are relevantly similar to the fact. The fact and foils should share many features, allowing attention to be focused upon the fewer differences. For example, whilst Skocpol’s question ‘why was there a revolution in France in the late eighteenth century, rather than in England?’ is enlightening, the question ‘why was there a revolution in France in the late eighteenth century, rather than in New Zealand?’ is anything but.

Barnes notes that the criterion of similarity between fact and foil is necessary but not sufficient to constrain sensible contrastive questions⁵⁰. Barnes asks us to imagine that if ‘Bush wins the ’88 election’ (P) had not obtained, then ‘Bush was depressed’ (Q) would have; further, in actual fact P turned out true. P and Q presumably have similar causal histories, right up to the point where Bush won the election, where the histories

⁴⁹ We can add that, though useful, the distinction between compatible and incompatible is not as sharp as Lipton seems to presume. The fact and foil may preclude each other to any given degree; we might call these ‘lifeboat cases’. If Jill and Jane are stranded in a lifeboat wherein only a small amount of food is available, the question ‘why did Jill survive, rather than Jane?’ will logically be a compatible one, yet a compatibility which is, in practice, unlikely since the survival of one radically reduces the chances of survival of the other.

⁵⁰ Barnes, 46-47

of P and Q diverge. Therefore, the similarity condition is fulfilled. Yet the contrastive question ‘why did Bush win the ’88 election, rather than getting depressed?’ is not a sensible one to ask; unlike ‘why did Bush win the ’88 election, rather than narrowly losing it?’. Barnes proposes to add a further criterion for good contrastive questions: that ‘P and Q are culminating outcomes of some single type of natural causal process’⁵¹. Of course, what ‘same type’ means will, in part, depend on pragmatics, specifiable according to investigative interest.

A contrastive question focuses attention on the *difference* between fact and foil(s). A first, approximate, formulation for a condition of adequacy that a contrastive answer must meet is that the answer must provide such information, and only that information, which distinguishes between the occurrence of the fact and the foil(s). But this is very rough; in particular, we need an accurate analysis of ‘distinguish’ in the previous sentence. I shall consider three possible explications: Lewis’ counterfactual analysis, Lipton’s Difference Condition (DC), and Van Fraassen’s (and Hitchcock’s⁵²) probabilistic relevance model.

Lewis claimed that to explain ‘why P rather than Q’ is to *provide a cause of P which would not have also led to Q*. For example, to explain

(MO) ‘why L. gave a talk at Monash University rather than giving a talk at Oxford University’

we might cite the fact that ‘the climate is good in Monash at that time of year’⁵³. Lipton pointed out⁵⁴, however, that we could imagine a particular cause of P (say, ‘being invited to speak at Monash’) which would not also have led to Q; yet could imagine such a cause to be non-explanatory. In terms of MO, this would be the case where an invitation also came from Oxford. Lewis’ analysis, therefore, does not pick out a sufficient condition for explanatory adequacy.

⁵¹ *Ibid.*, 50

⁵² Christopher Hitchcock, ‘Contrastive Explanation and the Demons of Determinism’ in *British Journal for the Philosophy of Science* 50(4) (1999): 585-612

⁵³ Lewis, ‘Causal Explanation’, 196

⁵⁴ Lipton, ‘Contrastive Explanation’, 215-216

The natural response to the sort of counter-example offered is that we should modify Lewis' formulation so as to demand a cause of P, *a type of which* would not have also led to Q. This would indeed solve the counter-example where both universities invited L., but Lipton sees a further problem. He envisages a case where there is a cause of Q which is a necessary cause: for example, suppose that an invitation is a necessary condition of speaking. In this case, an invitation could still have been the explanatory cause of speaking at Monash rather than at Oxford, since maybe only Monash issued an invitation. Yet a cause of that *type* certainly *would* have led to speaking at Oxford. Thus, the modified counterfactual analysis now does not pick out a necessary condition for explanatory adequacy.

The net result of this argument is that a counterfactual model will not do; we need to pay attention to the *actual* history of the foil. Lipton proposes that to explain 'why P rather than Q', we select an actual difference between the causal histories of P and not-Q : where not-Q is the actual history which represents the failure of the foil to be realised. There should be given a 'cause of P and the absence of a corresponding event in the case of not-Q'⁵⁵. This is the Difference Condition (DC) for adequacy of answers to explanation-seeking contrastive questions.

DC provides detail to the rough idea that answers should somehow distinguish between fact and foil(s), and as such is a major advance in our theory of contrastive explanation. It is, essentially, the analysis I will adopt for the subsequent use of contrastive explanation in this thesis. However, I do wish to make three comments which should sharpen understanding of DC.

Firstly, DC clearly applies only to causal explanation, due to the language used; yet I believe contrastive explanation to be a type of explanation applicable to any relation thought to be potentially explanatory. However, it would seem that only a simple modification would be required to extend the applicability of DC beyond causal explanation. Any ontic explanation will maintain a set of facts explanatorily relevant to the happening of P, and a further set relevant to the happening of not-Q. Thus,

⁵⁵ *Ibid.*, 217

generalising DC, an answer to ‘why P rather than Q’ requires an actual difference between the explanatorily relevant facts pertaining to P and to not-Q, where this fact is explanatorily relevant to P and a corresponding fact is absent in the case of not-Q.

Secondly, we should recognise that DC does not restrict differences between the fact and negated foil to a single one. There may be many relevant actual differences between the histories of P and not-Q. Some of these differences may be more salient than others – an idea developed in Part III.

Thirdly, DC handles cases of compatible contrast better than it does those of incompatible contrast. In the case of incompatible contrasts, it *seems* (given just what I have stated above) that DC asks us to find a difference between the history of the fact, and the history of the negated foil. Yet, where the two options are mutually exclusive, the history of the fact refers to much the same information as does the history of the failed foil. Consider

‘why did L. go to Monash, rather than not going?’

What is in the history of failing to ‘not go to Monash’, if it is not just going to Monash?! There is, in this case, no way for a cause to be in one history and not the other.

Lipton recognises this difficulty, and consequently responsibility falls to the notion of ‘corresponding event’⁵⁶. In cases of incompatible contrast, DC should not require us to simply find a cause of P absent from the history of not-Q. Rather, it requires a cause of P which does not have a *corresponding event* in the history of not-Q. A corresponding event in the history of not-Q is ‘something that would bear the same relation to Q as the cause of P bears to P’⁵⁷.

As Hitchcock points out⁵⁸, the notion of ‘same relation’ is not an exact one. Hitchcock’s example introduces Adam, who orders a lasagne rather than a salad. It would be plausible to explain this choice by stating that Adam was hungry; even though the hunger was in the history of both ordering the lasagne and not ordering the

⁵⁶ *Ibid.*, 217

⁵⁷ *Ibid.*, 217

⁵⁸ Hitchcock, 596

salad, and would have been a cause of eating the salad had Adam ordered this. The reason that the hunger is nonetheless a relevant difference is that Adam's hunger bore 'more of a relation' to ordering the lasagne than it did to ordering the salad. Recognising the vagueness of 'corresponding event' is not necessarily a fatal concession for DC to make. We only need recognise that answers to incompatible contrastive questions can be better or worse, due to the cause being more or less corresponding.

Finally, we should consider the 'probabilistically relevant' (PR) notion of contrastive adequacy as a possible competitor to DC. Hitchcock offers PR in response to the above criticism of 'corresponding event'; Van Fraassen's account is similar. Van Fraassen's contrastive condition of adequacy on an answer A to explain 'why B (rather than ...)' is 'on the basis of how well A redistributes the probabilities on the contrast-class so as to favour B against its alternatives.'⁵⁹ Hitchcock takes an answer to a contrastive question to be an attempt to provide a probabilistically relevant factor, given the presupposition that either the fact or (to be read inclusively) one of the foils must have been the case⁶⁰. The answer must 'redistribute probabilities' in the correct way: it must increase the probability of the fact happening over the foil(s).

One advantage of PR is claimed to be that it is easily able to accommodate the changes in degree which we saw were necessary in our discussion of Lipton's notion of 'corresponding', since an answer may increase the probability of the fact over the foil(s) by any degree⁶¹. The problematic notion of 'correspondence' is to be replaced by the tidier (and potentially quantitative) notion of 'probabilistically relevant'. It is worth examining this claim. At first sight, DC and PR perhaps seem equivalent, so long as causal explanation in general is equivalent to the citing of facts which raise the probability of the effect⁶².

⁵⁹ Van Fraassen, 147

⁶⁰ Hitchcock, sections 7 and 8

⁶¹ *Ibid.*, 602

⁶² We might question this assumption in two ways: firstly, by asking whether all causes raise probability. Secondly, we might ask whether all raising of probability is achieved causally. Other means include as a result of geometrical properties (as in Lipton's example of the orientation of thrown

Yet, on closer inspection, the two approaches yield different kinds of answers. In discussion of a structurally similar case to MO (Lewis' Monash/ Oxford decision), Hitchcock asks 'why Jones, rather than Smith, has paresis'⁶³. According to PR, the correct reply is 'the information that Jones had latent untreated syphilis, but Smith did not'. The subtle, yet crucial, difference between this formulation and that of DC is that the answer does not cite a relevant *cause*, but instead provides probabilistically relevant *information*.

This locution suggests counter-examples to PR. For example (back to L.'s decision), suppose that Monash has a good climate at this time of year, whilst Oxford has excellent real ale pubs. Suppose further (not implausibly) that these facts were equally likely to tempt Lewis to visit the respective universities. Then the information given in sentence S: 'Monash has a good climate at this time of year and Oxford has excellent real ale pubs' is not probabilistically relevant to (4). Therefore, according to PR, S is absolutely no good as an explanation. However, it seems that S *is* more of an explanation than the reply T: 'both Monash and Oxford invited Lewis'. The latter is not only probabilistically irrelevant (as is S), but, unlike S, provides no explanatory causes at all.

The difference between S and T cannot be found in their probabilistic relevance, which is equal: both are zero. Rather, the difference is that in the first statement the causes stated are of *different types*; in the second they are not. But to understand this feature as explanatorily relevant, we require a notion of 'correspondence'; probabilistic relevance will not account for the difference. The notion of 'correspondence', or 'same relation' remains necessary; it cannot be replaced by 'probabilistically relevant'. I suggest instead that probabilistic relevance may do some work as a condition of the degree of *goodness of contrastive answers*, requiring DC as a prior condition of adequacy. This suggestion will be examined in chapter eight.

sticks, 'Contrastive Explanation', 208; discussed in chapter six); or even as a result of logical relation. 'Abel's death' raises the probability of 'Brenda being a widow' (from 0 to 1), but does not thereby cause Brenda's widowhood.

⁶³ Hitchcock, 605

The account of historical explanation which I shall recommend draws primarily on the ontic and cognitive traditions. I will not attempt a defence of this account at this point, since it will be developed in conjunction with an examination of the case study explanations. However, the outlines can be sketched: an analysis of historical explanation must emphasise (i) the need for relevant causes (understood in a broad sense), or other relational ontic feature and (ii) the need for proper conceptualisation of those causes. (i) seems vital in marking out what is distinctive about *explanatory* information, and in dissolving problems of asymmetry which plague solely cognitive accounts. 'Relevant' in (i) is to be understood with reference to the constraints of contrastive explanation, and hence is primarily a pragmatic term.

Yet those who would place all the weight on ontic notions (even if supplemented by pragmatic rules) underestimate the role of re-conceptualisation in explanation. Even if historians cite correct and relevant causes, not every way of doing this will be explanatory. In addition, 'explanation by re-description' is particularly prevalent in historical scholarship, and this must be accounted for.

I will develop this dual understanding of explanation in chapters five and six. Chapter five will be concerned with the criterion of causal relevance: I shall rely heavily on the theory of contrastive explanation, and aim to demonstrate the applications of this notion with regard to explanatory exclusion. Chapter six will be concerned with the criterion of proper conceptualisation, and its link with explanatory competition. I will, at that point, investigate the cognitive virtues which underpin the historical explanations under examination. The overall aim is to adopt an account of historical explanation which can do justice to the wide variety of historical explanations found in the case study, and in historical scholarship more generally.

Part II

Chapter Four: Principles of Explanatory Exclusion

Part I was concerned with developing the methodological basis for the present work, introducing the historical material to be used as a case study, and providing an overview of approaches to analysing ‘explanation’. The primary focus of Part II will be *explanatory exclusion*.

It is helpful to begin by considering two previous attempts to analyse the idea of explanatory exclusion. Each proceeds from a different understanding of explanation, as introduced in chapter three. The first ‘Principle of Explanatory Exclusion’ (PEE hereafter) is Jaegwon Kim’s. It is based on the idea that explanation picks out relevant features of the world – it is therefore an ontic conception. The second PEE is based upon Bas Van Fraassen’s erotetic theory of explanation. After a brief examination of some examples of competition and non-competition in our historical examples, I will then develop my own, more general, notion of explanatory exclusion.

Ontic Principle of Explanatory Exclusion

A good first step towards comprehending explanatory exclusion is to focus on the targets (explananda) of the explanations. Where two explanations have the same target we might

think that their relationship is thereby a candidate for exclusion. However, not all explanations with the same target thereby exclude one another: a first explanation which cites the icy road as a cause of the accident does not exclude a second which cites the tiredness of the driver.

Having the same target seems relevant to explanatory exclusion then, but not sufficient (and, I will argue, not strictly necessary either). An attempt to specify what else is required for exclusion is given by Kim's Principle of Explanatory Exclusion (which I shall subsequently abbreviate KPEE):

'there can be no more than a single complete and independent explanation of any one event, and we may not accept two (or more) explanations of a single event unless we know, or have reason to believe, that they are appropriately related - that is, related in such a way that one of the explanations is either not complete in itself or dependent on the other.'¹

Kim deploys this principle to debate the status of causal mental explanations as compared to causal neuro-physiological explanations; yet his PEE is intended to be general towards both subject matter and nature of the explanatory relation. In particular, it is not intended to be restricted only to causal explanation. Clearly, in order to assess KPEE, we need to be able to properly interpret certain key terms: in particular, 'complete' and 'independent'. Kim's substantive attempts to define these terms are based on making apparent the link between the KPEE and explanatory realism.

Kim defines explanatory realism thus: '*C* is an explanans for *E* in virtue of the fact that *c* bears to *e* some determinate objective relation *R*. Let us call *R*, whatever it is, an "explanatory relation."² We should make clear at this point that, under Kim's version of explanatory realism, 'e' – what is explained – stands for an event. Given a commitment to explanatory realism of this kind, it is claimed that KPEE follows naturally. This is

¹ Jaegwon Kim, 'Explanatory Realism, Causal Realism and Explanatory Exclusion' in *Midwest Studies in Philosophy* 12 (1988), 233

² *Ibid.*, 226

shown by attempting an exhaustive consideration of all possible relationships between multiple ontological ‘relations’, and hence (given explanatory realism) between multiple explanations. The aim is to demonstrate that there are only three possibilities, corresponding to those mentioned in the definition of KPEE. Either relations (hence explanations) exclude, or they are incomplete, or they are dependent.

Consider two explanations of the same event e ; explanation one refers to fact c 's appropriate relation to e , whilst explanation two refers to fact c^* 's appropriate relation to e . There are five possible relations between c and c^* , and hence (given explanatory realism) five possible relations between explanations one and two:

1. $c = c^*$. The explanations refer to the same items, even though different terms may be used to pick these out.
2. c is supervenient on c^* .
3. c and c^* are not independent; for example, c is an earlier link in a causal chain which comes to include c^* , hence c^* is dependent upon c .
4. c and c^* are each components of a wider explanatorily relevant event; for example, c and c^* are both necessary but insufficient causes comprising a jointly sufficient set. Hence neither c nor c^* is complete.
5. c and c^* overdetermine e . In this case also, Kim claims that neither c nor c^* is complete.

Cases one and two are those in which we actually have only one explanation, not two. Case three describes dependent explanations, which are excluded from KPEE. Cases four and five are cases of incomplete explanations, also excluded from KPEE. To derive KPEE, therefore, we need to assume two things: firstly the realist principle that

explanatory relationships mirror ontological relationships; secondly, that the above five possibilities map all possible ontological relationships³.

Any defence of the latter premise is likely to be inconclusive, since it is only established to the extent that we cannot think of any possible relationships other than the five mentioned by Kim. I will mention only one potential counter-example here⁴. Two explanations may be linked by the relationship of *instantiation*. Such examples are each (potentially) complete, and are not dependent (in the sense of one being a cause of the other), yet do not lead to exclusion. As such, they appear to be a counter-example to KPEE. An example: ‘Why did Able die?’ Explanation (i) states that it was ‘Because of Briar’s passionate jealousy’; explanation (ii) that it was ‘Because of Briar’s stabbing of Able’. Explanation (ii) is an instantiation of explanation (i), specifying the detailed mechanism by which (i) took place.

Whilst an explanatory relationship of instantiation cannot be straightforwardly accommodated by KPEE, I will argue in chapter five that such examples can be shown to be comprehensible according to the general intuition behind KPEE. Still, the issue does demonstrate that development of Kim’s PEE in the way that I shall suggest is not an optional extra, but necessary if KPEE is not to be susceptible to counter-example. At this point, however, I would now like to question three more general features of the above model. These are: first, the treatment of over-determined explanation in KPEE; second, the applicability of a defence of KPEE outside the realm of causal explanation; third, and most importantly, the notion of a complete explanation.

³ Thagard’s comments on explanatory competition are similar in most respects to Kim’s. For example, in his most recent paper on the topic (Paul Thagard, ‘Probabilistic Networks and Explanatory Coherence’ in *Cognitive Science Quarterly 1* (2000): 91-114) Thagard provides two necessary and jointly sufficient conditions for explanatory exclusion: that the explanations explain the same explanandum; and that the explanations are not ‘explanatorily connected’ (Thagard, 97). This phrase is intended to cover cases where one explanation is a part of the other, or both are part of a greater explanatory whole; hence this condition corresponds to Kim’s demand for ‘completeness’. Thagard overlooks the necessity for a requirement of independence.

⁴ Suggested to me by George Botterill.

It is clear that Kim's treatment of over-determining causes is unsatisfactory. Indeed, given Kim's strategy of defining explanatory relationships with reference to ontological relationships, the claim that two over-determined explanations are individually incomplete seems inconsistent and unmotivated. The meaning of 'over-determined' is that the two causal processes are complete antecedents for the same effect, and are independent. If this is so, then over-determining explanations must also be excluded from KPEE (though, Kim does indicate at one point that he is prepared to consider this possibility⁵).

Of course, the relevance of this criticism depends entirely on whether there are many (or any) genuine cases of causal overdetermination in the historical subject matter. And this issue itself depends largely on whether we take explanatory targets to be broad or narrow: if the former, there is more scope for potential cases of overdetermination. The examination of the nature of explanatory targets, in the following chapter, will address this question.

Secondly: whilst it is Kim's intention that KPEE be general with respect to explanatory relation, it is not clear how this can be supported. As we have seen, substantive support for KPEE has only been provided through the investigation of possible causal relationships, thereby demonstrating that causal explanations of the same event are either exclusive, dependent, or incomplete. The five possibilities envisaged by Kim do perhaps seem exhaustive of causal relationships (though, as I have suggested, this can hardly be proved conclusively). Yet it is not at all clear that these correspond in any straightforward manner with explanations featuring other explanatory relationships – for example, teleological, functional, or rational. Still more problematic would be the comparison of different explanatory relationships; for example, where there are two explanations of the same event, one of which is causal, the second teleological. Are these to be excluded? Are they complete or not, independent or not? This is an issue which I

⁵ Jaegwon Kim, 'Mechanism, Purpose and Explanatory Exclusion' (originally 1989) in Jaegwon Kim, *Supervenience and Mind: Selected Philosophical Essays*, ed. Ernest Sosa (Cambridge: Cambridge University Press, 1993), 253

shall explore in chapter six. For now we must admit that, to the extent that it *is* a good theory, KPEE has only been shown to be plausible in accounting for causal explanation.

‘Man’s mind cannot grasp the causes of events in their completeness, but the desire to find those causes is implanted in man’s soul.’⁶ – Tolstoy

What are we to make of the notion of a *complete* explanation? Let us focus on causal explanations, in line with my second observation. In this case, two potential meanings of ‘complete’ immediately come to mind. First, we could take a complete causal explanation to be one which cites all antecedents of the target event. Famously, Mill regarded only such a complete set as the ‘real Cause’ of an event⁷. Yet, considerations similar to Mill’s suggest the impossibility of ever listing such a set, and therefore the impossibility of providing a complete explanation. In imagining a set of conditions sufficient to produce a particular death, Mill lists

‘the act of eating the dish, combined with a particular constitution, a particular state of present health, and perhaps even a certain state of the atmosphere ... all the conditions were equally indispensable to the production of the consequent; and the statement of the cause is incomplete, unless in some shape or other we introduce them all’⁸

Not only are there many causal antecedents of any given event, it is clear that causal explanation must also utilise negative causal information; in other words, causal explanations also mention the absence of certain factors. If the unfortunate person in Mill’s example usually took a particular medicine before eating, but before the fateful meal in question did not, then an explanation of the death could not be considered complete without including amongst the antecedents the fact that ‘the person did not, before eating the dish in question, take their regular medicine’. Yet if this is to be

⁶ Leo Tolstoy, *War and Peace*, trans. Louise and Aylmer Maude (Oxford: Oxford University Press, 1991; originally 1868-9), 1055

⁷ John Stuart Mill, *A System of Logic, Ratiocinative and Inductive* (Charlottesville: Lincoln-Rembrandt, 1996; originally 1843), Book III, Chapter V, Section 3

⁸ *Ibid.*

allowed, then the causal information required for a complete explanation of this sort is not only lengthy, but unlimited: for there can surely be no exhaustive list of absent causes. In that case, it is unclear how completeness can function even as an ideal standard for explanations. Finally, the notion of a complete explanation can appear still more ridiculous when we consider that causal chains stretch back indefinitely in time: should a complete explanation therefore mention *all* causes of a given event?

A second, and more promising, interpretation of ‘complete’ is to make use of Deductive-Nomological explanation. Consider the explanation

(1): ‘JS ate dish x and had condition y; everyone who eats dish x and has condition y dies; therefore JS died.’

Or (2): ‘JS ate dish x and had condition z; everyone who eats dish x and has condition z dies a painful death’.

If such explanations are available then they are also, in one sense, complete explanations of JS’s death. The sense in which these are complete is that the premises are *sufficient* to produce the target. This reading of ‘complete’, as meaning ‘sufficient’, is at least more workable than ‘complete’ as meaning ‘all causal information’, since sufficient explanations are certainly possible in practice.

Yet Kim is not keen to interpret ‘complete’ in this way: he is concerned that to do so would be in contradiction with the principles of explanatory realism⁹. Kim’s reason for believing this appears to be the general one that

‘a preoccupation with the deductive or inferential character of explanation leads to a form of explanatory irrealism ... [it] point[s] away from explanatory realism with the causal relation serving as objective correlate of the explanatory relation’¹⁰

⁹ Kim ‘Mechanism, Purpose and Explanatory Exclusion’, 240-243

¹⁰ Kim, ‘Explanatory Realism, Causal Realism and Explanatory Exclusion’., 236

Why would Kim think this? There are three possibilities which seem reasonable, given Kim's other comments on the issue. First, we may be concerned about the status of the laws relied upon. Under a realist, ontic model of explanation, do we therefore have to defend nomological realism in addition to causal realism? Second, we may be concerned about the notion of an explanation as an *argument*; Salmon and Kim, in particular, have often emphasised the irrealist consequences of such a conception. Third, it would no longer be correct to think that explanatory targets are events. Hempel makes it clear that, according to the D-N analysis, what is explained are not events, but

'kinds or properties of events. ... For the object of description and explanation in every branch of empirical science is always the occurrence of an event of a certain *kind*. ... What is sometimes called the *complete description* of an individual event ... would require a statement of all the properties exhibited by the spatial region of the individual object involved, for the period of time occupied by the event in question. Such a task can never be completely accomplished.'¹¹

However, I am not inclined to regard these three issues as particularly serious. The question of nomological realism is, in the same way as the question of causal realism, outside the scope of this thesis. To the extent that we can assume (as I shall) some form of the latter doctrine, there is no reason to reject, *prima facie*, the former. In chapter six, I will make it clear that I do not regard the inferential structure of certain explanations to be contrary to their causal (ontic) nature. Finally, the argument of chapter five will be that explanatory targets are not (despite, perhaps, first appearances) events.

That does not mean that I regard the analysis of 'complete explanation' by means of 'sufficient explanation' as a fully satisfactory answer. Whilst explanation may be regarded both from a causal-realist and an inferential perspective, the inferential does have a tendency to introduce causally irrelevant explanans. This was demonstrated in chapter two, particularly in the example of John and the birth control pills. In that example the irrelevance was very easy to spot (that's what made it such a clear counter-example to a purely D-N analysis), but this is not always the case. To take a well known

¹¹ Carl G. Hempel, 'The Function of General Laws in History', in *Journal of Philosophy* 39 (1942), 37

example of Hempel's¹², the expansion of the copper rod can be deduced (hence, according to the D-N model, explained) from *both* the heating of the rod, and the stretching of the rod. Yet, these inferential claims do not imply that the causes were equally efficacious.

More seriously, explanations can be complete and yet not sufficient. If the causal mechanism is genuinely stochastic, as it is thought to be in the realm of quantum mechanics, then citing all the causes will not be enough to imply that the effect *had* to happen. On the other hand, an explanation may be sufficient, yet not possess the sort of features we might intuitively believe to be necessary for complete explanations. For example, if the law which guarantees that the explanandum had to happen applies *only* to the case in question (only covers a single case), what does this law really add? Complete explanations would then be trivially easy to achieve.

The notion of a 'complete explanation' will turn out to be relevant to much of the material of this thesis. I will attempt to develop a more satisfactory understanding of the purpose and nature of the concept through further comment at the end of this chapter, in chapter five, and in chapter eight. The interim conclusion is that specifying what a complete explanation consists of is a non-trivial task, and one which seems to lead to a rather different theory of explanation to Kim's ontic approach. However, we cannot simply abandon the notion; as I will argue later in this chapter, it does useful work in the explication of explanatory exclusion.

¹² Hempel, 'Aspects of Scientific Explanation', 418-419

Erotetic Principle of Explanatory Exclusion

Recall that Van Fraassen's erotetic theory analyses explanations as being answers to (explanation seeking) why questions. Such questions ask

'why P rather than Q (or Q₁, Q₂ ...)?'

where Q (Q₁, Q₂ ...) are the 'contrast class'. The presuppositions of such a why-question are that P is true, and that the contrast cases are false. In addition, a range of acceptable answers is chosen, or implied, by the question: those which meet the 'relevance relation'.

The answer, A, should meet three criteria. First, A should be true. Second, A should *discriminate* between P and the contrast class (in a sense best given by the Difference Condition, introduced in chapter three). Third, A should be a relevant reason for P; for Van Fraassen, this means simply that A is one of a class of propositions fulfilling the relevance relation in the correct way with respect to the particular question in hand.

Under the erotetic conception, the provision of an answer which meets the truth requirement, the Difference Condition, and the relevance requirement is sufficient for a perfectly good explanation. Note that there is no role for 'completeness'.

Mark Risjord uses this analysis in order to develop an erotetic principle of explanatory exclusion. He states that two explanations conflict when either

- '1. the presuppositions of one (kind of) why-question are inconsistent or not compossible with the presuppositions of another (kind of) why-question, or
2. the relevance relation for the two (kinds of) why-questions isolate classes of answers that are always or typically inconsistent or not compossible'¹³

¹³ Mark Risjord, 'No Strings Attached: Functional and Intentional Action Explanations' in *Philosophy of Science* 66(3 Supplement) (1999), S302

It is important to note that Risjord's principles are dependent upon features of the question (explanandum), rather than the answer (explanans). This is to be contrasted with KPEE, which focused on features of the respective *explanans*. As we have seen, Van Fraassen's theory holds that a why-question has two presuppositions – concerning the truth of the topic of explanation, and the falsity of the contrasts relied upon by the explanation – and in addition specifies a relevance relation. These features are the basis for Risjord's two principles.

It does no harm to reformulate Risjord's two criteria to make their purpose easier to grasp. The first criterion states that, given that topic and contrast jointly constitute the target of the explanation, where explanatory targets contradict, the explanations are not compatible. (The meaning of 'compossible' in this context is not entirely clear, but most likely refers to natural compatibility of targets, over and above logical compatibility.) This situation was not envisaged at all by KPEE, but does seem reasonable. Where there are contradictory targets (whether or not we believe that those targets are to be understood contrastively), it is not possible for both targets to be true; in that case the respective explanations are also contradictory, therefore exclusive. I take it that the spirit of Risjord's first criterion is undeniable, even if we are not to analyse 'target' in terms of 'why-questions', as Risjord does.

Risjord's second criterion states that where two explanations conflict over the proper relevance relations (i.e. which facts are to be treated as genuinely explanatory), then the explanations are exclusive. The fact that explanations *disagree* over relevance relations is not sufficient for exclusion, since it is a feature of the erotetic theory that different explanations of the same target can quite legitimately have different sorts of relevance relations. To use one of Van Fraassen's own examples¹⁴, question Q may ask

Q: 'why is this conductor warped?'

¹⁴ Van Fraassen, 142

Yet in different pragmatic circumstances (dependent upon what the questioner is particularly interested in, what they already know, and so forth) the relevance relations may differ. For example, Q asked by a supervisor wanting to know which worker to blame may limit relevance to only those answers referring to human error: call this Q₁. However, where one worker asks another, Q may limit answers to those referring to the micro-physical properties of the conductor itself: call this Q₂. But in this case it is not true that questions Q₁ and Q₂ are exclusive; indeed, this seems a paradigmatic case of explanations legitimately focusing on different aspects of the relevant history of the conductor's warping.

Of course, Risjord states that for explanatory exclusion the relevance relations have to be not just different, but incompatible (or non-compossible). But note that Van Fraassen's example does, in one sense, demonstrate incompatibility of relevance relation: for the relevance set of Q₁ includes answer A (human error) and not B (micro-physical properties), whilst the relevance set of Q₂ includes B and not A. What more is required for exclusion? It seems that a question has to not only imply a different relevance relation than does its competitor, but must also *actively disallow the particular answer given by the other*. In our example, Q₁ must include A (human error) and also maintain that B (micro-physical properties) *cannot possibly* be a good answer to a question concerning the same target.

This interpretation of the second criterion appears to be that intended by Risjord in his analysis of Macdonald and Pettit's rejection of functional explanations of social phenomena: 'They think that the explanans of social explanations, when taken as explanatory, conflict with our conception of ourselves as agents'¹⁵. In other words, all answers to questions involving social action which are in functional terms – those which conflict with our conception of ourselves as agents – are excluded.

However, such an interpretation requires that content be given to 'relevance relations' which goes well beyond that envisaged by Van Fraassen's in his use of the term. For Van Fraassen, the relevance relation represents a pragmatic limitation on what could

¹⁵ Risjord, S301

possibly be an answer to *that particular* question, in that particular context. Why-questions themselves lay no restrictions on what counts as a good answer in general¹⁶: the analysis of the latter requires a separate body of theory (as I noted in chapter three). Risjord's use of 'relevance relation' requires that the specification of such relations go beyond pragmatic considerations; for Risjord, to specify a relevance relation is to make commitments which transcend particular question and answer situations. Yet the very essence of a pragmatic commitment is that it does not extend beyond the particular situation in which it is uttered, and that it does not lead to formal relations of inconsistency with other pragmatic commitments.

Risjord's criteria are therefore ill fitted to the explanatory theory which they were intended to serve. On the contrary, Risjord's second criterion, as I have interpreted it by using non-contextual criteria, bears much similarity to the ontically motivated KPEE. In what follows, I intend to build on such similarities to develop a revised set of principles. First, however, I want to put into practice the methodology outlined in chapter one by considering the case study explanations.

Historical exclusion and competition

Intuitively, when are historical explanations regarded as exclusive, and when are they regarded as competing? I will consider four fairly clear cut cases drawn from the pool of eight explanations introduced in chapter two. Two of these are relationships are naturally regarded as competing, and exclusive; two as compatible and non-exclusive.

¹⁶ In particular, see Van Fraassen, 142

There is nothing like a historical orthodoxy to inspire criticism. Lefebvre, as one of the prime representatives of the Marxist approach to the French Revolution, attracted much criticism and his explanation was, more or less explicitly, the foil for many rivals. In particular, Cobban intended his alternative social explanation to be a direct competitor to Lefebvre's. This challenge was developed in three clear ways. Firstly, Cobban criticised Lefebvre's reliance on an overly a priori theory: the Marxist theory of the priority of social class hierarchy as based on control of the means of production, which provides support for Lefebvre's central nomological claim¹⁷. Secondly, he challenged Lefebvre's interpretation of the historical records, with respect to certain key topics such as the composition and behaviour of Lefebvre's bourgeoisie¹⁸. Thirdly, and very much related to the second point, he questioned Lefebvre's understanding of certain core concepts, in particular 'bourgeoisie'.

From a philosophical point of view, we need to ask what general principles guarantee that Cobban's and Lefebvre's explanations are indeed exclusive. First, though, we should allay a potential source of confusion. We must distinguish between (i) direct criticism of an explanation, and (ii) the relation of exclusion or competition between two explanations. The criticisms cited above are an example of (i); strictly speaking, they should not be taken as a case of one *explanation* excluding another. However, neither should we imagine that (i) and (ii) are unrelated. They are related in the sense that elements of direct criticism in a historians' work help us to interpret the implicit claims of their actual (positive) explanation. In chapter two, I argued that recognising the negative and critical features of explanations is vital in interpreting the intended scope and purpose of the positive features of those explanations.

This is clearly the case for an understanding of Cobban's explanation. His negative criticism of Lefebvre's social explanation makes it clear that his own positive social explanation is intended to be in competition with Lefebvre. The direct criticism primarily takes the form of *criticism of Lefebvre's explanans*: the notion of 'bourgeois', and the

¹⁷ Cobban, Chapters I and II

¹⁸ For example, *Ibid.*, 163-167

Marxist law. (I will postpone consideration of the criticism of Lefebvre's theoretical stance itself until chapter six.)

The relationship between Cobban's and Lefebvre's positive explanations is more subtle, but the key point is that it must be interpreted in the light of Cobban's direct criticism. Whilst neither historian would claim that their explanation was 'the whole story' of the origins of the revolution (and therefore, each leaves room for other explanatory accounts), the fact that Cobban prepares the ground for his explanation by criticising Lefebvre's tells us that Cobban does intend to cover the same elements of that story as Lefebvre. In other words, the historians focus on similar parts of the revolution's history, and arrive at different conclusions: that is why they are in competition, and are exclusive.

Schama's explanation is less clearly a direct alternative to Lefebvre's, though I would argue that it is a partial competitor. One of Schama's aims is to provide an alternative explanation to the Marxist account¹⁹, though he doesn't mention Lefebvre by name. The way that Schama's explanation excludes Lefebvre's is similar to Cobban's: core elements of Lefebvre's explanans are challenged, particularly those related to the composition, motivation and role of the bourgeoisie. Because Schama's account is, in some ways, of a different sort to Lefebvre's, the exclusivity is not so clear cut as that existing between Cobban and Lefebvre. In particular, Schama's explanation is more detailed, covers a somewhat different chronological period, and has a different target (partly as a result of the differing chronological focus). We might think that the relationship between Schama and Lefebvre is one of *partial exclusion*.

The most clear cases of *compatible* explanation are of two sorts. Firstly, there are explanations which agree over methodological and theoretical background, and which simply investigate different aspects of the origins of the French Revolution. Furet's and Chartier's explanations seem to be of this sort. Furet assesses the political nature of the revolution, whilst Chartier emphasises the cultural. Crucially, we do not find in Chartier

¹⁹ For example, Schama, xiv, 290

a denial of the importance of other factors. Both writers adopt a broadly ‘interpretative’ approach, and are influenced by similar social theorists (Habermas, for example) and philosophers (such as Foucault; the influence of whom will be investigated in chapter six).

Secondly, there are explanations which do not seem to agree on much at all, but which adopt such different approaches that they may plausibly be seen as complementary rather than exclusive. Three examples stand out in this regard: the explanations of Michelet, Furet, and Skocpol. Whilst Michelet constructs a narrative leading to the French Revolution of the traditional, moralistic, type, Furet adopts a more historiographically aware and critical approach to the revolution’s origins. Skocpol adopts an equally different approach: her explanation is part of an explicitly scientific attempt to understand the origins of social revolutions in general.

What can we learn from this brief examination? One important point is the need to use the historians’ wider (negative and critical) comments to interpret the intended extent of their actual explanation. In particular, this helps the reader to gauge the extent to which their explanation is intended to eclipse other particular explanations of the same event. In addition, we can discern certain features which are relevant, positively or negatively, to historical explanatory exclusion: (i) straightforward conflict between explanans; (ii) the role of differing explanatory types in leading to complementarity between those explanations; and (iii) the role of differing targets.

The fundamental principles

My strategy in the remainder of this chapter will be to provide three principles of exclusion. These are derived from different types of logical inconsistency; I take it that this fact alone is sufficient to support the principles. A consequence is therefore that

explanatory exclusion can be reduced to logical inconsistency. However, I will argue that even principles developed from such a minimal basis can account for the principles expressed by Kim and Risjord.

The following Principles of Explanatory Exclusion are to be regarded as individually sufficient for explanatory exclusion. They are not individually necessary for explanatory exclusion, though they are intended to be jointly exhaustive. The claim that one of the three principles is necessary for any case of explanatory exclusion is supported by the thought that an explanation is constituted (only) by a target, explanans, and a claim of relevance of some kind between the two. Each principle focuses on one of these features.

PEE₁: *The targets of the explanations are inconsistent.* If we accept that targets are constituted by a combination of topic and contrast(s), as in chapter five I argue that we should, then inconsistency can derive from either feature. Should the explanations either *disagree over the truth of the topic*, or *over the falsity of the contrasts*, then they are inconsistent, hence exclusive. (Note that PEE₁ has similar import to Risjord's first criterion.)

PEE₂: *The truths of the explanans are inconsistent.* The most basic source of such inconsistency (exclusivity) is due to straightforward factual disagreement. But given that a set of explanans always implies further presuppositions – in particular, that the concepts used in the explanans are comprehensible and appropriate – disagreement over these presuppositions is sufficient to produce exclusive explanations.

PEE₃: *The explanatory relevance of the explanans are inconsistent.* Under a causal theory of explanation, this should be taken to mean that the explanations disagree over whether the events or facts cited stand in an appropriate causal relationship to the target²⁰.

²⁰ It may be argued that PEE₃ is a special case of the more general PEE₂, since the explanans of an explanation contains both factual and causal claims. PEE₃ should therefore not be listed as a 'fundamental principle'. Even if this is so, PEE₃ has an importance sufficient to warrant emphasising this particular type of exclusion, as we shall see. And in any case, the case is not clear cut: it would not be implausible to

To demonstrate these principles, consider a simplified version of Lefebvre's explanation for the coming of the French Revolution. Suppose that this explanation states that 'the revolution of 1789 took place in France (rather than in some other country) because, previously to that date, the bourgeoisie held economic power but not yet legal/ political power'.

PEE₁ would be met by an explanation which either disputed the truth of the topic – that there was a revolution in France in 1789 – or the falsity of the contrast(s) – that there were no revolutions in other countries in 1789. It is worth repeating that PEE₁ demands incompatibility of explanatory target, not simply difference in target.

PEE₂ would be met by an explanation which disputed the truth of the explanans – that the bourgeoisie held economic power but not legal/ political power before 1789. As suggested above, this could take the form of a direct factual disagreement, or disagreement with presuppositions, such as that there was a 'bourgeoisie' in the sense required by Lefebvre. Note that the explanans must be strictly relevant for PEE₂ to be applicable; not just any factual contradiction in a historical account would suffice. If two accounts differed, for example, in ascribing Louis XVI as 'exceptionally good looking' or 'pig ugly', this would probably not be sufficient to lead to exclusion.

Finally, PEE₃ would be met by an explanation which disputed the relevance of the explanans. In such a case the *fact* of the bourgeoisie's position would be accepted, but not that this fact was explanatorily relevant. Most straightforwardly, if we assume causal explanation, this would amount to denying that the bourgeoisie's position was a relevant cause of the French Revolution in 1789 ('relevant' in a sense most fully given by the Difference Condition, considered in chapter three).

Note that we don't need to actually investigate the truth of the explanations (i.e., the historical record itself) to be able to understand relations of exclusion. It is sufficient to

restrict the explanans proper to factual claims, which are then, by implication, claimed to be linked in the relevant way (usually causally) to the explanandum. I shall remain agnostic over this issue in this thesis.

discover that the explanations disagree in the relevant ways. Of course, to actually know which of two competing explanations is better we do need to know something about the truth of the explanation (i.e., something about the historical record); but this latter task is clearly not a job for the *philosophy* of history.

Application and extension of the principles

Two questions allow us to connect PEE₁₋₃ to the previous considerations of this chapter. First: given that there is no explicit mention of sameness of target in PEE₁₋₃, what should we make of the apparent importance of this restriction seen in regard to both KPEE and the survey of the case study explanations? Second: where do considerations of ‘completeness’ of explanation, introduced in the discussion of KPEE, fit in?

PEE₂ and PEE₃ do not require that the exclusive explanations have the same target. Should they? Kim’s PEE required that exclusive explanations be directed at the same event. In addition, in the previous section we saw that it was natural to take into account the relationship between explanatory targets when considering exclusion (over and above the role of incompatibility of target suggested by PEE₁). On the other hand, Risjord’s principles do not require this; not explicitly, at any rate.

I want to stick with PEE_{2,3} as they stand: explanatory exclusion does not require those explanations to have the same target, or even a similar target. This is the case so long as explanatory exclusion in general is understood as the incompatibility of different explanations. If an explanation of the British victory at Trafalgar contained the statement ‘the French navy decreased during the 1780s’, and an explanation of the French Revolution contained the statement ‘the French navy increased during the 1780s’ then the two explanations are incompatible – hence exclusive – no matter that they have different targets.

Yet explanatory exclusion in a narrower sense – a sense which is likely to be more relevant to actual historiographical consideration – does require sameness (or at least similarity) of targets. To see why, I want to focus attention on PEE₃, and to ask: under what conditions are we to judge that there is a dispute between explanations over the *relevance of explanans*?

This is a potentially tricky problem because explanations themselves (as opposed to direct critiques of those explanations) do not appear to contain (negative) propositions of the form ‘c was not relevant’. Yet, PEE₃ requires incompatibility of relevance claims, an incompatibility which is not satisfied just by replacing ‘c was relevant’ with the alternative claim ‘d was relevant’. This is, in fact, the same problem addressed in relation to assessment of Risjord’s second criterion. I argued that the mere fact that the relevance claims of two explanations *differed* is insufficient to produce explanatory exclusion. Rather, what is required for exclusion is that the acceptance of one relevance claim rules out another. But how are we to know when this is the case?

I think an answer can be formulated using two ideas. The first was introduced in the explication of the exclusive relationship between Cobban and Lefebvre’s explanations. The key to this was the need to rely on the more general critical framework of the histories in question. In the case of Cobban’s explanation, the critique of Lefebvre which preceded it made it clear that the explanation was intended to cover the same ground as Lefebvre’s, and therefore should be exclusive of it. Similarly, Risjord’s example of explanations which exclude due to inconsistent claims of relevance relied on a comprehension of the wider theoretical framework within which the explanations ‘proper’ are situated.

The second idea required for a proper philosophical understanding of PEE₃ is a rehabilitation of the notion of ‘completeness’ (which, as I have suggested, can only be fully achieved with reference to material in chapters five and eight). The purpose of the notion of a ‘complete explanation’ is made clear by PEE₃. PEE₃ requires that the relevance of a competitor’s explanans is challenged; often, this is achieved through

implicit denial. And ‘completeness’ gives us the resources to understand how this is done. The key point is that where one explanation is complete, then there is simply no room for another explanation. By providing a complete explanation, the existence of a rival explanation is precluded: both cannot be true, hence they are incompatible. This is subject to the two provisos noted by Kim: that the explanations are not in a relationship of dependence, and that *the explanations are directed towards the same target*.

In the more general model represented by PEE₁₋₃, KPEE is therefore a special case. It is a way to understand the implicit denial of a competitor’s claims to explanatory relevance. And in this special case, there is a need for sameness of explanatory target (or at least similarity, as I will shortly argue). A remaining concern is why, given that KPEE is only a special case of explanatory exclusion in general, it appears to occupy such a paradigmatic role. I would conjecture that the reason is that the other principles (PEE₁₋₂) are essentially straightforward criteria of propositional incompatibility. Only PEE₃ (with Kim’s principle as the most useful means to give substance to this) is uniquely directed at *explanatory* exclusion. In addition, I would suggest that, although explanations with quite different targets may be regarded as exclusive, only those explanations with similar targets are generally regarded as truly in *competition*. (Of course, this claim has only an intuitive basis; however, if the reader disagrees, they will be relieved to know that little of the subsequent material depends upon it.)

Whilst I am convinced that PEE₁₋₃ collectively exhaust the criteria for strict exclusion, I am not convinced that they are sufficient to do justice to our everyday, fuzzy, judgements of exclusion. For this reason, we need to find room for a notion of ‘partial exclusion’. I aim to demonstrate, however, that cases of partial exclusion can be accounted for on the basis of *modified* understandings of PEE₁₋₃.

PEE₁₋₃ can mislead in two, opposite, ways. Firstly, the principles can suggest that explanations are exclusive where we would naturally regard those explanations as being, for the most part, compatible. Secondly, the principles can suggest that explanations are

not exclusive where we would naturally regard those explanations as being so, again for the most part.

It is easy to see how PEE_{1-3} could be satisfied, and yet explanations remain *fundamentally* consistent. Firstly (PEE_1), the explanatory targets may be inconsistent to only a minor degree. In particular, there may be two contrastive explanations which share the same target ('why did the French Revolution come about?') and have only minimally differing contrast classes ('rather than the British/ German/ Russian' / 'rather than the British/ German/ Austrian'). Secondly (PEE_2), the explanans may be inconsistent to only a minor degree. Thirdly (PEE_3), the relevance of the explanans may likewise be inconsistent to only a minor degree. The meaning of 'minor degree' in this context is best explicated not by simply counting causes which differ, but by attending to the relative importance of these causes. The task of explicating 'explanatory importance' is undertaken in Part III.

On the other hand, there are cases where PEE_{1-3} are not satisfied, yet where we would wish to regard explanations as partially exclusive. One way for this to come about is where it is likely, though not certain, that the truths of explanans are inconsistent. For example, it may be that it is likely that the explanans of one explanation would have causes or effects which were in contradiction with the explanans of a second. 'The fact that the aristocracy owned the majority of the factories in pre-Revolutionary France' most likely contradicts with 'the bourgeoisie had the economic power in pre-Revolutionary France', though the two do not logically contradict (here, perhaps, is room for Risjord's notion of 'compossibility'). It should be noted that this condition of 'partial exclusion' is based on *epistemological* features, rather than on logical features which are the core concern of PEE_{1-3} . We may judge explanations to be 'partially exclusive' because we think it *likely*, but not certain, that explanations are incompatible.

Focusing on KPEE (which we saw was a special case of PEE_3), there are two interesting ways for this to be partially, though not strictly, satisfied. The key requirements were for *complete* explanations, leading to the *same* target. I suggest that partial exclusion may be achieved either where there are two explanations which are *largely complete*; or where complete explanations lead to a *similar* target (as well as, conceivably, some combination

of these two elements). The meaning of ‘degrees of explanatory completeness’ will be investigated parallel to the explication of *completeness* itself, particularly in chapters five and eight. ‘Similarity of target’ can be examined using the resources of a contrastive theory of explanation, as is done in chapter five.

The possibility of nearly complete explanations perhaps helps to make sense of Thagard’s claim that ‘we should assume that hypotheses that explain the same evidence compete with each other unless there is reason to believe otherwise’²¹. Whilst most explanations are not entirely complete, it is perhaps plausible to suppose that the majority of decent explanations at least approach completeness, to the extent that another ‘nearly complete’ explanation is at least unlikely to be consistent. In addition, the necessity of adopting a more vague understanding of KPEE was demonstrated with regard to potential counter-examples to KPEE, particularly cases of ‘instantiation’, in the first section of this chapter.

To summarise the claims made thus far: explanatory exclusion is always, and only, a result of explanatory incompatibility. Explanations are exclusive either where the *topics* of those explanations contradict, or where the *truth of the explanans* contradict, or where the *explanatory relevance of the explanans* contradict. It is perhaps the latter possibility which is of most interest in practice. Conflicting relevance claims may be more or less implicit. One important way to understand implicit conflict over explanatory relevance is to make use of the notion that complete explanations exclude other (independent, not over-determined) explanations of the same target.

Where one of these conditions is met, it is reasonable to suppose that such exclusivity can be of more or less consequence – from a fundamental conflict to one of little significance. Conversely, it may be the case that none of these conditions is met with certainty, yet that the content of the explanations lead one to think that compatibility is unlikely. Two possible instances of the latter phenomenon are the competing claims of two near-

²¹ Thagard: not included in the published version, but in an earlier version of the paper on his web-site: <http://cogsci.uwaterloo.ca/Articles/Pages/%7FTraditions.html>

complete explanations; and the competing claims of independent, complete explanations with only slightly differing targets.

What theory of explanation do the principles presuppose?

What model of explanation, or elements thereof, are presupposed by my three principles of explanatory exclusion? Risjord states that ‘clearly no advance [in understanding explanatory exclusion] could be made without adopting some model of explanation’,²². To some extent, of course, this is true; though I have endeavoured to construct (and phrase) the principles in such a way as to leave them with as wide an appeal as possible. For the principles to be applicable, all that is required is that explanations may be incompatible. If one agrees that explanations consist of a target, explanans, and claims of explanatory relevance, then PEE₁₋₃ follow naturally.

Even such weak requirements provide PEE₁₋₃ with appropriate bite. Compare Raymond Martin’s purely epistemological criterion of ‘competing explanation’ (Martin’s term ‘competing’ should be read as ‘exclusive’, in the sense of this thesis):

‘two explanations of an event are competing explanations of that event *if* it is more likely, on the available evidence, that one or the other explains that event than that both explain it’²³

This doesn’t appear to get to the root of the matter: how do we judge whether it is more likely that one or both explanations explain it, and how is it possible to be correct or incorrect about this? To the extent that there aren’t definitively correct answers, Martin’s criterion may be regarded as an attempted explication of an *irrealist* principle of

²² Risjord, S300

²³ Martin, *The Past Within Us*, 51-52

explanatory exclusion. To what extent, though, is it possible to provide substantive principles of explanatory exclusion from the basis of an irrealist theory of explanation?

The explanatory irrealist, broadly speaking, is one whose criteria for a ‘good explanation’ refers solely to provision of epistemological links, not ontological truth. Kim, in discussion of this issue, cites the causal irrealists. For Hume, Russell and Wittgenstein, amongst others, the causal relation is derived from the explanatory relation. The direction of this derivation leads to a rejection of explanatory realism, since the required notion of a metaphysical relation which the explanatory relation should mirror is conspicuously absent. We might add to these the position of Hayden White. As we saw in chapter three, White regards historical explanation as consisting of the arranging and privileging of elements of the historical record in a manner which cannot be constrained by the historical past (explanations are ‘invented, not found’). Finally, Van Fraassen’s erotetic theory of explanation is irrealist, since the ‘relevance relation’ is defined purely pragmatically.

It is natural to suppose that the explanatory irrealist should not feel constricted by explanatory exclusion. Indeed, it is interesting that this is indeed the effect of White’s treatment of different historical explanations. White regards the differing narrative structures imposed by historians in order to explain their targets indicative only of the fact that they have approached the historical record in different ways. For White, there is no question that one narrative excludes any other.

However, Kim does consider it possible that a plausible account of explanatory exclusion could be based on a position of explanatory irrealism, even if it is more natural to ally his principle with explanatory realism. He argues on epistemological grounds for what might be described as a *cognitive* or epistemological principle of explanatory exclusion:

‘There is a corresponding *epistemological* exclusion principle: *No one may accept both explanations unless one has an appropriate account of how they are related to each other. ...* Even if we abandon the idea that there are objectively explanatory relations in the world, we may still find something cognitively unsettling and dissonant about having to face, or accept, two or more independent explanations of the same phenomenon. The explanatory premises of one

explanation need not logically contradict those of another, and there may be sufficient evidential warrant for thinking each set to be true. However, accepting the two sets of premises as constituting explanations of the same event (or any one thing), each complete in itself and independent of the other, may induce a sort of incoherence into our belief system.’²⁴

As Kim makes clear in the passage which precedes this quote²⁵, such an idea depends to a great extent on the plausibility of ‘the maxim of explanatory simplification’: that our explanatory scheme is preferable to the extent that it relies on fewer explanatory premises. This clearly relates to the idea of explanation by unification, introduced in chapter three, as one type of cognitive model of explanation. Without going into the detail of such an account, and without accepting that reduction in the number of independent explanatory premises *always* leads to better explanations, we can, I think, accept that to unify is generally an explanatory virtue.

However, whilst Kim’s general motivation is clear, the crucial demand is that we have an *appropriate* account of how the explanations are related. Kim attempts to substantiate this demand by reference to the terms ‘complete’ and ‘independent’. Yet, as we have seen, these terms have only been explicated with reference to the ontological structure underlying explanations. These terms, and the resulting principle, therefore place the explanatory irrealist under no compulsion to identify cases of explanatory exclusion.

One advantage of explicating explanatory exclusion in the very general terms of PPE₁₋₃ is that those principles point to constraints which even the explanatory irrealist should recognise. What is ‘irreal’ about these theories are the claims of relevance which explanations purport to make. Whilst the realist regards relevance claims to be properly based on ontic (primarily, causal) structure, the irrealist makes no such demand. PPE₃ will, therefore, have no meaning for the explanatory irrealist, simply because there is no way for relevance claims to stand in logical contradiction. Since KPEE is a special case of PPE₃, it is no surprise that this principle requires a thesis of explanatory realism. However, PPE₁₋₂ should present the explanatory irrealist with no metaphysical qualms.

²⁴ Kim, ‘Mechanism, Purpose and Explanatory Exclusion’, 257-258

²⁵ *Ibid.*, 254-257

At the end of chapter two I summarised the findings of the case studies by drawing attention to the different aspects of a historical explanation. In particular, these included the explanatory *explanans*, *target*, *negative or critical claims*, *theory*, *concepts used*, and *judgements of importance*. I have already outlined how the explanans, target, and negative/ critical claims are relevant. But the fundamental principles need to be extended in order to properly take account of each of the explanatory elements.

The explanatory target is important for three reasons. Firstly, PEE₁ referred directly to incompatibility of targets as a criterion of explanatory exclusion. Secondly, I have suggested that we are usually most interested in explanatory exclusion where the explanations are directed at the same (or similar) target: only in such cases do we regard the explanations as genuinely *competing*. Thirdly, to apply and extend PEE₃ in the direction indicated by KPEE, we require explanations to have the same (or similar) target. In chapter five I will focus on a contrastive notion of explanatory target in order to challenge Kim's version of explanatory realism, and then to show how this analysis can aid comprehension of explanatory exclusion.

Consideration of explanatory theories, and of concepts used in explanations, raise the issue of incommensurability. A concern is that PEE₁₋₃ appear not to cover the possibility that certain explanations are incommensurable. However, for example, the consideration of Michelet's, Furet's and Skocpol's explanations above may lead us to regard these as incommensurable (incomparable). In chapter six I will examine the possibility of conceptual and methodological incommensurability, and, in general, the effect that the adoption of different types of explanation has on explanatory competition.

Finally, 'explanatory competition' seems to be, in some sense, a broader notion than 'explanatory exclusion': there can be competing explanations which do not exclude. Such explanations can be better or worse, in the relative fulfilment of explanatory virtues, without thereby excluding one another. This possibility will be the topic of chapter seven.

Chapter Five: Target and Contrast

In chapter four, I derived three fundamental Principles of Explanatory Exclusion from the presupposition that explanatory exclusion is derived from explanatory incompatibility. Given that the most basic form of an explanation is of a target, one or more explanans, and claims of relevance connecting explanans and target, it is natural to formulate three exclusion principles. Conflict over relevance claims is perhaps of most interest, and Kim's ontic PEE (KPEE) illuminates an important way for this to occur. There can be no more than one complete and independent explanation of the same target, since there can be no more than one complete and independent set of explanatorily relevant facts pertaining to the same target. Justifying KPEE in this way relies on an assumption of explanatory realism.

In this chapter, I will argue that contrastive explanation should be adopted as the means to analyse the nature of explanatory targets. A contrastive understanding provides certain unique advantages for the comprehension of explanatory practice in general. In particular, I will demonstrate how contrastive explanation can aid understanding of both explanatory exclusion, and the notion of a 'complete' explanation, introduced in chapter four.

KPEE and explanatory target

KPEE, and the more general PEE₃ (though not PEE₁₋₂) relies on the prior acceptance of explanatory realism. Kim describes the thesis of explanatory realism thus: C is an explanans for E *in virtue of* the fact that c bears to e some determinate objective

relation R. Let us call R, whatever it is, an “explanatory relation.”¹ Crucially, Kim makes it clear that ‘c’ and ‘e’ stand for events. Kim therefore claims that

if and only if - C names event c, and

- E names event e, and

- c stands in an appropriate relation (e.g. causal) with e

Then C explains E.

The problem is that these conditions do seem necessary for explanation, but not sufficient. Not any way of *naming* appropriately related events leads to explanation. Using the conceptualisations introduced in chapter three, the problem is that Kim has identified the ontic requirements for explanation, but has disregarded the necessary cognitive requirements. The following, well known, example illustrates this point:

(1) To cite the hurricane as an explanation of the catastrophe is appropriate. Yet to claim that the event reported on page 5 of Tuesday’s Times explains the event reported on page 13 of Wednesday’s Tribune would be incorrect. This even though the hurricane is the event reported on page 5 of Tuesday’s Times, and the catastrophe is the event reported on page 13 of Wednesday’s Tribune².

The problem stems from the fact that explanation depends on how the relata are described. In other words, the explanatory relation is intensional; unlike, most would claim, the causal relation. This fact leads to a difficulty not only with Kim’s definition of explanatory realism, but his principle of explanatory exclusion.³

(2) Suppose I am at a football game and feel hungry; I eat a prawn sandwich. Then the event ‘my hunger being satisfied’ and the event ‘my eating a prawn sandwich’ are one and the same. ‘I ate a prawn sandwich because I was in the mood for seafood’ is one explanation; ‘my hunger was satisfied because I consumed some food’ quite

¹ Kim, ‘Explanatory Realism, Causal Realism and Explanatory Exclusion’, 226

² The example originates from Davidson: Donald Davidson, ‘Actions, Reasons and Causes’ (originally 1963) in Donald Davidson, *Essays on Actions and Events* (Oxford: Clarendon Press, 1980), 17

³ Marras also criticises KPEE on the basis of its dependence on an unsatisfactory statement of explanatory realism: Ausonio Marras, ‘Kim’s Principle of Explanatory Exclusion’ in *Australasian Journal of Philosophy* 76(3) (1998): 439-451

another. Note that ‘my hunger was satisfied because I was in the mood for seafood’ is not a good explanation’.

Example (2) shows that two explanations of the same event may properly focus on different aspects of that event. As it stands, this provides a counter-example to KPEE. For, as (2) demonstrates, we may have two explanations which are complete, independent, focused on the same event, and yet not exclusive. A similar counter-example was suggested in the previous chapter, resulting from explanations bearing a relation of ‘instantiation’. In such cases also, two explanations can be complete, independent, directed at the same *event*, and yet not be exclusive.

There are two ways to respond to these sorts of criticism. We might, firstly, object to the tacit notion of either ‘explanation’ or ‘event’ at work in the problematic examples. It could be suggested that the ‘bad’ explanations, such as ‘the event reported on page 5 of Tuesday’s Times explains the event reported on page 13 of Wednesday’s Tribune’ are not strictly incorrect, just odd. In particular, Davidson makes it clear with reference to this example that his supposition is that explanations presuppose laws. There must be some kind of plausible regularity from the type of event mentioned in the explanans to the type of event mentioned in the explanandum. There might be some sort of regularity linking hurricanes and catastrophes, but not events on page 5 with those on page 13. Why can’t this presumption simply be denied?

Perhaps the supposition that all explanations imply a nomological relation *can* be denied (a question I consider in chapter six). However, the more general requirement that explanans and target always be cited in *appropriate* ways can not. The necessity of regarding explanation as a partially epistemic notion will be the subject of chapter six. For now, let us simply note that Kim himself does not agree with the denial:

‘Just as knowledge requires more than truth, explanations presumably must meet further requirements (‘internal’ conditions – perhaps logical and epistemic ones), although exactly what these are does not concern us here.’⁴

⁴ Kim, ‘Explanatory Realism, Causal Realism and Explanatory Exclusion’, 226

Kim's preferred response would be to challenge the notion of 'event' in the problematic examples. To the extent that the criterion of event identity is 'narrow', the events mentioned (my hunger being satisfied and my eating of a prawn sandwich, for instance) may not, after all, be identical. Kim's theory of events leads to particularly fine grained identity conditions, such that events which are intuitively identical turn out not to be according to Kim.

The essence of Kim's analysis is that an event consists of 'a concrete object exemplifying a property at a time'⁵. Adopting such a view would lead to a rejection of the event identity claims made in examples (1) and (2), since different properties are picked out in the different cases. But if the events aren't the same, then the explanatory target differs, meaning that KPEE is not intended to apply. Whilst Kim's 'property exemplification' analysis of events therefore allows KPEE to avoid the putative counter-examples, the analysis has been criticised on independent grounds. Most obviously, it leads to judgements of event identity rather drastically at odds with common intuition: Brutus' killing of Caesar turns out to be a different event to Brutus' stabbing of Caesar, for example⁶.

Whether or not Kim's analysis of events is tenable, we *can* surmise that KPEE does not function if we are to regard explanatory targets as events in the more intuitive sense of the term (if, for example, Brutus' killing of Caesar is intuitively thought to be the same event as Brutus' stabbing of Caesar). Indeed, Kim's analysis seems closer to another, more popular, analysis of explanatory target: in terms of *properties* of events. (One difference is that, for Kim, though an event is constituted by a single property, it may exemplify other properties: 'Brutus' stabbing of Caesar' may have the property of being 'performed outside'. However, Katz suggests that Kim's account suffers difficulties with such 'adverbial modification' of events.)

⁵ Jaegwon Kim, 'Causation, Nomic Subsumption and the Concept of an Event' (originally 1973) in Jaegwon Kim, *Supervenience and Mind: Selected Philosophical Essays*, ed. Ernest Sosa (Cambridge: Cambridge University Press, 1993), 8

⁶ See Richard Swinburne, 'Are Mental Events Identical with Brain Events' in *American Philosophical Quarterly* 19 (1982), 176. See also Bernard Katz, 'Kim on Events' in *Philosophical Review* 87 (1978): 427-441

Regarding explanations as directed at properties of events, rather than events as such, is Marras' preferred approach⁷. As examples (1) and (2) suggest, it often seems more accurate to qualify the analysis of explanatory targets in this way. There are, of course, questions which can be raised concerning an analysis of explanatory targets in terms of properties. How should properties be individuated? How are we to accommodate the fact that the property 'deep red' is a part of 'red'? And how are we to understand the intuition that 'ambling' overlaps with 'strolling'? These are serious issues: a conception of what explanatory targets are should be able to make sense of the identity of those targets. However, I will not consider these issues in relation to *properties*, for the reason that the property analysis of explanatory target still does not get things quite right, whether or not the questions concerning individuation and the like can be resolved.

The *spirit* of KPEE is not endangered by the above criticisms. For, given a presupposition of explanatory realism, some form of KPEE is justified⁸. A modified version alters the categorisation of explanatory target from event, which affects the scope of KPEE to the extent that the preferred relata differ from events.

Why contrasts?

I suggest that we analyse explanatory targets in contrastive terms. I introduced a brief history of the theory behind *contrastive explanation* in chapter three. There I made it clear that I regard the fundamental claim of that theory to be one regarding the nature of explanatory targets: contrastive explanation regards the target as being of the form 'p rather than q'. I shall add detail to this analysis of contrastive targets in this chapter.

⁷ Marras, 448-450

⁸ Kim recognises this: 'if it is aspects of events, rather than events *simpliciter*, that are explained, then explanatory exclusion would apply to these event aspects' (Kim, 'Mechanism, Purpose and Explanatory Exclusion', 257)

There are four main reasons why contrastive explanation is ideally suited for analysis of the targets of historical explanation. Firstly, it enables a satisfactory understanding of explanatory exclusion, particularly KPEE, including the notion of partial exclusion. Secondly, it enables us to make sense of ‘complete explanation’ as it appears in Kim’s principle. These two advantages will be explored in the penultimate section of this chapter. In addition, contrastive explanation has the flexibility to apply to a variety of explanatory targets. Finally, it can be seen to fit well with the historical case studies, through its ability to illuminate salient aspects of explanatory practice (over and above the main focus of this chapter, explanatory exclusion).

Contrastive explanation provides a necessary flexibility in categorising explanatory targets. Whilst explanatory targets are often more accurately categorised as properties, or aspects, of events rather than events per se, this does not *always* seem to be the case. Explanations are sometimes directed at properties of events; but sometimes simply at events, and sometimes at absences.

As Woodward has argued, sometimes we do want to explain just the occurrence of an event, and not one of its properties⁹. Using Woodward’s example, we can explain not just ‘that the short-circuit caused the fire to be purple’ but ‘that the short-circuit caused the purple fire’. A defender of a thorough going property analysis might reply that we should regard ‘existence’ as a property; one which is necessary for the presence of other properties of that event. But this view is beset by notorious problems, raised by, for example, Kant, Frege and Russell. We can also explain absences: ‘why was there no fire?’ ‘Because the circuitry did, after all, remain intact’. What could be the *property* referred to in such explanations?

A further set of examples provides difficulties for an analysis of explanatory target based purely on event properties. In conjunction with an examination of the case study explanations, I will argue that mention of event properties does not always provide the necessary detail in characterisation of explanatory targets. Whilst properties of events generally lead to more fine-grained explanatory targets than

⁹ James Woodward, ‘A Theory of Singular Causal Explanation’ (originally 1984) in *Explanation*, ed. Ruben 262-265

events (which is why the examples (1) and (2) in the previous section are problematic for an analysis based on the latter but not the former concept), there are occasions where even the relevant properties are insufficiently fine-grained to fit the intuitively applicable target.

The following brief study assesses the suitability of contrastive explanation as an analysis of the explanatory targets of the case study explanations, prior to a more theoretical investigation of the above claims made in support of contrastive explanation. I will focus on three ways that historians' contrasts shape and affect their explanation. As elsewhere in this thesis, I attempt to use this descriptive study not only to exemplify aspects of my chosen theory of explanation (in this case contrastive explanation), but also to shape understanding of that theory itself.

The most obvious dimension of explanatory target which depends upon the historian's choice of contrast is which *aspects of an event's history* are relevant. Recall that the Difference Condition for answers to contrastive questions, introduced in chapter three, states that the only relevant causal history is that which has no analogue in the history of the contrasts. The choice of contrasts therefore 'screens off' as irrelevant elements of the causal history. If the choice of contrast or contrasts is a good one, the remaining elements are manageably few. The examples of Lefebvre, Hegel and Skocpol demonstrate the difference in foci which derive from differences of contrast.

Lefebvre uses different kinds of revolution as contrasts: the English revolution (of 1688), and the nineteenth century continental revolutions¹⁰. In this way, two aspects of the French Revolution are picked out for explanation by Lefebvre. By comparison with England, the salient difference is the violence of the revolution; by comparison with the continental revolutions the salient difference is that the French Revolution was a revolution 'from below'. Corresponding to these salient differences, the contrasts select certain parts of the (causal) history of the French Revolution as relevant; those parts which do not correspond to anything in the histories of the English or other continental revolutions.

¹⁰ Lefebvre, 2

At one point, Hegel uses as a contrast a revolution which never happened: an imagined German revolution at a similar time to the French¹¹. The salient difference picked out by Hegel's contrast is, of course, that of nationality. Hegel therefore must locate a cause of the French Revolution which has no corresponding factor in German history; quite plausibly, he focuses on the differing Ecclesiastical relationships in each country.

Skocpol uses contrastive cases to shape her explanatory target more explicitly. The features of the French Revolution to be explained are those which differ from the contrastive cases of the 'failed revolutions' of Prussia, Japan and England¹². In accordance with this choice, the relevant aspects of the French Revolution and its history are, for Skocpol, primarily concerned with attempted reformations carried out by the French state. (Notice that to the extent that Skocpol's overall explanatory target is a general fact, rather than a particular, contrastive explanation is not restricted to explanation of particulars.)

However, we should note that Skocpol's explanandum is also shaped by *comparison* with other cases, in accordance with Mill's Method of Agreement. The Method of Agreement states that where two circumstances produce the same outcome, the cause is a common factor. Skocpol therefore compares the French Revolution to other revolutions which are of the same type as the French: those of China and Russia. This strategy – which we might call *convergent explanation* – is easier to take account of than is the use of *contrastive* explanation. We should say that to explain P *compared* to Q is to provide a cause of P which is also a cause of Q. Convergent explanation mirrors contrastive explanation in many ways. For example, just as a rule of thumb guide to selection of contrast is that the contrast be as similar as possible to the fact to be explained, so the facts *compared* should be as different as possible (whilst exhibiting the same outcome), so as to minimise similarities and hence 'relevant' causes. We might wonder whether such use of convergent explanation counts against my overall recommendation of contrastive explanation: I will return to this in the final section of this chapter.

¹¹ Hegel, *The Philosophy of History*, 443

¹² Skocpol, 37

A second use of contrast in relation to targets of explanation is in accounting for the *detail of explanatory targets*. (This aspect of contrastive explanation was first pointed out by Garfinkel¹³, but so far as I am aware, has received no attention since.)

Garfinkel introduced the example: ‘why did the car skid?’. What is the implicit contrast to this question? To explicitly complete the question by adding ‘rather than 5cm further on’ would usually be inappropriate, since it would require an answer too detailed for most purposes. The answer to ‘why did the car skid (there), rather than 5cm further on?’ would need to refer to aspects of the history of the skid which related only to its occurrence in that precise place. These are not aspects which would normally be cited in answer to ‘why did the car skid?’. This is because we have an alternative, if approximate, idea of what the contrast should be: perhaps ‘rather than not at all in that journey’, or ‘rather than around the next bend’.

Lefebvre’s account provides an example of this phenomenon. Lefebvre states his intention to explain the ‘distinctive features’ of the French Revolution¹⁴. Yet he clearly does not intend to explain why the Estates General first met on the 5 May rather than the 4 May; nor why the violence of the revolution resulted in precisely the number of executions that it in fact did. Each of these explanatory targets rely on contrasts which are (pragmatically) *too close* to the fact to be explained, and which therefore call for causal history which is inappropriately detailed. The permissible contrasts are kept at a pragmatically appropriate ‘distance’ from the event to be explained. Explanations do of course vary in the detail which they require. The varying degree of detail required can be regarded as resulting from the (usually implicit) selection of contrasts which are closer to or further from the target fact.

Sometimes, a detailed explanation – one with a more specific target – is better. But not always: for an explanation with an overly specific target may miss broader patterns. An explanation such as Lefebvre’s, or Hegel’s, aims to explain why an event *of a certain broad type* came about: a social change, or a change in conception

¹³ Garfinkel, 30

¹⁴ Lefebvre, 2

of freedom. If we agree that the explanation of more general targets may be fruitful in this case, then that choice of level of generality may be preferable.

Thirdly, our understanding of contrastive explanation can make sense of *explanatory refusal* (suggested again by Garfinkel¹⁵, also by Van Fraassen¹⁶). Where there is no substantive (non-trivial) difference between the fact and the contrasts of question Q, then there is no good answer to Q. In this case, rather than attempting to answer Q, the question should be dismissed. How, for example, would we answer the question ‘why are there *nine* planets (rather than some other number)?’¹⁷ We should dismiss it: there is no non-trivial difference between the history of there being the number of planets which there actually are, and the possible situations in which there were different numbers. We might call this sort of explanatory refusal ‘refusal due to *applicability* of target’.

Pertinent applications of the idea of explanatory refusal can be found in the historical works. *Factual* explanatory refusal can be seen where the truth of the explanandum is challenged. For example, Schama’s reply to Cobban’s question ‘How was the bourgeois revolution related to the role of the peasantry?’¹⁸ would be to challenge the presupposition that there was a ‘bourgeois revolution’. Such cases point to exclusion of the type PEE₁, as outlined in chapter four.

Refusal due to *applicability* presents a further class of explanatory refusals. Skocpol, in viewing the revolution as essentially political, finds the key question to understanding the origins of the French Revolution to be ‘why did France devolve from European dominance to bankruptcy?’¹⁹ In contrast, Lefebvre, in accordance with his emphasis on the social aspect of revolutionary origins, would regard the above question as, if not irrelevant, at least peripheral. The disagreement depends on background theory, just as did the disagreement over the relevance of asking ‘why are

¹⁵ Garfinkel, 4-10

¹⁶ Van Fraassen, 111-112

¹⁷ Adapted from Garfinkel, 6

¹⁸ Cobban, 24

¹⁹ Skocpol, 52

there *nine* planets?’ According to Lefebvre’s understanding of the origins of the French Revolution, there is less of a coherent story to be told regarding the difference between the French state being bankrupt and being solvent than there is between the bourgeoisie being in a politically dominant or a politically subservient position.

Schama’s question ‘why was the revolution *violent*?’ provides a second example; the question is of prime importance to Schama, but less important to most others. Again, the opinions differ on the basis of background theories, implying that a given contrastive difference is trivial or fundamental, accidental or inevitable. Explanatory refusal due to applicability of target is an extreme case on the scale from explanatory relevance to irrelevance, and is well accommodated once the intended contrasts and concomitant differences are recognised.

Contrastive focusing

We are now in a position to develop a fuller understanding of contrastive explanation, initially by investigating how contrastive explanation relates to non-contrastive formulations. My central idea will be that contrast is an extremely useful *tool* in understanding any explanatory practice, particularly in relation to the specification of explananda. By ‘tool’ I mean to exclude the idea that ‘contrast’ is in any sense an irreducible, ontological feature of the world; as compared to, for example, events or properties.

In order to clarify this idea, I will first compare *contrast* to *explanatory presupposition*: are they equivalent notions?²⁰. There is, after all, a clear relation between that which is presupposed by contrastive answers, and explicit presupposition. To explain: in chapter three it was recognised that contrastive

²⁰ Hitchcock, for one, believes that they are (Hitchcock, 608)

questions presuppose that the fact to be explained is true, whilst the contrasts are false. In addition, we have seen that according to DC, the similarities between fact and contrast(s) are not relevant to a good answer. Therefore, these similarities are *presumed* by the explanation. Only differences are relevant. It is as a result of the latter presupposition that contrastive explanation achieves its primary function of focusing explanatory attention on a small amount of the enormous source of explanatorily relevant material. The first set of examples drawn from the case study in the previous section demonstrated this in practice. So long as the contrasts are reasonably similar to the fact, the vast majority of explanatory features are thereby excluded.

A contrastive target therefore presupposes that only certain facts are relevant. This suggests the possibility of being able to restrict relevance non-contrastively: ‘directly’. The request ‘why P rather than Q?’ is equivalent to ‘why P, with respect to facts {x}?’; *so long as facts {x} are those which are in fact differences between P and Q*. The same question can, therefore, be asked in different ways, contrastively and non-contrastively. Why, then, should we bother with the contrastive formulation?

It is important to recognise that though the contrastive and non-contrastive formulation may be empirically equivalent, they are not formally equivalent: for their equivalence is dependent on the empirical proposition ‘facts {x} are those which are differences between P and Q’. In a similar way, ‘the sky is blue’ is equivalent to ‘the grass is green’. They have the same truth value, but this equivalence is dependent on contingent (non-logical) factors.

The most important reason why we might prefer the contrastive formulation is precisely because it allows ignorance of such contingent or empirical facts. Asking the contrastive question ‘why was there a revolution in France, rather than Germany?’ *may* be to ask the same question as the non-contrastive ‘focusing on church-state relations, why was there a revolution in France?’. But the latter presumes more knowledge than the former. The contrastive formulation only requires the knowledge that the two cases are reasonably similar²¹.

²¹ Lipton recognises this as an advantage of contrastive explanation: Peter Lipton, ‘Making a Difference’ in *Philosophica* 51(1) (1993), 50

Next, compare contrastive explanation to the analysis of explanatory targets in terms of properties. I suggested that not only are some explanations seemingly *not* directed towards properties of events, but in some cases the analysis of targets in terms of properties is insufficiently detailed.

Marras suggested that explanatory relata are *properties*; I have suggested that they are *contrastive facts*. Aren't these ideas equivalent? In particular, in framing contrasts between events, do we thereby pick out properties, such that the difference between two cases results in the properties in question? Not quite. The question 'why did the French Revolution come about?' implies, from its surface form alone, that the explanandum is an *event*. In asking 'why was there a revolution in France?', a *property* of the event the French Revolution is picked out for explanation: that it is French. The contrastive question 'why was there a revolution in France rather than Germany?' further restricts the answer. It does this according to the presupposition of difference, in that the question admits only of answers which focus on differences between the French and German case.

In this way the relevant *history* of an event may be focused, in addition to the relevant *properties* of that event. The difference between the non-contrastive and contrastive formulations can be seen in the fact that more history is relevant to the former question than the latter. For example, the fact that France is a continental country, surrounded by geo-political rivals, may be relevant to the question 'why was there a revolution in France?' But this information is not relevant to 'why was there a revolution in France, rather than Germany?', because the geographical factor is also in the history of the contrast.

The upshot is that contrastive explanation is a device for focusing explanatory information, over and above the focus on *aspects* of events. This focusing *can* be done using straightforward command or exclusion; however, the use of contrast carries the unique advantage of being tolerant of ignorance, fitting well with common usage, and being flexible enough to encompass a range of types of explanatory target.

Contrastive explanation suggests that explananda are constituted according to contrastive focusing of events and their histories. In short, explananda are individuated according to what a good explanation of those explananda should *exclude*. Any given historical question is shaped by a (more or less implicit) range of contrast cases; ‘shaped’ according to the requirement that a certain answer compare the fact and foil(s) in a manner given by DC.

I believe that to understand properly the way that explanatory targets are constituted by what they exclude, we need to make use of Garfinkel’s term *contrast space*. Whilst the standard examples used to introduce contrastive explanation in chapter three – Jones/ Smith, and Monash/ Oxford – made use of determinate, and indeed a solitary, contrast, that is not usually the case. Where we apply the notion of contrastive explanation beyond such clear cut cases as these, we find that contrasts must be seen as at least *partially indeterminate* for the idea of contrastive explanation to make any sense. It is for this reason that I find ‘contrast space’ appropriate, as the term implies that contrast cases need not be precisely specified. Instead, contrasts are better represented by regions of possibility than by discrete cases.

To be more precise: contrast spaces are collections of possibilities – those possibilities which an adequate answer to the contrastive question must exclude, by pointing to a relevant difference between the histories of the fact and the possibilities constituting the contrast space.

We could use the now familiar modal notion of ‘possible worlds’ to make ‘contrast space’ more precise, but it is important to note that contrast spaces are not simply collections of possible worlds. Firstly, contrast spaces may vary by including differing numbers of possibilities from the same possible world. For example, contrast space A may include a possible eighteenth century English Revolution, whilst contrast space B may include both this and a Russian Revolution of the same time; both Revolutions belonging to the same possible world. Secondly, the possibilities included in contrast spaces are usually too broad to belong to a single possible world: the possibility indicated by ‘English Revolution’ will not be a sharply delineated logically unique possibility. Garfinkel suggests that contrast space be regarded not in

terms of ‘possible worlds’, but in terms of Wittgenstein’s ‘logical space’ (developed in the *Tractatus*)²².

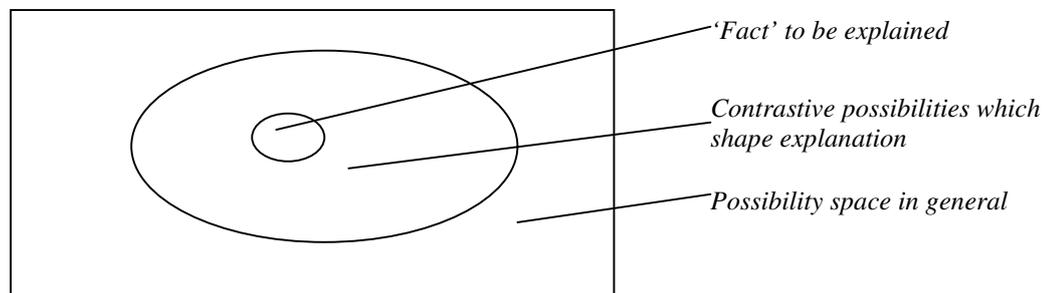
Two historical examples previously introduced demonstrate the need for an idea of ‘contrast space’. Hegel asked ‘why was there a revolution in France rather than in Germany?’. This question is an intelligible contrastive demand, yet its intelligibility does not rely on there being one definite possibility corresponding to ‘the German Revolution’. Rather, a spread of possibilities best grounds the intended contrast. We do intuitively have an *approximate* idea of the kind of event Hegel had in mind – that it is roughly similar to the French Revolution, but taking place in Germany. Yet it would be difficult and, for most purposes, quite pointless to delineate these possibilities precisely.

A still clearer demonstration of the need for ‘contrast space’ can be seen in the second use of contrastive explanation, in constructing a target of the appropriate level of detail. This aspect of contrastive focusing was demonstrated with reference to Lefebvre’s desire to explain ‘the distinctive features’ of the French Revolution. Any given event contains as much detail as the investigator is willing, or able, to keep discovering. Yet the explanations we demand only require a certain degree of detail; a degree which varies according to pragmatic considerations. As the contrast space is ‘moved’ closer to the fact to be explained (in other words, as contrasts are chosen which are more similar to the fact), appropriate answers must refer to finer details of the causal history. Yet, the distance between fact and contrast space can hardly be defined precisely; we have an intuitive, approximate idea of what level of detail is appropriate for the particular explanation.

The idea of a contrast space allows us to specify the identity conditions for explanatory targets precisely. An explanatory target is defined according to what it is intended to exclude. The resulting explanation ignores the similarities between the explanatory fact and the contrasts, and focuses only on differences. Two key ways for contrast space to vary are according to the inclusion of more, less, or just different regions of contrast space (‘outward’ variation); and according to a greater or lesser encroachment into the possibility space of the fact itself (‘inner’ variation). These

²² Garfinkel, 39

two types of variation are demonstrated, respectively, in the examples of Hegel and Lefebvre, above. As an aid to understanding, the following diagram represents these ideas, in terms of regions of possibility:



The understanding of the relation of contrastive explanation to non-contrastive, and of the role of ‘contrast space’, can now be put to work in three ways. Firstly, in relating contrastive explanation to explanatory exclusion. Secondly, in relating contrastive explanation to the concept of ‘complete’ explanations. Thirdly, in answering the arguments of those who would reduce contrastive explanation to non-contrastive.

Exclusion and completeness

We have seen that KPEE does not fare well when targets of explanation are taken to be *events*. For it seems that there *can* be two complete, independent explanations of the same event, as they may legitimately focus on different aspects of that event. However, analysis of targets of explanation in terms of *properties* introduces a different sort of concern for KPEE. In this case, the degree of exclusion recommended by KPEE can appear much too weak: for only those explanations which focus on *exactly the same properties* are thereby candidates for exclusion under KPEE. Our intuitive judgements of explanatory exclusion are, it would seem, stronger than this: KPEE, conjoined with a ‘property’ analysis of explananda, does not seem to yield as many examples of explanatory exclusion as we would like.

How does a contrastive analysis affect KPEE? Using the notion of contrastive space, we should say that the contrastive spaces should be equivalent in order for the explanatory targets to be equivalent. Where the contrast spaces differ, so do the explananda; in this case (as far as KPEE goes), there is no explanatory exclusion. With regard to the difficulty facing a property analysis that explanatory exclusion is too hard to come by, a contrastive analysis appears to suffer still more. After all, I have previously argued that contrastive targets are often more specific than targets constituted purely according to properties of events.

Where contrastive explanation may offer an advantage in explicating KPEE is in offering an analysis of *partial explanatory exclusion*. The suggestion made in chapter four was that where Kim's criteria – completeness, independence and identity of target – are only partially realised, then the result is partial explanatory exclusion. In particular, we saw that it was unrealistic to suppose that (complete, independent) explanations which were directed at only slightly different targets thereby could not be exclusive. Whilst this is a general consideration concerning KPEE, the contrastive analysis of explanatory target is an ideal model by which to more accurately interpret the suggestion.

Given that an explanatory target is identified according to the contrast space which it excludes, similarity of target can be explicated *geometrically*. Partial exclusion may be the result either of the *overlap of contrastive spaces*, or as the *subsumption of one contrastive space within another*. The relation between Lefebvre and Skocpol's explanatory targets is an example of the former: the explanatory space of each includes the example of the English 'revolution', yet in other ways they diverge. Subsumption of contrast space enables us to understand more specific explanations such as Chartier's, which focuses purely on cultural origins, as compared to a wider focus of explanation such as Schama's, which focuses on cultural origins and more. Other types of origin in Chartier's explanation are 'screened off', or presupposed. Thus Chartier's contrast space is smaller than Schama's, and hence the resulting explanation has a smaller pool of relevant differences to include.

Recognising *partial* exclusion is an important advance, since it provides KPEE with some much needed 'bite'. KPEE may thereby fulfil a substantive role in considerations of explanatory exclusion; in addition, our intuitions that explanatory

exclusion is not such a rare occurrence are respected. Ultimately, the justification of partial exclusion due to KPEE must rest on similar considerations to the justification of KPEE itself. I argued that the role of KPEE was to be found in its ability to provide substantive conditions for when PEE₃ was satisfied; in other words, where explanations contain incompatible relevance claims. From this perspective, it is clear that explanations may focus on somewhat different targets, and yet contain implicitly contradictory relevance claims.

Contrastive explanation also helps to make sense of the problematic notion of *complete explanation*. Whilst the concept is not strictly necessary for an analysis of explanatory exclusion, it does feature in an important type of exclusion: KPEE. The benefit to be derived from a contrastive analysis is clear once we recall that a good contrastive question drastically restricts the range of causal history which is relevant. The question ‘why did the French Revolution come about?’ allows an indefinite degree of information to be relevant: any of the causes of the French Revolution. The contrastive question ‘why did the French Revolution, rather than a German Revolution, come about?’ excludes as irrelevant all similarities between the French and German cases. And it is to be expected that such similarities actually comprise the vast majority of the causal history of the French Revolution. In any case, even if the range of relevant causes remains too wide, the contrast space may be reformulated so as to further restrict this.

Whilst DC does not necessarily restrict relevant differences between fact and foils to a single one, a good choice of contrasts leads to at least a small number of differences. Given that a contrastive question only demands information which satisfies DC, complete *contrastive* explanations are, therefore, practically possible. ‘Complete’ is therefore made a *practically* tractable notion, which is clearly required for considerations of explanatory exclusion due to KPEE. It should be noted that, even utilising a contrastive analysis, a complete explanation cannot be one which mentions *all* causes of the explanatory target. This is so because it is highly likely that a causal difference implies the existence of earlier differences in addition (for all but some genuinely stochastic cases of causal mechanism). However, we can at least make

sense of a complete explanation *at a given time slice* of the history of the explanatory target.

It is of interest to compare this contrastive understanding of completeness to another potential formulation. In chapter four I pointed out that interpreting ‘complete explanation’ as meaning ‘sufficient explanation’, in the sense of providing a Deductive-Nomological explanation, raises certain difficulties. However, in the sense that they are sufficient, D-N explanations are complete. This ‘completeness’ is achieved through the inclusion of a relevant law, stating that the explanandum *had to* come about, given the existence of the particular fact (the cause).

Yet no law is universally true: in other words, no law actually guarantees, in all situations, that the presence of the cause is sufficient for the explanandum. All laws are *ceteris paribus*; they are of the form ‘all x’s are Y, all else being equal’. For example, ‘all masses attract’ (the law of gravity), would not be true in a situation where some other type of force restrained those masses. Less universal laws than the law of gravity have further degrees of presupposition. ‘Any social class with economic power but not legal/ political power will soon gain the latter’ clearly holds in a fairly limited range of situations (there must be no undue extra-national influence, the leadership of the bourgeoisie must be sufficiently dynamic, and so forth).

What is interesting about the *ceteris paribus* nature of laws is that this feature has the effect of *presupposing* much of the causal history of the explanandum. It is only through these presuppositions that sufficiency – hence, in a sense, completeness – is possible. And we have seen that it is through presupposing those aspects of the fact’s history which are similar to those of the foils that contrastive explanations may be complete. In each case, much of the causal history is presupposed, or ‘held fixed’, allowing a complete explanation of the remainder. Whilst the production of an explanation of D-N form is one way to provide complete explanations, given a well formulated contrastive scenario, it is not the only way.

Is contrastive explanation always applicable?

In the final section I want to deploy the understanding of contrastive explanation developed in this chapter to two related questions. First, is contrastive explanation *eliminable*? Second, to the extent that it is not eliminable, is contrastive explanation applicable to all explanatory episodes? Much of the substance of an answer to these questions has already been provided. I need only add here one further idea.

Contrastive explanation can be both a *method* and a *means of analysis*. Very often, contrast is the most suitable means to restrict relevant explanatory information. The use of contrast enables a historian to focus on which aspects of a target event need explaining, and which elements of the history need citing. As such, contrastive explanation is a method of requesting explanations.

However, I have noted that the same – empirically equivalent – question can be asked in a number of different ways. It may be asked contrastively. It may be asked by *explicitly* directing attention towards certain aspects of the event and its history (‘why did the French Revolution come about? I only want to know about ...’). It may be asked by presupposing aspects of the causal history, which are therefore to be excluded from the information (if you knew that your audience did not want to hear about economic aspects of the history of the French Revolution, you would exclude these from an explanation). It may be asked in still other ways: Skocpol, for example, restricted relevance according to the requirement that only those facts which found an analogue in the comparison cases were relevant. As I have suggested, contrast is a device which is tolerant of ignorance (useful where we simply don’t have the knowledge to be explicit about what is to be excluded or included), which is flexible and which corresponds well to actual practice. Yet we cannot pretend that all explanations are asked for contrastively and answered with reference to explicit contrasts.

Even where contrast is not, and perhaps should not be, used in practice, contrastive explanation remains appropriate as a means of analysis of explanatory targets. In this sense, contrastive explanation is a ‘meta-theory’ of explanation. The explanation may

not *actually* have been asked for and provided contrastively, but analysing the explanation contrastively can still carry the benefits previously listed: of making sense of explanatory relevance, appropriateness and refusal, the detail of explanatory targets, explanatory exclusion (particularly with reference to KPEE) and completeness.

This dual notion of contrastive explanation enables us to claim that the analytical benefits of contrastive theory can be applicable even where the explanatory form is not explicitly contrastive. Contrastive explanation seems particularly inappropriate in two types of explanation found in the case studies, which I shall call (somewhat roughly) ‘what’ and ‘how’ explanations. ‘What’ explanations are those which, like Cobban’s, approach their target through means of factual (rather than ‘why’) questions: such as ‘who were the bourgeois? What was the feudalism which they are supposed to have overthrown?’²³. ‘How’ explanations are narrative explanations, exemplified here by Michelet and Schama. These explanations are those which aim to provide the detailed mechanisms by which an event came about.

In fact, ‘what’ explanations are equivalent to those questions which *explicitly* limit relevance, considered previously: those of the form ‘why P, with respect to facts {x}?’ The elements of the event in question which are not queried are those which are presupposed by a ‘what’ question. Given what I have said about such explicit restrictions, this is analytically equivalent to a contrastive formulation; where the contrast space shares with the target fact those features which are presupposed. (I shall say more about ‘what’ explanations in the next chapter.)

‘How’ explanations – narratives – fit awkwardly with an idea of contrastive explanation due to the detail provided, and due to multiple foci. A narrative, at its most basic, explains the set of temporally ordered facts {c, d, e, f} such that c explains d, d explains e, and e explains f. No summary reason or set of reasons is usually provided for the final state (certainly this is the case with Schama’s narrative, for example). For this reason, it may not be possible to provide one particular contrast

²³ Cobban, 24

space which is appropriate for the narrative explanation as a whole, simply because there is no one target for the explanation. However, it is still true that contrast is applicable at different points in the narrative.

I wish to consider two arguments whose conclusion is that contrastive explanation is either not applicable, or less applicable than I have suggested. The former, ‘strong’, claim is that there are really no contrastive explanations. This argument is based on the idea that the contrastive question ‘why P rather than Q?’ can be reduced to non-contrastive notation; such that ‘why P rather than Q?’ is equivalent to ‘why P?’ and ‘why not Q?’. Further, it is pointed out that in many cases, there is no contrast intended; the questioner may simply want an explanation for ‘why P?’. Temple raises this problem for the contrastive theorist:

‘Thus, ‘Why P & ¬Q’ seems to be a fully adequate reading for questions of the form ‘Why P rather than Q?’. ... if the contrast theorist wants to avoid this result by insisting that there is, after all, more to ‘Why P rather than Q?’ than is captured by ‘Why P & ¬Q’, then it is incumbent on him to provide an account of this excess meaning.’²⁴

I believe resources exist in the above account of contrastive explanation which enable an account of this excess meaning to be provided. The contrast space shapes those aspects of P which require explanation, and which elements of the history of P are relevant in this explanation. To translate the contrastive question as a conjunctive question is to miss this point entirely, for the connective ‘rather than’ is not truth functional (as ‘&’ is). ‘Rather than’ depends on the empirical content of the connected terms, as can be seen in the demand that the answer to a contrastive question provide *only differences between fact and contrast space*.

Ruben’s objection is similar to Temple’s: he states that ‘explaining the conjunction of facts is to explain the contrastive’²⁵. Yet the crucial problem with this formulation is that we must be careful about the idea of explaining a fact alone; does it mean to provide all information relevant to that fact? We have seen that this is impossible. Or

²⁴ Dennis Temple, ‘The Contrast Theory of Why-Questions’ in *Philosophy of Science* 55 (1988), 150-151

²⁵ David-Hillel Ruben, ‘Explaining Contrastive Facts’ in *Analysis* 47 (1987), 36

does it mean to provide any information relevant to that fact? But this will lead to pragmatically irrelevant explanations; exactly the unfortunate result which contrastive explanation is introduced in order to avoid. True, target facts could be properly limited non-contrastively; I have provided some examples above. Yet we have seen that contrastive explanation is still not analytically redundant, as Ruben implies, not least due to the unique advantages of a contrastive analysis previously mentioned.

It must be said that Ruben is sympathetic to uses of contrast which ‘only’ serve to define precisely what is to be explained; he cites Lewis’ paper, ‘Causal Explanation’, as a defence of this type of contrastive theory. To this extent, Ruben’s position is similar to that defended in this chapter: contrast is to be viewed as a *tool*. However, no writer to my knowledge has defended the idea that contrastive explanations *can’t* be translated into non-contrastive terms. Even such a paradigmatically ‘contrastive’ theorist as Garfinkel, for example, strongly emphasises the equivalence between contrast space and the presuppositions of explanation. Does Ruben’s conception of contrastive explanation, therefore, differ to mine? Perhaps in two ways: I locate unique advantages which are to be found in analysing explanation contrastively; and I would claim that a contrastive approach is always a potential analysis of explanatory target.

The second, ‘weak’, argument against contrastive explanation denies this latter claim. Whilst some explanations may be seen as irreducibly contrastive, not all can be; in particular, where no obvious contrast suggests itself, and so where the default contrast ‘rather than not’ is added. And the problem with this is simply that ‘why P rather than not-P?’ is equivalent to ‘why P?’; thus the contrast is dispensable. As Ruben claims:

“‘Explaining why e occurred rather than not’ is just a tedious pleonasm for ‘explaining why e occurred’, which is to explain a non-conjunctive (and non-contrastive) fact. Indeed, if ‘e’s obtaining rather than not’ did name a contrastive fact, the very distinction between contrastive and non-contrastive facts would be at risk.’²⁶

²⁶ Ruben, ‘Explaining Contrastive Facts’, 37

A less expected source of concerns regarding contrastive explanation can be found in Lipton's 'Contrastive Explanation'. Lipton's argument for contrastive explanation is based on the fact that it is sometimes harder to answer a contrastive question than a non-contrastive question; whilst it is, in other cases, sometimes harder to answer a non-contrastive question than a contrastive one. The two formulations cannot, therefore, be equivalent. Non-contrastive questions are pre-supposed by Lipton; yet we are not told what it is to explain a non-contrastive fact. Further, Lipton's examples of explanation of non-contrastive facts can be interpreted rather as explanations of different, yet still contrastive, facts.

Lipton's first example of a contrastive explanation is 'Why did Jones, rather than Smith, get paresis?'²⁷. The reply is that 'Jones had syphilis'. (The relevant background knowledge is that syphilis is necessary, but not sufficient for paresis; and that Jones, but not Smith, had syphilis.) The example of a similar, but non-contrastive, explanation is 'Why did Jones get paresis?'. However, this question suggests two potential (implicit) contrasts. We could have in mind the contrast 'Jones *without* syphilis'; the question would therefore suggest the answer 'because he got syphilis'. Or, we could have in mind 'Jones *with* syphilis'; the question would therefore suggest the response that 'we don't have the medical information to know!'. Lipton assumes the latter contrast, which is why he says that explaining the contrastive fact (I would say, the explicitly contrastive fact) is easier. The important point is that, even in discussion of a seemingly 'non-contrastive' explanatory target, an implicit assumption of relevant contrasts is still required. I take the moral to be that contrast is often implicit; many more questions are contrastive than appear to be so by their surface form.

What of seemingly more clear-cut examples of non-contrastive explanatory targets? In particular, what of the cases located by Temple and Ruben where, if asked for an appropriate contrast we could think of nothing better to add than 'rather than not'? I agree that *taken at face value* such a contrast adds nothing to simply asking why the fact obtained. The contrast 'rather than not' does not restrict relevant explanatory information, since in this case contrast space is equivalent to possibility space in general. To answer such a question, we would need to differentiate the history of the

²⁷ Lipton, 'Contrastive Explanation', 211

fact from the histories of all other possibilities; a task which amounts to the provision of (at least!) a full set of conditions of the type envisaged by Mill.

What I suggest is that we don't take the contrast 'why P rather than not-P?' at face value. Although we are sometimes inclined to suggest that the contrast space is simply the negation of the fact to be explained, it is clear that implicitly the contrast space is more determinate. Garfinkel made the same point: we need to make room for '*limited* negation, a determinate sense of what will count as the consequent's "not" happening',²⁸.

We have already found examples of what 'limited negation' means in practice: particularly in Lefebvre's explanation of the French Revolution's 'distinctive features'. The French Revolution's not happening – the 'not-P' of the contrast – excluded from the contrast space such possibilities as the revolution starting a day later, the violence resulting in a single difference in numbers of executions, and so forth. Whilst these exclusions are rarely stated explicitly, we know that they are implicitly present simply by considering that a response which did not pay heed to them would be pragmatically incorrect. Although the contrastive formulation will frequently be abbreviated as 'why P rather than not-P?', in practice no demand for explanation is expressed with the expectation that the answer will provide information sufficient to differentiate all other possibilities. All historical explanation, in practice, may be beneficially analysed as contrastive explanation.

Summary

Our starting point was Kim's PEE, which is one salient way to guarantee explanatory exclusion. Explanations exclude each other where they are explanations of the same event, and they are not related as dependent, or are two incomplete parts of a wider,

²⁸ Garfinkel, 30

complete explanation. Assuming explanatory realism, this idea must be along the right lines. But it goes wrong in specifying explananda as events.

I have argued that explananda are best seen as contrastive facts. The target, or explanandum, of an explanation is a combination of the fact to be explained and a contrast space which shapes the appropriate range of explanatory information. Contrastive explanation states that explananda are individuated according to the entity (event or property) which is the target fact, *and those possibilities which a good answer should exclude*.

Using this notion, we can gain insight into some of the ways that explanations can be consistent, exclusive or overlapping. Where explanations share both fact and contrast space (and Kim's two other criteria hold), they exclude each other; both cannot be correct. Where explanations differ in fact or contrast space, they can be (according to KPEE, at least) consistently held. Where the fact is the same and contrast spaces overlap, the explanations are partially excluded. Yet, the necessarily indeterminate nature of contrast space precludes entirely definite answers to the question of explanatory exclusion.

I have also developed an account of contrastive theory, such that contrast plays a dual role: methodological and analytical. Many explanations are framed in contrastive terms, explicitly or implicitly, and it makes sense to understand the structure of these. Broadly speaking, I accept that Lipton's Difference Condition, when combined with an understanding of contrast space more similar to Garfinkel's, is the best way to understand contrastive explanation.

But, in addition, contrast can be a useful tool by which to analyse all explanations. Its uses include, but are not restricted to, comprehending the nature of explanatory focus, explanatory detail, explanatory refusal, and explanatory exclusion. Such demonstrations of the utility of contrastive theory, supplemented with a developed understanding of the determinate nature of contrast space, refute the eliminativist arguments purporting to show that contrastive explanation is dispensable.

Chapter Six: Theory and Concept

To what extent is there a concept of historical explanation?

The principles of chapter four are clear cut: explanations are exclusive whenever the targets, explanans, or, in particular, the relevance claims of explanans contradict. However, there remains a concern that these fairly clear cut principles may not be appropriate where historical accounts seem to operate according to quite different notions of what ‘explanatory relevance’ amounts to.

Where the competing explanations are both causal, we have made some headway in providing criteria according to which the claims of relevance are contradictory. In particular, we have seen that one important way to specify when causal explanations disagree over claims of causal relevance is to make use of the notion of a ‘complete explanation’. KPEE made use of the presupposition, feasible when conjoined with a viable statement of explanatory realism, that there cannot be more than one complete and independent explanation of any given target. KPEE can be defended with regard to causal contrastive explanations, with the possible exception of cases of over-determination. But is KPEE true of explanatory competition in general? In other words, is it true that there can be no more than one complete and independent explanation *of any kind* for any given target?

No argument for this conclusion has been provided by Kim. Recall that his argument for the ontic PEE proceeded by considering all possible relationships between explanations, based on the assumption that explanatory relationships mirror appropriate ontic relationships. Yet the five possibilities are implicitly causal, as Kim recognises. Indeed, not only has an argument for a more generalised principle of the form of Kim’s not been given, but there are additional reasons for being sceptical of the possibility of such a task. It seems quite possible that there could be, for example, a teleological explanation for E, and a causal explanation for E, each being complete

and independent, yet not mutually exclusive. It could even be maintained that a one-sided version of events results *unless* different kinds of explanation (like causal and teleological) are both given. It is clear that more detailed consideration needs to be given to the extent to which there are different types of explanation, and how the adoption of different types affects explanatory exclusion.

In general, it seems that explanations which hold radically different conceptions of 'explanatory relevance' are not potential candidates for exclusion. We might think that nomological (scientific) and rational explanation, for example, do not come into competition with one another because they are *incommensurable*. The concept was introduced (independently) by Kuhn¹ and Feyerabend², who both argued that (at least some) scientific theories were incommensurable. The meaning of the term differs somewhat between Kuhn and Feyerabend, but the core is that incommensurable theories are *not comparable in a deductive or axiomatic manner*³. It is important to see that incommensurability is different to incompatibility; incommensurable theories are those which cannot be said to be compatible or incompatible. Similarly, incommensurable theories cannot be compared as better or worse in a neutral or rational manner.

Incommensurability is a concept that may have applicability outside the realm of scientific theories, for which it was first introduced. In particular, it seems reasonable to regard certain explanations as incommensurable. We would not seek to compare (in a deductive fashion, at any rate) a poetic and a biological explanation for the occurrence of love, nor would we wish one explanation to exclude the other. Likewise, it is plausible that religious and scientific explanations of the formation of the universe may be complementary, rather than exclusive, in virtue of having very

¹ Kuhn, 111-135, 198-204

² For example, Paul Feyerabend, 'On the 'meaning' of scientific terms' in *Journal of Philosophy* 61 (1964): 266-273, and 'Consolations for the Specialist' in *Criticism and the Growth of Knowledge*, eds. Imre Lakatos and Alan Musgrave (London: Cambridge University Press, 1970): 197-200

³ In more detail: a Kuhnian definition would be in terms of the Kuhnian notion of a *paradigm*, such that incommensurable theories are those which can only be judged within the terms and standards of their governing paradigm. Therefore, since for Kuhn all mature science is governed by paradigms, incommensurability is the norm. Between paradigms, there can be little, if any, communication. Feyerabend, however, sees it important to find a way to compare incommensurable theories. He regards it possible for incommensurable theories to be compared, just not in a deductive manner. Feyerabend regards incommensurability as being less common than Kuhn.

different purposes and standards of explanatory relevance; at least under non-literalist interpretations of creation stories. Both of these examples illustrate a wider notion of incommensurability. Do any of our historical examples share similar characteristics?

The fundamental task for this chapter is, therefore, to ask whether there are ‘different types’ of explanatory relevance relations at work in historical explanations. In answering this question, I shall outline a general conception of what historical explanation consists of. The investigation will initially be conducted in the descriptive mode, by examining the theory at work in the case study explanations. The philosophical issue of ‘types of explanation’ was first raised in chapter three, but at this point the issues can be specified more precisely. The issue raises two related questions: those of *reduction* and *pluralism*.

Firstly *reduction*: are the different modes of explanation (in particular, those identified in the descriptive study to follow) representative of totally different types? Or, perhaps despite appearances to the contrary, should we regard the differences as more superficial? The latter judgement could be supported either through a demonstration that certain ‘types’ of explanation are actually reducible to another, more fundamental type; or through a demonstration that the different ‘types’ are capable of subsumption under a single, more general notion.

Secondly *pluralism*: to the extent that this reduction (or unification) is not possible, we are faced with a choice. Either we take the non-pluralist view that a certain type, or certain types, of explanation are appropriate, whilst others are not. Or we adopt the more pluralistic standpoint that each type is different, yet each performs a respectable explanatory task. The import of the second question depends on the first; it is only relevant if the first question is answered in a non-reductivist way. In all cases, ‘different type’ should be primarily understood with regard to differing conceptions of explanatory relevance, as they relate to the potential for explanatory exclusion.

In this section I will outline possible responses to these questions which have actually been adopted in the philosophical literature: reductivist and non-reductivist, pluralist and non-pluralist, as well as more intermediary positions. This survey will lay the foundations for my own understanding of historical explanation, which will be stated

and defended subsequent to the descriptive study. I will use this model to suggest that whilst incommensurability of explanatory theory may reasonably be thought to be relevant to explanatory competition in general, historical explanations may be regarded as commensurable when considering explanatory exclusion.

In chapter three I outlined the *explanatory reductivist* position, particularly with regard to the causal reductivist arguments of David Lewis and Eric Barnes. In this context, causal reductivism is the thesis that all good scientific explanation is causal explanation. If this is the case, then there is only one ‘type of explanation’, and so the principles of explanatory exclusion can be applied in a straightforward manner. A corollary of the causal-reductivist analysis is to yield an answer to the question of *pluralism*: for where scientific explanations cannot be reduced to causal explanations, they are therefore inadmissible.

The causal reductivist must confront the fact that many good explanations certainly do not *seem* to be causal. The most relevant examples of this phenomenon for the present study will be found in the case-study historical explanations, considered in the next section. But the following are examples of the sort which pose particularly stark problems for a causal reductivist view:

(i) (from Lipton⁴): ‘Why, when sticks are thrown into the air and their instantaneous positions recorded at some point in their descent, are appreciably more near the horizontal axis than the vertical? Because there are two horizontal dimensions and only one vertical (hence there are more ways for a stick to be near the horizontal than the vertical).’ The explanation cites a geometrical fact, not a cause.

(ii) (from Ruben⁵): ‘Why is element A of a high pH? Because item A has a high concentration of H ions.’ In this case, the explanans and explanandum refer to the same item. However, whilst self-explanation may be appropriate, self-causation is – presumably – ridiculous.

⁴ Lipton, ‘Contrastive Explanation’, 208

⁵ Ruben, *Explaining Explanation*, 218-219

(iii) (from Ruben⁶): ‘Why did this salt dissolve? Because it was placed in holy water.’ The explanation cites a true fact, which was indeed causally relevant. However, the explanation appears odd – if not downright false – due to the description of the water in question. Whilst it may be true that placing the salt in water caused it to dissolve, describing that water as *holy* water weakens the causal explanation, perhaps to the extent that it is no explanation at all.

My final model will take a line similar to the causal reductionist, so I will have to say something about such supposed counter-examples. Where my conclusions depart most clearly from those of the causal reductionist is in asserting that citing a cause is insufficient to provide a good explanation. Recognising that causes can be described in different ways, ways which are more or less explanatory, will be crucial in taking account of the explanations that historians actually give. A ‘bare’ theory of causal explanation, without the necessary focus on differing ways of describing causes, simply fails to make sense of the range of types of historical explanation that there is.

Even if a universal reduction of explanation to causal explanation is not thought possible, *particular reductions* may be advocated. One example is the reduction of intentional explanation to causal explanation. Roughly speaking, the intentional explanation E_i : ‘A x’d in order to y’ can be taken to mean ‘A intended y, and this intention caused them to x’. The intentional explanation is re-interpreted so that it no longer points towards the future, but towards the past, as good causal explanations should.

There are notorious problems with such attempted analyses, the most fundamental of which is demonstrated by Davidson’s rock climber⁷. A climber in a precarious position intends to let go of the rope holding his colleague in order to stop himself falling too. Yet the very thought of letting go of the rope makes the climber’s hands perspire to the extent that he slips accidentally. In this case the climber intended to stop himself falling by letting go of the rope, and the intention caused the letting go. Thus the example is one of intentional action, according to the proposed causal

⁶ *Ibid.*, 193

⁷ Donald Davidson, ‘Freedom to Act’ (originally 1973) in Davidson, *Essays on Actions and Events*, 79

analysis. But the intention to stop himself falling caused the climber to let go of the rope *accidentally*; so the climber cannot truly be said to have let the rope go in order to stop himself falling. The twists and turns resulting from such counter-examples are too numerous to detail here, though most do assume that a causal analysis of teleological (intentional) explanation must, in principle, be possible⁸. I shall return to this issue towards the end of this chapter.

The diametrically opposite position is that of the *non-reductivist*. Different types of explanation found in different historical accounts really are as methodologically separate as they appear, and for that reason they are incommensurable. Hayden White's typology of differing modes of 'formal argument', introduced in chapter three, exemplifies a non-reductivist philosophy of explanation. Recall that we struggled to say what these 'types' had in common: the weak answer could only be that they somehow made the historical record 'meaningful'. Another, more limited but clearer, example is provided by Dray's understanding of descriptive explanation. Dray believes that 'explaining what' (descriptive explanation) and 'explaining why' are radically different types of explanation. It would be pointless to attempt to make one do the job of the other, since the two activities have complementary and irreducible roles. Descriptive explanation and 'why' explanation cannot, therefore, exclude one another. Again, I shall return to Dray's idea, and descriptive explanation in general, later in this chapter.

A rather different sort of non-reductivism provides a further possible position in relation to the questions of reduction and pluralism. A strong tradition in certain fields of historical scholarship is to disparage the role of causal explanation in history. Traditionally, such a position has been motivated by the arguments of the nineteenth-century hermeneutic theorists (for example, Winch's theory, introduced in chapter three), but is today more commonly reliant on post-structuralist theory. Different types of explanation – in this case, the *causal* and the *interpretative* – are therefore asserted to be separate, not reducible to one another. Yet, only one of those types –

⁸ See, however, Scott Sehon, 'Deviant Causal Chains and the Irreducibility of Teleological Explanation' in *Pacific Philosophical Quarterly* 78(2) 1997: 195-213 as an example of one who affirms the non-reducibility of teleological explanation

the interpretative – is regarded as a proper basis for explanations in history and the social sciences. With regard to historical and social scientific explanation, this position is therefore non-pluralist. There can be no question of explanatory exclusion between explanations of two different types if this view is correct, simply because causal explanations are not really respectable historical explanations at all.

Explanatory Theory

In the following descriptive survey, I use various conceptions of explanation to categorise the historical works: *Deductive-Nomological*, *Nomological-Comparative*, *narrative*, *teleological* and *descriptive explanation*. These (embryonic) analyses will not necessarily be a part of the more considered, general theory of historical explanation which I shall subsequently propose. I hope that by relying on such a wide range of theories of explanation for my initial typology, I can escape the charge that I am simply *presuming* that the different explanations are all of the same type. To some degree, the categorisation is artificial; both in the sense that all the historical works display a mixture of the different types, and in that the different types are, in any case, not always distinct. The concepts provide, however, a descriptive starting point; and it will be a virtue of my approach to historical explanation that it can make sense of such overlapping theoretical methodologies.

There are, in addition, other types of explanation which are not represented in what follows, for the reason that I have selected only that which appears to be the most distinctive or pervasive mode of explanation in each account. All the accounts contain causal claims (despite the reluctance of some of the historians to recognise this; an attitude which will be examined in the following section). All, with the exception of Skocpol, contain ‘rational explanations’: explanation of action by reference to reasons which a particular historical individual, or social groups, had for performing the actions they did.

Deductive-nomological

The core of *Lefebvre's* explanation of the French Revolution is the application of a Marxist law of history:

‘These groups [the nobility and clergy] preserved the highest rank in the legal structure of the country, but in reality economic power, personal abilities and confidence in the future had passed largely to the bourgeoisie. Such a discrepancy never lasts forever. The Revolution of 1789 restored the harmony between fact and law.’⁹

The explanation reconstructed in deductive-nomological form is as follows:

- (P1) The nobility and clergy had the highest legal rank; the bourgeoisie had the economic power
- (P2) Whenever one class has economic power but not the highest legal rank, its legal rank will change (for the better)
- (C) The bourgeoisie’s legal rank changed (for the better)

Such an explanation is explanatory due to its demonstration of *expectability*. Given a deep (long lasting, robust) fact about French society, something like the French Revolution as it actually was had to happen. It is undeniable that, if such an explanation is available, it is satisfying – we see why the French Revolution had to happen.

However, we should closely examine the *target* of Lefebvre’s explanation. The above explanation does *not* explain the actual event which was the French Revolution, only that there was a change of legal rank in the bourgeoisie, at some time. There was, however, a change in legal rank of the bourgeoisie in Britain before the French Revolution, but no revolution there remotely akin to the French. I mention this not as a criticism of Lefebvre, for he clearly recognises the need in his wider account for further detail. Rather, the important point to note is that one significant way that ‘types of explanation’ can differ is in the precision or generality of their targets. This

⁹ Lefebvre, 2

aspect of explanatory targets has been explicated, in chapter five, using a contrastive theory of explanation.

Nomological-comparative

Skocpol's theoretical basis is consciously scientific. Her method, as we have seen, is to select both positive and negative comparison cases. These cases are then compared according to Mill's methods of Agreement and Difference, in order to discern causes. We must remember that Skocpol's overall aim is to develop a general theory of social revolutions. She does wish to explain why the French Revolution took place, as do the other historians considered here, but only in the context of an account of why, in general, social revolutions take place. I take it then that Skocpol's explanation of the French Revolution will, if correct, be satisfying in part because of its ability to be applied to further cases.

Skocpol's account relies on the citing of laws just as Lefebvre's, hence it is also nomological. However, Skocpol cites three nomological factors, not one: state organisation (greater susceptibility to administrative and military collapse leads to a greater chance of revolution), pressure from more developed countries abroad (more pressure leads to a greater chance of revolution), and agrarian socio-political structures (those which facilitate peasant revolt against landlords lead to a greater chance of revolution).

With regard to the relationship between the generality of laws cited in an explanation and generality of explananda, Skocpol's target is correspondingly more precise than Lefebvre's. It is worth noting that whilst this may quite justifiably be regarded as an advantage for Skocpol's nomological explanation, we saw in chapter five that it is not correct to think that the more detailed the explanation, the better. Consider again Garfinkel's example: 'why did the car skid there (rather than 5cm further on)'. In the context of contrastive explanation, there will always be some level of detail which is quite properly made irrelevant according to the selection of contrast space.

One commentary on Skocpol's work which is of particular relevance to the concerns of this chapter is Manicas'¹⁰. Manicas claims that though Skocpol's explanation is sound by its own standards, it is not really a historical explanation at all. Historical explanations, for Manicas, must be 'processive'; in other words, they must trace a chain of particular events in a narrative fashion. In the context of the eight historians considered in this work, I am not sure that Skocpol is so different in this respect. In any case, we are owed an argument as to why only 'processive' explanations should be considered historical explanations. Manicas' notion of 'historical explanation' is a good example of the sort of restrictive conception against which I contrast my own conception, to be developed in this chapter.

Teleological

The core idea of *Hegel's* theory of historical explanation is that causal explanation alone is insufficient to provide a proper understanding of historical phenomena. Whilst the true Philosophical Historian should not depart from empirically determined facts, the sequence of facts must be configured in such a way as to make clear the internal essence of that sequence.

I believe that this demand is best understood with reference to Lakatos' theory of Rational Reconstructions, considered in chapter one. One demand of Lakatos' theory – the need for normative or rational explication – is that a philosophical history of science which does not make apparent the essential rationality of that history is, on that basis, deficient. Hegel simply widens this notion: the requirement to make clear the rationality of a sequence is justified not only for histories of science (or, perhaps, other intellectual developments), but for history in general¹¹.

The internal essence of human history is, for Hegel, the development of the spirit of the age, or 'Spirit'. This means the development of *freedom*; since Hegel claims that

¹⁰ Peter Manicas, 'Review Article: Theda Skocpol, States and Social Revolutions' in *History and Theory* 20(2) (1981): 204-218

¹¹ This interpretation gains some support from the fact that the type of history immediately preceding (and therefore conceptually the closest to) Hegel's preferred methodology of 'Philosophical History' is 'Specialised Reflective History'. This type of history is 'for example, the history of art, of law, or of religion'; Hegel, *Introduction to the Philosophy of History*, 9

‘the essence of Spirit is its freedom’¹². Hegel’s history therefore consists primarily of the tracing of the changing idea of freedom, that idea which motivates the French Revolution being near the end of the story. In tracing the notion of freedom, Hegel claims to make apparent the internal and rational thread of history, understanding of which can alone lead to a proper understanding.

Hegel’s theory of explanation is teleological in the sense that the explanation refers to the end-state of Spirit. Hegel’s account of the French Revolution only makes sense in the light of both where Spirit has been, and where it is going: towards perfect freedom. Hegel’s understanding of teleology clearly owes much to an Aristotelian metaphysics, wherein each thing has a potential which it struggles to achieve. To cite this potential is to cite the Aristotelian ‘final cause’, which is thereby one way of explaining the activities of that item.

However, Hegel also regards ‘usual’ (efficient) causation not just as complementary to ‘final’ causation (as a more orthodox Aristotelianism such as that suggested in chapter three would assert), but as the ‘servant’ of the latter¹³. I take this to mean that a teleological understanding operates by marshalling physical causation in such a way that causes are arranged so as to exhibit teleology (an idea which will be developed later in this chapter).

Hegel’s teleological explanation gains its support from the belief, acquired philosophically and non-empirically, that since there is only one true and right way of thinking and acting (with respect to ‘freedom’), history will eventually reach that point. It is for this reason that Hegel’s theory has been seen as paradigmatic of ‘speculative’ history. If this justification for Hegelian teleology is along the right lines, it shows the explanation to be *external rational*.

To explain: in chapter one, I argued that the most fundamental question in categorising theories of rationality is to what extent the participant’s own aims and beliefs are considered sufficient to demarcate rationality of action. An explanation based on a theory of rationality which aims to make sense of actions only within their

¹² *Ibid.*, 20

¹³ Georg Wilhelm Friedrich Hegel, *Encyclopedia of the Philosophical Sciences*, trans. W. Wallace (Oxford: Oxford University Press, 1975; originally 1817) Book I: Logic, Section 204

own context of beliefs and aims can be called an instrumental or *internal rational* explanation. Hegel, however, clearly conceives of ‘rationality’ as being a more demanding notion: the end state of Spirit is the most completely rational state, no matter what any participant of history believes or wants. Hegelian rationality is therefore *external* to the circumstances of those participants.

Narrative

Michelet is a ‘classic’ narrativist historian, writing at a time when to present a moralistic narrative was the standard means of presenting a historical account. *Schama* is a representative of the modern return to narrative, fashionable since the 1970s. Narrative explanations are, due to their very form, harder to summarise than nomological explanations; the latter can be summarised with reference to the key laws and initial conditions, as both Lefebvre and Skocpol actually do. Whilst D-N explanations are paradigmatic of *abstractive* explanation, narrative explanations are deliberately complex and detailed in an attempt to remain true to the perceived complexity of the historical record.

A few pertinent features of the respective narratives can, however, be noted. Michelet narrates the coming of the French Revolution ‘from the inside’; as (he thinks) the participants actually conceived of the event. The narrative is governed according to the belief that the revolution was a unitary phenomenon, an extraordinary event which embraced and brought together the entire French nation. And the revolution was fundamentally a change of basic ideals: the replacement of the Christian principle of grace with the Enlightenment principles of justice and reason.

These fundamental features demonstrate that elements of what we might regard as other types of explanation are also present in Michelet’s narrative explanation. Part of the explanatory worth of the narrative is to be found in its ability to make sense of the participant’s own actions (‘from the inside’). In other words, the worth of Michelet’s narrative should be judged, in part, according to its demonstration that the events have an ‘*internal rationality*’ to them. In addition, the unity of the French nation exerts a *teleological* pull on events in Michelet’s narrative. Finally, Michelet’s categorisation

of the Revolution in terms of Enlightenment and Christian ideas suggests similarities to ‘descriptive explanations’, to be considered next.

Schama adopts a very different understanding of the revolution. He regards the French Revolution as an illiberal ‘blip’ en route to a modern, liberal French nation. The causes of the revolution were accidents; the origins haphazard and chaotic. Whilst the revolution appears (to the participants and initially to us) to be of ‘epochal significance’, what in fact brought it about was a mixture of individual ambition, rivalry, and unfortunate short-sightedness. Further, and in direct contrast to Michelet, Schama insists that in a very real sense there was no one revolution, but a multiplicity.

How, in general, do narratives explain? In the basic sense, a narrative could be said to be a succession of causal explanations, each of which provides the basis for a single narrative step. (This idea was suggested with regard to ‘how’ explanations, in chapter five.) But the central question is how far, and in what way, a narrative’s explanatory power goes beyond a summation of the causal claims which constitute the narrative.

One way to claim that the explanatory power of a narrative does exceed that of the causal statements would be to make the case for the dependency of explanatory power upon the overall *configuration* of the set of events, such that those events are thereby given a comprehensible meaning (Hayden White’s theory represents one way of developing an intuition of this kind). Whilst a narrative cannot abstract in the way that a D-N explanation can, the narrative does in some sense unify the events in the account – whether this unification leads to the understanding that the revolution was a fundamental and inevitable change of ideas, or that it was essentially a mistake.

Descriptive explanation

Chartier, *Furet* and *Cobban* all, in somewhat different ways, structure their explanations through a mixture of descriptive and causal focus. All three make numerous causal claims, despite Chartier and Furet’s professed dislike of this mode of explanation (a theoretical attitude which will receive attention in the following

section). The more interesting feature of these accounts is, however, the use of description in an explanatory context. Whilst description is sometimes conceived as an activity distinct from explanation, we do need to account for the idea of explanation as *re-description*, or alternatively, *conceptual explanation*.

Descriptive explanation proceeds by describing the material under examination in a certain way, according to certain concepts. The common factor is that those concepts bring explanatory understanding of some sort: at a minimum, they should bring order to, or unify, the phenomena under examination. This is the role of the ‘*key concepts*’, identified in chapter two. In each historian’s work, there is room for one or two concepts which serve a particularly conspicuous explanatory role. A more detailed account of the explanatory work done by key concepts follows. One interesting question which I shall not raise in what follows is the extent to which the key concepts are *justified* by the historical material: I shall restrict attention to the explanatory import of the concepts (no matter to what extent they be empirically justified).

I would be tempted to identify ‘descriptive explanation’ with ‘interpretative explanation’ were it not for the fact that the latter phrase is usually understood with reference to a philosophy of social science inspired by the later writing of Wittgenstein. This ‘interpretivism’ maintains that explanation of social action must make reference to the social rules which the participants themselves use to orientate their behaviour. Further, this sort of explanation is explicitly regarded as being non-scientific and non-causal. Interpretivism is the guiding principle of Winch’s theory of social explanation, mentioned in chapter three. I would understand ‘interpretivism’, in this sense, as one type of descriptive explanation.

Much of Furet’s thinking on the origins of the French Revolution is shaped by his understanding of the revolution in terms of the two concepts of ‘state’ and ‘society’. Thus, ‘In the dialogue between societies and their States that is part of the underlying texture of history, the Revolution tipped the scales against the State and in favour of

society.’¹⁴ Furet’s fundamental characterisation, based around his key concepts, allows him to understand and interpret certain important features of the revolution, of which I shall mention three.

Firstly, the characterisation explains such fundamental facts as that French ideology and politics altered enormously during the revolution, despite (in Furet’s view) the fact that little changed economically or administratively. This is because where the balance of power swings in such a fundamental way from state to society, everything is up for grabs. Ideology could, and did, penetrate any aspect of life, and no ideological axiom remained any longer sacred.

Secondly, the concepts serve to frame Furet’s more detailed investigations. For example, he uses the concepts to emphasise the fact that the role of representatives of society in late eighteenth century France passed from the *parlements* to the *philosophes*. This is then of interest in explaining the abstract nature of subsequent political debate¹⁵. Finally, Furet’s interpretation of state and society allows a fundamental characterisation of the French Revolution to be provided: the revolution *was* the event which swung political power from the state to society. Thus the interpretation ‘has the ... advantage of restoring to the French Revolution its most obvious dimension, the political one’¹⁶.

Cobban’s key concepts are rather more numerous than Furet’s, though they do share a similarity in all being primarily social: ‘town’/ ‘country’, ‘poor’/ ‘rich’, in particular¹⁷. The resulting interpretation is, for this reason, less tightly unified. Cobban conceives of his task primarily as one of critical description. In particular, he wishes to describe the nature of social divisions (particularly, through an examination of ‘bourgeoisie’), the rural economic system (whether or not it was ‘feudal’), the motivations to incite revolution, and the economic effects of the revolution. Through this description, the various antagonisms which constituted the revolution can be stated. The revolution

¹⁴ Furet, 24

¹⁵ *Ibid.*, 37: ‘Thus the French, deprived as they were of true liberties, strove for abstract liberty; incapable of collective experience, lacking the means of testing the limits of action, they unwittingly moved toward the *illusion of politics*.’

¹⁶ *Ibid.*, 27

¹⁷ Interestingly, like Furet’s ‘State/ Society’, Michelet’s ‘Christianity/ Revolution’, and others, Cobban’s concepts come in opposed pairs. Perhaps this is because the explanandum is a revolution?

consists fundamentally of the grievance of the poor against capitalism, and the grievances of the poor against the rich in general.

Apart from stressing the necessity (but not the sufficiency) of cultural origins of the French Revolution, the most significant aspect of Chartier's work is to challenge prevailing definitions of culture. In reversing the assumed relation between practice and ideas, Chartier presents an alternative interpretation of the cultural origins of the revolution. This interpretation privileges practice and the less conscious aspects of social life, over the clearly articulated ideas of the intellectual elite. Chartier's approach to cultural origins shapes his understanding of which features of eighteenth century French life should be emphasised as leading to the Revolution: cultural practice, not explicit ideology.

In addition to the interpretation of the cultural aspects of eighteenth century France (the 'origins' of the Revolution), Chartier's work allows a conceptual understanding of the course of the Revolution itself. Chartier is concerned with the question of continuity or discontinuity: whether the excessive violence of the Revolution ran counter to the 'process of civilization' of the eighteenth century; and whether the dominance of the public sphere seen as the Revolution progressed is paradoxical when compared to the invention of the private sphere prior to the Revolution. Chartier concludes: 'the French Revolution had roots in the century that it brought to an end, even where it spectacularly seemed to run counter to the old evolution.'¹⁸

Causal historical explanation

The remainder of this chapter defends a general understanding of historical explanation. This understanding is formulated with the twin aims of being *permissive* enough to do justice to the preceding descriptive account; yet being *unifying* enough to be able to demonstrate that explanatory exclusion can be relevant across seemingly

¹⁸ Chartier, 197

different explanatory types. I claim that historical explanation paradigmatically consists of two elements: (i) claims of causal relevance; (ii) specification of those claims in such a way as to exhibit a certain explanatory virtue or virtues. The explanation must, therefore, meet two jointly sufficient criteria: the ‘cause’ must be actually occurrent and causally efficacious, and it must be cited in an appropriate way. This section will examine the causal element of this analysis, the following section the descriptive element.

Crucial to a defence of causal historical explanation is attributing the correct meaning to ‘cause’. I will develop this conception of causation by considering what appears to be an obvious problem for a general theory of historical explanation as causal: that certain historians explicitly repudiate the notion of causation.

In contradiction to my condition (i), Furet and Chartier wish to explicitly disavow any idea that causal explanation is possible, or useful, in history. Chartier states that ‘History has become more circumspect in the designation of causality’¹⁹, Furet that ‘a phenomenon like the French Revolution cannot be reduced to a simple cause-and-effect schema’²⁰. At the very least, such statements imply that not all good historical explanation is causal; further comments suggest that Furet and Chartier perceive the role of causality in historical explanation to be minimal. I shall select five central elements to their critique, the first three of which, drawn from Chartier, are explicitly based upon Foucault’s comments on historical explanation²¹.

The search for ‘causes’ is equated with the search for ‘origins’, but this latter notion makes certain undesirable assumptions. ‘Origins’ are brought together and unified, but this process ‘demands a retrospective reconstruction that gives unity to thoughts and actions supposed to be “origins” but foreign to one another, heterogeneous by

¹⁹ *Ibid.*, 1

²⁰ Furet, 22

²¹ Michel Foucault, ‘Nietzsche, Genealogy, History’ in Michel Foucault, *Language, Counter-Memory, Practice*, ed. Donald Bouchard (Oxford: Blackwell, 1977): 139-164

their nature'²². The unification of 'causes' is given by hindsight, whereas the events in fact had no real unity.

The "chimera of origins" also presupposes that

'historical becoming is organized as an ineluctable continuity; that events are linked together, one engendering another in an uninterrupted flow of change that enables us to decide that one is the "cause", another the "effect"'²³

The distrust of assumptions of continuity derives directly from Foucault (and, indirectly, from Nietzsche). Foucault would prefer history to abandon the attempt to trace a continuous sequence of cause and effect, and instead to embrace the idea of history as a sequence of accidents, deviations, reversals, errors, faults and fissures²⁴.

Connected to these critiques is a distrust of teleological assumption, which is, it is claimed, presupposed by the term 'origins'. Teleology is suspect because it assumes 'an ideal continuity': 'the notion of origin entails the further risk of proposing a teleological reading of the eighteenth century that seeks to understand it only in relation to the phenomenon deemed to be its necessary outcome: the French Revolution'²⁵. Revealingly, Chartier later accepts that, despite his criticism, a teleological treatment is to some degree necessary, since this is the only way to bring order to the historical material: 'History stripped of all temptation to teleology would risk becoming an endless inventory of disconnected facts abandoned to their teeming incoherence for want of a hypothesis to propose a possible order among them.'²⁶. The assumed equation of teleology and causality at this point is stark.

Furet and Chartier both regard 'cause' as insufficient, and perhaps unnecessary, in explanation of the French Revolution because 'cause' is limited to concrete, physical items. Discussing revolutionary action and ideology, Furet states that 'These 'events'', being political and even ideological in nature, invalidate by definition a causal

²² Chartier, 4

²³ *Ibid.*, 4-5

²⁴ Foucault, 146

²⁵ Chartier, 5

²⁶ *Ibid.*, 7

analysis based on economic and social ‘contradictions’²⁷. It is clear that Furet has in mind a Marxist analysis as his paradigm of a causal explanation. Likewise, Chartier warns us that ‘...attributing “cultural origins” to the French Revolution does not by any means establish the Revolution’s causes; rather, it pinpoints certain of the conditions that made it possible because it was conceivable.’²⁸ ‘Cause’, for Chartier, cannot be applicable to cases where the historian simply ‘makes an event conceivable’.

Finally, Furet and Chartier believe causal explanation to be insufficient, since in many ways the revolution was an entirely new event. Furet maintains that ‘The mere fact that the Revolution had causes does not mean that they are all there is to its history. ... the revolutionary event, *from the very outset*, totally transformed the existing situation and created a new mode of historical action that was not intrinsically a part of that situation.’²⁹ Chartier agrees: ‘Even assuming that the Revolution had many origins (intellectual, cultural, or other), its own history still cannot be limited to them.’³⁰

I cannot enter into detailed criticism of Chartier’s and Furet’s philosophical position, which draws on elements from Foucault, Nietzsche, and even from a classic Humean critique of causality. Fortunately, I don’t think this is necessary. All I intend to point out is that Chartier and Furet do in fact provide causal explanations *in the sense in which I understand them*; and that their theoretical worries are aimed, by and large, at a rather different target.

The minimal sense by which I understand ‘cause’ is simply that of *an antecedent which was empirically necessary in the circumstances for the phenomenon to be explained*. That antecedent may be intellectual, economic, social, or of any other type. Causes need not be rationally or teleologically linked to effects. There is a ‘continuity’ of some sort existing between cause and effect, but this continuity need not be ‘ideal’, fundamental or even particularly interesting. It is a ‘continuity’ which

²⁷ Furet, 23

²⁸ Chartier, 2

²⁹ Furet, 22

³⁰ Chartier, 8

does not preclude understanding the causal sequence as accidental or productive of the radically new. By contrast, in general, Furet's and Chartier's theoretical comments either presuppose that the 'cause' makes the effect rationally and teleologically comprehensible; or that 'causes' should be limited to socio-economic factors. There is no reason to accept either of these restrictions.

Furet's and Chartier's final criticism is somewhat different: that the revolutionary development cannot be accounted for simply by what preceded it. But once we recognise the revolution as a complex particular (extended over time), we can agree with the statement without thereby abandoning the demand for causes. For it is quite possible that earlier parts of the revolution causally explain later elements, in a way that could not have been foreseen by limiting attention to causes operative before the revolution. In this way we can make sense of the claim that the Revolution's history need not be limited to its (pre-Revolutionary) causes.

In general, my intention is to critique Furet's and Chartier's theoretical comments through an application of Reflective Equilibrium. Furet and Chartier do, despite claims to the contrary, provide causal explanations (seen both in the preceding descriptive survey, and in the following section). There is, therefore, logical tension between elements of the same historical account: in the juxtaposition of the *practice* of citing causes and the *theoretical* aversion to causal explanation. This tension can be resolved by specifying a viable meaning of 'cause' which can make sense of what Furet and Chartier do in their explanations.

The theoretical comments can then be read as a warning that causes *of a certain type* may not be assumed in explanation of the Revolution. As I have suggested, a causal explanation is not just a list of causes: those causes must be specified in a particular way. The target of Furet and Chartier's criticism are particular ways of specifying causes. Their arguments should be interpreted as an attempt to show that causal explanations phrased in rational or teleological terms will not be possible in historical explanation in general, and for the French Revolution in particular.

Clearly, my proposed proto-analysis of 'cause' raises many issues. Firstly, I use the term '*empirically necessary*' to exclude logical antecedents as causes. However, a

long philosophical tradition questions whether we can make sense of empirical necessity, over and above logical necessity. Secondly, I say that a cause must be ‘necessary in the circumstances’ in order to make sense of the idea that, whilst two causes may lead to (roughly?) the same effect, in the specific circumstances which obtained, only one cause could have led to that effect. However, genuine cases of over-determination will pose a problem for an analysis of this sort.

I don’t intend to explore these issues in any detail, since it is to be hoped that if a better (though similarly permissive) analysis of ‘cause’ can be provided, then that analysis could replace my own without affecting what I say about causal *explanation*. With regard to my development of the latter idea, it is of more use to return to the three examples, introduced in the first section, which seem to suggest difficulties with a causal approach to explanation.

Example (i), ‘Lipton’s sticks’, seems not to be a causal explanation, since the explanation is in terms of a global feature of our space-time structure. It is commonly thought that ‘causes’ should be restricted to events, or changes of state; a type of entity which ‘a global feature of our space-time structure’ manifestly is not. However, according to my definition of ‘cause’, example (i) does turn out to be causal, since it refers to ‘an antecedent empirically necessary in the circumstances’ (so long as the meaning of ‘antecedent’ does not restrict the types of entities which may be ‘causes’). The structural properties are indeed necessary, in the circumstances, for the position of the sticks to be as they are. Does this point to a weakness in my conception?

I would argue that it does not, considering that my conception of ‘cause’ is something of a term of art. As I have freely admitted, it is no thorough-going analysis of the term. Rather, the purpose of arguing that all historical explanations are (partially) causal is to allow us to conceive of the possibility that seemingly different types of historical explanation may yet be exclusive. For this task, the prime requirements for a definition of ‘cause’ are that it allows historians to substantively – empirically – disagree over the truth and relevance of causes; and that it can be used in an understanding of contrastive completeness of explanation. I am therefore inclined to

insist that we stick with my conception of ‘cause’, even where that conception is applicable to items which would not normally be regarded as causes. However, if this recommendation is too unpalatable, the core argument of this section may be rephrased accordingly, with no dire consequences.

Rather than claiming that all historical explanations feature a cause or causes, described in certain appropriate ways, the causal case may be regarded as a paradigmatic example situated within a more general schema. The more general schema would be that all historical explanations cite an *ontic* relationship. This ontic relationship may be causal, structural, dispositional ...: indeed, the sorts of ontic relationships mentioned in chapter three. These relationships may all produce ‘antecedents empirically necessary in the circumstances’; hence they are on a par with regard to explanatory exclusion. It is neater to regard all historical explanations as having a causal element; but if causal relationships are regarded as only one possible type of ontic relationship, and ‘ontic relationship’ is analysed broadly as I have analysed ‘cause’, then it makes no difference.

Example (ii), Ruben’s ‘self-explanation’, is intended to be incompatible with causal explanation, no matter what conception of causation is utilised. The example is compelling to the extent that we agree that *no* analysis of causation should permit the same item to be both cause and effect. I think that examples such as these present one sympathetic to causal explanation with a choice; my preference being for the second option. We might refuse to accept the example as a genuine case of explanation: this is Lewis’ strategy, in reply to an analogous case³¹. Lewis’ response makes it clear that, despite perhaps appearances to the contrary, the example is one of *description*, not explanation.

Alternatively, we might point to the fact that no sensible theory of explanation would hold that *all* explanations – including mathematical, philosophical, and religious – must contain a causal element. Indeed, some explanations surely have no ontic basis at all, having standards of adequacy which are purely cognitive. A clear case is explaining ‘why x is an A’ (using another example of Ruben’s, ‘why is a cow a

³¹ Lewis, ‘Causal Explanation’, 190-192

ruminant'³²): relevant facts about x must be given, but these can hardly be causal facts. Whilst scientific and historical explanations are not typically of the same sort as mathematical or philosophical explanations, there seems no reason to suppose that an explanation more of the form of the latter might not appear in a scientific or historical account. This seems to be a good case for recognising the 'family resemblance' nature of *explanation*.

Example (iii) can be dealt with more quickly. I accept that this demonstrates that citing causes of E is not sufficient to explain E. The causes must be described in the correct manner: a requirement which is the subject of the following section.

Explanatory concepts

In addition to more or less explicitly citing causes, historical explanations must conceive of these causes in the correct way. This is clearly demonstrated by Ruben's example, (iii), and demonstrated less clearly but in more detail by the case study historical explanations. Causal explanation, and using description, concepts or interpretation to explain, are not contradictory activities, as is sometimes claimed. (I understand 'descriptive explanation', 'conceptual explanation' and 'explanation through interpretation' – in a sense not restricted to Winch's use of the term – to be roughly equivalent notions). Indeed, not only does causal explanation *always require* the proper use of concepts in explanation, but the use of critical description in bringing understanding to the historical material *frequently* involves or implies causal claims.

My second claim to be substantiated in this section is that there is no *one* correct way to conceive of, or describe, causes in a historical explanation. Indeed, it is because there are multiple ways to provide the necessary conceptual backing in a historical explanation that we are faced with what seem to be different 'types' of historical

³² Ruben, *Explaining Explanation*, 220

explanation. In particular, I want to reject the commonly held view that causes *must* be cited so as to exhibit a nomological relation between cause and effect. Maybe this is a requirement of *scientific* explanations, which usually provide the reference class in philosophical discussions of explanation; but it is not a universal requirement for historical explanation. Whilst there are multiple means to properly conceptualise causal claims in history, I will make some general comment as to the nature of these historical concepts. I take this task to be that which is necessary to tackle the question, raised in chapter three, concerning how we are to demarcate *explanatory* modes of understanding.

My third task, in the final section, will be to comment on the implications for explanatory exclusion of the dual understanding of historical explanation recommended here. Whilst exclusion resulting from claims of explanatory relevance relies on the causal element of an explanation, it would be wrong to think that the conceptual type of explanation has no effect on judgements of exclusion.

In order to substantiate my claim that causal and conceptual explanation are complementary, rather than necessarily alternative, activities, we need to examine *descriptive* or *conceptual* explanation in more detail. Explanation may sometimes be achieved simply by re-describing the item in question. Under what conditions is re-description successfully explanatory? Davidson³³, Hempel³⁴, Ruben³⁵ and others have pointed to the importance of laws in explanation. A re-description is explanatory where that re-description demonstrates a nomological connection between explanans and explananda. Explanation through re-description may be considered an elliptical version of D-N explanation, where the law is suggested rather than explicitly formulated. For example:

‘Why did that saucepan break? Because it was made of glass’

³³ Davidson, ‘Actions, Reasons and Causes’

³⁴ Hempel, ‘The Function of General Laws in History’; in particular 40-41

³⁵ See, for example, David-Hillel Ruben, ‘Singular Explanation and the Social Sciences’ in *Philosophy* 65 (1990): 95-117

Whilst no law is mentioned explicitly, the re-description implies a plausible, if approximate law: that most objects made of glass break upon contact with a hard surface. What appears to be a *conceptual* or *descriptive* explanation is really nothing distinctively of the sort. If this theory of descriptive explanation is maintained universally, then any alleged descriptive explanation which did not imply a nomological connection would be on that basis not really an explanation at all. Such candidates ('the peoples of Europe rapidly migrated as stars in an exploding supernova') might be considered rather as metaphors; perhaps useful for understanding, but not truly explanatory.

But there are counter-examples to the nomological model. Useful examples can be derived from Walsh's work on *colligation*: such as that Hitler's reoccupation of the Rhineland in 1936 might be explained by locating it within a more general policy of German expansion pursued from 1933 onwards³⁶. The salient feature of this example is that the German expansion is a *singular*, though complex and extended, event. The explanation does not, therefore, seem to be even tacitly nomological. Rather, the example appears to be explanatory in the sense that it unifies the explanandum within a wider particular.

Dray³⁷ provides a similar argument against a nomological analysis of descriptive explanation. Dray's analysis of the explanatory value of 'explaining what' is that it is *summative*: 'it allows us to refer to x, y and z collectively as "a so-and-so"³⁸. In defending descriptive explanation as being distinct from D-N explanation, Dray emphasises the fact that the former had an entirely different purpose to the latter. Descriptive explanations are answers to "what" questions; Deductive-Nomological explanations are answers to "why" questions.

Walsh's notion of *colligation* leads to a similar analysis. Walsh developed the idea of colligation through such examples as the aforementioned 'Hitler's occupation of the Rhineland'. He emphasises the way that historians frequently explain by ordering individual events into wider processes. Colligatory wholes are complex particulars:

³⁶ W. H. Walsh, *An Introduction to the Philosophy of History* (London: Hutchinson, 1951), 59

³⁷ W. H. Dray, "Explaining What" in History' in *Theories of History*, ed. Patrick Gardiner (New York: Free Press, 1959): 403-408

³⁸ *Ibid.*, 406

like ‘the policy of German expansion’, ‘the Dark Ages’, or ‘the French Revolution’, they are intelligible wholes whose comprehensibility does not require multiple instantiation. In his later work on the subject, Walsh made it clear that explanation through colligation has a similar function to Dray’s notion of ‘explaining what’: ‘Dray and I are concerned ... more with *interpretation* than with explanation’³⁹. The two activities should be seen as separate, and complementary.

I agree with Walsh and Dray that descriptive explanation may be justified even where no nomological claim is implicit. However, I disagree with their claim that descriptive explanation is a radically different sort of activity to causal-scientific explanation. Whilst it would be implausible to maintain that all description and use of concepts in historical writing implies causal claims, the case study accounts do demonstrate a striking interplay between descriptive and causal aspects.

One clear example is provided by Cobban, in whose account the overall conceptual interpretation shapes our understanding of the causes and effects of the Revolution. The critical description of the divisions in pre-Revolutionary French society leads to the idea that peasant resentment of the rich, and the growth of capitalism, brought about the revolution. The same overall description provides an understanding of the outcome of the revolution: the fact that, in many ways, it was more obstructive of capitalism than supportive. Cobban’s critical description serves to *summarise and unify* causal claims, relating to events both before and after the revolution. A primary use of descriptive explanation is, as Dray and Walsh suggest, the unifying effect of such a description. Yet what is unified are frequently causal claims themselves. In this typical case, causal and descriptive explanation are two, complementary, aspects of historical explanation.

A second example is provided by Furet’s explanation. His key concepts, ‘state’ and ‘society’, allow an understanding of why political life altered so rapidly and so fundamentally during the revolution. The concepts provide the basis for the implicit *nomological* claim that ‘alterations of the power balance between state and society tend to lead to extreme political upheaval’. But these key concepts also *summarise*

³⁹ W. H. Walsh, ‘Colligatory Concepts in History’ (originally 1967) in *The Philosophy of History*, ed. Patrick Gardiner (Oxford: Oxford University Press, 1974), 136

such causal claims as that the state's suppression of the traditional representative bodies of society (the *parlements*) came to profoundly influence the nature of revolutionary politics. Furet's interpretation is intimately related to implicit causal claims, and in part depends upon the plausibility of these claims.

I now intend to make good my second claim: that there are multiple ways to 'correctly' conceptualise causal claims in historical explanation. The following list is not intended to be exhaustive, though does seem to adequately explicate the differing nature of the eight case study explanations. It is by no means the case that a single historical account always utilises a single type of conceptualisation, though there is often a general tendency in a single account to use one or another. Indeed, it is possible for a single causal claim to be conceived in such a way that it may fit more than one of the following types.

One way of conceptualising causes is to present cause and effect so that they exhibit a *nomological* relationship. This is what Lefebvre does: the cause 'the bourgeoisie had economic but not legal-political power' is, if explanatory, made so because there is a plausible, empirical law which connects this cause to the explanandum, the French Revolution. A variation of showing the explanans to connect nomologically is to present cause and effect in such a way that the terminology explicitly *subsumes other cases*. This is Skocpol's primary aim in conceptualising the causes of the French Revolution. Notice that an explanation may be fully nomological, without it thereby applying to any other cases. Perhaps the particular combination of laws operative in bringing about the French Revolution simply has not been instantiated in the case of any other historical event. Wider applicability is plausibly a virtue of laws, but is not necessary for a nomological approach.

A second way to present causes is by showing there to be a *rational* link between cause and effect. In the terminology developed previously, such an explanation would be an 'internal rational' explanation. Michelet attempts to explain the events of the French Revolution in such a way as to show that the Revolutionaries were justified in behaving as they did. Schama, on the other hand, disagrees: he takes it as a virtue of his explanation that he does not cite a rational link between the causes of

Revolution and the Revolutionary events. The Revolutionary origins are, for Schama, haphazard, chaotic, and frequently accidental. A variation on the conceptualisation of action as being rational is to conceptualise the action as being *morally justified*; an approach attempted by both Michelet and Lefebvre (though explicitly renounced by Schama and Cobban).

Where it is considered that the *only* way to explain in history is to demonstrate the internally rational connection between actions, the approach comes into line with the Wittgenstinian-Winchian theory of social explanation previously introduced. Of course, one of the key tenets of that theory is that such demonstration of rationality cannot be the same activity as citing causes. My model of historical explanation depends on a rejection of this claim. We should recognise that, of course, not all (Winchian) interpretative *description* involves causes. The example given in chapter three demonstrates a non-causal interpretation: ‘why did that person exchange a piece of paper for some money?’, answered by describing what ‘cheque’ means in our society, and in particular, the rules governing how cheques are treated by banks. Yet where the explanation is in terms of *reasons*, causes are presupposed. I have no new argument for this claim, but would be inclined to follow Davidson:

‘some philosophers have concluded that the concept of cause that applies elsewhere cannot apply to the relation between reasons and actions ... But suppose we grant that reasons alone justify actions in the course of explaining them; it does not follow that the explanation is not also – and necessarily – causal. ... it is necessary to decide what is being included under justification. It could be taken to cover only ... that the agent have certain beliefs and attitudes in the light of which the action is reasonable. But then something essential has certainly been left out, for a person can have a reason for an action, and perform the action, and yet this reason not be the reason why he did it. Central to the relation between a reasons explanation and an action it explains is the idea that the agent performed the action *because* he had the reason.’⁴⁰

There have been some who have regarded rational conceptualisation of cause as simply a special case of nomological conceptualisation in general. The rough idea would be that to describe a cause in the terminology of ‘reasons’ is to claim that the agent had a belief B and a desire D, and that sometimes/ often/ always when agents

⁴⁰ Davidson, ‘Actions, Reasons and Causes’, 9

have B and D they perform action A. Again, I am inclined to follow Davidson's assessment of the proposed reduction⁴¹. Whilst it is possible to formulate a sort of law which is equivalent to a claim that an agent acted for a reason, this law will be of a particularly weak sort; for very often agents have B and D and yet do not perform A. (As Davidson says: 'what is the ratio of actual adulteries to the adulteries which the Bible says are committed in the heart?'⁴²) Furthermore, explanations in terms of reasons seem far more convincing than an explanation in terms of the equivalent, very weak, law.

A third means of conceptualisation is more contentious among historians; that of demonstrating a *teleological* connection between cause and effect. Hegel represents this approach most fully: his whole approach to historical explanation is centred around the need to 'discover' teleology. Michelet advances a more cautious teleological approach, based on the natural pull of 'national unity'. A 'strong' teleological approach, by which I mean one which establishes the teleological law non-empirically, is intimately linked to a claim of 'external rationality'. For Hegel, history must develop in broadly the way that it has, because the end point is the only truly rational way to think, behave, and organise our societies. Of course, such a claim of external rationality has fallen very much out of favour amongst historians since Hegel's time. Many historians now frame their explanations with the explicit aim of discouraging lingering notions of teleology in relation to the French Revolution; amongst them, Furet and Schama. A weaker teleological approach, where the teleological endpoint is established empirically, draws near to the nomological to the extent that the endpoint can be predicted according to laws.

It may be thought that I need to say more to defend my claim that teleological explanation is comprehensible within my general understanding of historical explanation. Do I need to consider the extensive literature questioning the possibility of reducing intentional or teleological explanation to causal explanation? I think not. Suppose that intentional or teleological explanation is not able to be reduced to causal explanation in the manner suggested towards the start of the chapter. The difficulty

⁴¹ Davidson, 'Hempel on Explaining Actions', in Davidson, *Essays on Actions and Events*, particularly 264

⁴² *Ibid.*, 264

for proposed reductions was that an intention might cause an action, but do so in the ‘wrong’ way – i.e., accidentally. And it was (it is claimed) not possible to explicate what is the ‘right’ way of causally intending something, without making use of the term being analysed: ‘intention’. But this conclusion is quite compatible with my understanding of explanation as causal *and* conceptual: it could even be argued that such a conclusion is to be expected. What is not suggested by a defence of the irreducibility of intentional or teleological explanation is that the latter types of explanation do not refer to the underlying causal structure, *in some way*. Such a defence does have the consequence that intentional or teleological explanation can’t be analysed using only causal notions; but this is a consequence which I can be content with.

Finally, causes are presented in such a way as to show that they are, in some sense, *unified*. Clear examples of such a means of conceptualisation have been provided in the context of colligatory explanation: describing causes such that they are seen to be part of a unified process is a common historical mode of explanation. Explanations which conceptually unify demonstrate the explanatory virtue of simplicity; at the least, they allow helpful summaries of why the French Revolution took place. And if no summary *at all* is possible, to what extent is the account really explanatory? Most historical accounts therefore unify to some extent, but a particularly clear example can be found in Furet’s explanation. However, certain writers would deny that simplicity is particularly a virtue to be admired; Schama, for example, wants to chart the multiplicity of revolutions constituting what we think of as *the* French Revolution.

I wish to emphasise again that these means of conceptualisation may be combined within a single historical account, and even within a single element of that account. In particular, the explanations classified initially as ‘narrative’ rely on a multiplicity of types of conceptualisation; Michelet’s narrative, for example, is to some extent teleological, rational, and unifying. This lack of distinction between the different modes of conceptualisation demonstrates that historical explanation differs by degree, not by kind.

As I have suggested, the above list will not be exhaustive. One possible way of conceptualising causes might be to present a causal chain as *meaningful*; an analysis which is suggested by, for example, Hayden White's philosophical approach. The reason I have not included 'meaning' as a conceptual category is simply that, without further specification, the term is somewhat vacuous. Indeed, explicating what is meant by 'meaning' in this sense is likely to lead to one of the conceptualisations already suggested: that the causal chain is *rational*, or *teleological*.

Further types of conceptualisation may be more appropriate for non-historical explanations. One example is to demonstrate that the causal chain is *familiar*. In the context of my general analysis of explanation, we should regard the virtue of 'making the explananda familiar' as one possible virtue of conceptualisation, but not the basis of a universal *analysis* of explanation. Lipton, for example, notes that 'we often explain familiar phenomena ... surprise is often a precursor to the search for explanation, but it is not the only motivation.'⁴³ But this observation raises no difficulties if we accept that presenting a *familiar* account is only one explanatory virtue amongst others.

Explanatory type, exclusion and competition

I have defended a conception of historical explanation such that explanations cite a relevant ontic relation (paradigmatically, a causal relation), which is conceptualised in a certain way (nomologically, rationally, teleologically amongst others). What remains is to tie this analysis clearly to explanatory *exclusion* and *competition*.

Explanatory exclusion may result from incompatibility of explananda, explanans or claims of relevance. To understand exclusion due to relevance claims (PEE₃), we need to be able to claim that different types of historical explanation understand 'relevance' in broadly the same way. It was for this reason that I have argued that even different 'types' of historical explanation are causal, in a minimal sense. All

⁴³ Peter Lipton, *Inference to the Best Explanation* (London: Routledge, 1991), 29

historical explanation aims to pick out certain elements of the causal history of the explanandum, though the *way* that these elements are conceived varies widely. It is, therefore, the causal, rather than the conceptual, element of historical explanations which it is necessary to attend to in judging cases of exclusion over relevance claims.

Given that the ‘type’ of explanation is determined by the conceptual mode of the historical explanation, is it therefore the case that the *type* of explanation is of no interest when judging explanatory exclusion? No. The type of explanation, or mode of conceptualisation, often affects judgements of exclusion (specifically, PEE₃) in two ways.

Firstly, different types of explanation often lead the explanans to refer to correspondingly different aspects of the causal (in general, ontic) history. Most obviously, rational explanation focuses on *reasons* – psychological causes. Often, nomological explanation leads a historian to focus on *structural* elements – social, economic, political or geopolitical. If different elements of the same history are focused upon, contradiction between relevance claims are less likely: so long as there is no implicit claim that those elements are all there really is to the history.

Secondly, different types of explanation often lead to correspondingly different sorts of target. In particular, the *generality* of the target is liable to vary according to explanatory type. Nomological and teleological explanations tend to lead to broader targets; a tendency clearly evident in Lefebvre’s and Hegel’s explanations. Narrative explanations, eschewing abstraction, tend to focus on a very specific explanatory target. In chapter five, I demonstrated how the notion of contrast space allowed us to understand this phenomenon of explanatory practice, and how similarity of target related to explanatory exclusion.

Where explanations are not exclusive, they may still be in competition. (And I have noted that explanations may not be in competition, but may be exclusive.) Whether or not explanations are to be regarded as ‘in competition’ is not so amenable to a precise explication as ‘exclusive explanation’ has been shown to be. The phrase is somewhat vaguely applied. However, I have suggested that explanations are likely to be regarded as ‘in competition’ when there is a similarity of explanatory target. In this

case, explanations compete according to the relative possession of explanatory *virtues*. These virtues may be of a familiar kind⁴⁴, but it is advantageous that they can be seen to be connected to the previous claims which I have made concerning historical explanation. They are of three kinds.

First, there are those which relate to degree of confirmation: *confirmatory virtues*. In particular, the explanation is good to the extent to which it can be argued that the explanans are (were) true. In historical practice, this type of argument is common; utilising historical evidence to demonstrate the likely truth of one's explanans (or the improbability of one's opponents'). Second, there are those which I will argue are *ontological virtues*; demonstrating that an explanation corresponds well to a way the world actually is. (This is where 'importance' fits in, as I will argue in Part III.) The difference between these two types can be seen in regard to the demand for causally relevant explanans: those explanans can be shown to be more or less *likely* to be causally relevant, and in addition can be shown to be actually more or less causally relevant.

Third, there are the virtues which, I have argued, are not universally applicable as the first two types are. These virtues are connected to the different modes of description, outlined previously; for instance, in demonstrating a rational or nomological connection. If a historian attempts to give causes in a nomological manner, for example, then their explanation is virtuous to the extent that they are successful in doing this. I shall call the virtues corresponding to the modes of conception *cognitive virtues*. This third type of explanatory virtue is perhaps the most mysterious; the nature of these virtues, and the effect this has on explanatory competition, requires further attention. Though this is not the place for a detailed examination of these points, I believe it is appropriate to sketch a possible response. A more thorough examination would have to tackle the issue of how, and to what extent, historical concepts were *empirically justified*. This is an issue which I will consider in detail with relation to the specific virtue of *citing 'important' causes*, in Part III, but not for explanatory virtues in general.

⁴⁴ For example: Martin, *The Past Within Us*, chapter three. Like me, Martin attempts to develop a list of explanatory virtues from a prior empirical examination of historical accounts

It might be thought that explanatory ‘cognitive’ virtues cannot be justified by the historical record, and are therefore incommensurable. Kuhn suggested three distinct ways for theories to be incommensurable. Alongside incommensurability of *observation* and incommensurability of *concept* is the relevant sense for our purposes: incommensurability of *methodology*. In Kuhn’s theory, this type of incommensurability results from the fact that the standards used to judge scientific theories (that they be, for example, simple, accurate and comprehensive) are themselves theory dependent. This is so for two reasons. Firstly, the background paradigm determines the *interpretation* of those standards; whether or not a theory in question can be considered ‘simple’, for example. Secondly, the paradigm governs the *relative weighting* of one virtue against another: whether, in a particular case, theory A’s advantage of being more simple will count for more than theory B’s advantage of being more accurate. In general, the thesis of methodological incommensurability claims that there is no neutral way to judge and compare different theoretical approaches.

It is likely that arguments similar to Kuhn’s would apply equally to historical explanatory virtues; to the cognitive virtues, at any rate. Indeed, it seems that historical debate is in a still more perilous position than Kuhn’s scientific debate. For not only do historians disagree over relative weighting and interpretation of virtues: unlike scientists, it is common to even disagree over what the virtues are. As we have seen, the fact that an explanation is, say, *rational* may be a virtue to one historian, of no consequence to another, and a reason for suspicion to a third.

A conception of explanatory (at least the cognitive) virtues as incommensurable would naturally lead to a relativistic attitude to explanatory competition. Hayden White well exemplifies such a position:

‘These correlations of the tropological strategies of prefiguration with the various modes of explanation used by historians in their works have provided me with a way of characterizing the styles of given historians. And they have permitted me to view the various debates over how history ought to be written, which occurred throughout the nineteenth century, as essentially matters of stylistic variation ...’⁴⁵

⁴⁵ White, *Metahistory*, 427

I believe that we ought to avoid the idea that the explanatory virtues are purely of *aesthetic* or *stylistic* import. To do this, we have to consider what else the cognitive virtues could be based upon, other than simple intellectual need. One way of making clear the idea that explanatory virtues somehow correspond to an underlying reality is to tie all explanatory virtues to causal inference. An explanatory virtue, *if* it really is, is so because it reports something true about the causal history of the target.

Arguments of this sort can be found in Barnes' and Rappaport's (previously cited) papers. However, given that I have already made it clear that my understanding of historical explanation requires *both* causal elements *and* further 'cognitive' or 'conceptual' elements, it will not be surprising that I wish to resist the claim that all explanatory virtues can be reduced to the causal.

Perhaps all we can say of cognitive explanatory virtues in general is that they provide *information about what brought the explananda about*. 'Information' is to be understood in three possible ways. In part, the way that causes are conceptualised in (hence, the cognitive virtues of) a historical explanation leads to empirically justifiable information about the causal history, as Barnes, in particular, argues⁴⁶. Most clearly, this includes information about the degree to which causes are unified. Such a claim is justified by similar information to straightforward causal claims.

Secondly, cognitive virtues link the causal history in question with wider causal histories. Presenting a nomologically supported causal relation groups events under types in a certain way. The connection specified according to the law is not causal, but neither is it purely aesthetic: it is primarily a relation of *similarity*. Finally, cognitive virtues can simply present the causal history in an interesting way. In particular, cause and effect may be conceived as providing a rational justification under a particular conception, but not under another. It is frequently of interest to present the causal chain in this way, where it is possible.

⁴⁶ Barnes, 256-269

Summary of Part II: Analysis, exclusion, competition

I present here the conclusions of Part II in a stark form, without the caveats and subtleties required for a more complete understanding of explanatory exclusion.

A historical explanation is aimed at a contrastive *target*. In the case of the historical accounts considered here, that target is loosely ‘the French Revolution’. More precisely, the target is constituted according to those possibilities that the explanation should exclude – the contrast space. The nature of that contrast space determines which aspect(s) of the target event is relevant, which aspect(s) of the event’s history is relevant, and what level of detail is required. As a general rule, the contrasts should be similar to the target event.

As a minimum, the *explanans* should meet four criteria. (i) They should cite facts which are true. (ii) They should cite facts which are indeed explanatorily relevant to the target. (iii) The facts should not only be relevant, but should locate differences between the target fact and the contrast cases: there should be no correspondingly relevant fact in the history of the contrast cases. If we understand ‘cause’ as ‘an empirically antecedent necessary, in the circumstances, for the explanandum phenomena’, then we can replace ‘explanatorily relevant’ in (ii) and (iii) with ‘causally relevant’. Finally, (iv) the causally relevant facts should be conceived of, or described, in an appropriate way. Such ways include demonstrating nomological, rational or teleological connection, or demonstrating the unity of various causal elements.

Seemingly ‘*different types*’ of explanation follow primarily from two variables in the above conditions. Firstly, explanations may differ in the type of targets presumed. They may be broad (Hegel) or narrow (Schama); they may be particular (Schama, Chartier, Cobban) or generalised beyond the particular (Skocpol). Secondly, explanations may differ widely in the way that causes are conceived, such that the explanation is, for example, more or less nomological, more or less teleological, and more or less simple or unified.

The principles of *explanatory exclusion* follow from the basic conditions of adequacy for historical explanations. Explanations strictly exclude one another where elements of those explanations are in contradiction. In particular: (a) The targets are inconsistent: such that either targets disagree over the truth of the respective target facts, or disagree over the falsity of the respective contrast cases. (b) The truths of the explanans are inconsistent. (c) The relevance claims (causal claims) of the explanans are inconsistent. A particular way for (c) to be realised is to demonstrate that the relevance claims of two explanations cannot be consistent because they are complete and independent explanations of the same target. ‘Complete’ must be understood in the context of a contrastive theory of explanation.

Actual historical practice, and ordinary intuition, dictates that we find a space for *partial exclusion*, in addition to strict exclusion. Clearly, explanations may contradict in any of the above respects to a greater or lesser extent. In addition, an important class of cases of partial exclusion is in the context of the aforementioned particular way for (c) to be realised; targets may be overlapping, rather than the same. This can be understood geometrically, according to the idea that contrast spaces may overlap to a greater or lesser extent.

Competition between historical explanations is broader than simply exclusion, or even partial exclusion. In other words, two explanations may compete, and yet not exclude. Like the principles of exclusion, principles of competition can be seen to follow from the basic conditions of explanatory adequacy. Corresponding to (i), an explanation is better to the extent that it is likely that the explanans are true. Corresponding to (ii), an explanation is better to the extent that it is likely that the explanans are relevant. Both these virtues are epistemological. In addition, (ii) leads to the idea that an explanation is better the more relevant those explanans are: a seemingly ontological criterion, which will be the primary focus of Part III. Corresponding to (iii), an explanation is better to the extent that the causes cited have less correspondence to those in the history of the contrast cases.

Finally, corresponding to (iv), an explanation is better to the extent that the causes are conceptualised in the correct manner. This latter criterion is, however, likely to be quite different to the remainder. This is for the reason that explanations differ over

what 'correct conceptualisation' means, whilst they do not differ in substance over what 'truth' or 'relevance' of explanans is taken to mean.

Part III

Chapter Seven: Importance

In the final part I will examine a feature of historical explanations which plays an essential role in explanatory competition, as well as being of interest in its own right: the notion of *explanatory importance*. I take it that citing important explanatory features is a virtue of explanations. My aims in this chapter will be four-fold. First, I will argue that ‘importance’ is a central virtue. Second, I will show that the adoption of a realist understanding (in a sense to be specified) of importance opens up a significant space for genuine competition between historical accounts, over and above providing more or less true explanations. Third, I will tackle various arguments against such a realist understanding, to show that there is no reason in principle why we should not understand historical importance in this way. Further, I will show that such an understanding makes sense of existing historical argument and assertion. In chapters eight and nine I will develop and defend a realist model of historical importance in detail.

The concept of importance

The notion of *importance* is a vital one for historical scholarship. As will be demonstrated towards the end of this chapter, historians tend to do more than just describe events and causal relations: in addition they make judgements of relative importance. ‘Importance’ is a historical explanatory virtue in two related senses: an

explanation is better to the extent that it attributes historical events with the *correct* level of importance; secondly, an explanation is better to the extent that it mentions important events, and omits the unimportant.

The locutions vary; I shall treat the following as fulfilling similar roles: ‘salient cause’, ‘deeper cause’, ‘causes’ when contrasted with ‘conditions’, ‘*the* cause’, and ‘historical events’ when contrasted with ‘non-historical’ events. (The use of some of these phrases will become clear in the context of the arguments which follow.)

An alternative technique in ascribing importance to events or causes is to rely on pragmatic emphasis; related more to the second understanding of ‘importance as virtue’, cited above. In this way, the historian tacitly demonstrates their judgements of importance through consistently emphasising event (or type of event) a, whilst ignoring event (or type of event) b. It will not always be obvious when such importance claims should be read into a historical account. Simply mentioning ‘x’ or that ‘x caused e’ *can* imply no disagreement with other accounts failing to mention ‘x’ or that ‘x caused e’, and which instead mention other facts. As I suggested in chapter four (in the context of the icy road and the tiredness of the driver leading to the crash), it is a commonplace that accounts can differ in what they mention without thereby contradicting. On the other hand, in many cases it is plausible to interpret an account which consistently emphasises a certain (type of) fact as thereby judging that fact to be important. An explicit affirmation of importance is not always necessary.

In general it will surely be a delicate issue to discern where emphasis implies judgements of importance. Yet if we regard it possible that judgements of importance may be implicit, this has the advantage of making it possible to understand those philosophical debates relating to *selection* and *interpretation* in terms of ‘importance’. Descriptions and interpretations are, sometimes rightly, criticised not only for what they say but for what they don’t say. Why is this? Not because of omission *per se*; for no account can mention everything, even where the topic is limited. A plausible answer is that criticism is justified where the description or interpretation has omitted something *important*.

Of course not all historical selection is grounded in judgements of importance; much will be based simply on demands resulting from the interests of and relevance to the intended audience, the need to provide examples of general historical claims, and the desire for anecdotes and interesting detail in the historical account. Yet, given that *some* selection decisions appear to be capable of being defended in more robust ways (by which I mean in ways which are not explicitly relative to interest), it is clear that there may be a role for using considerations of importance to ground choices of selection. Clearly, the question of what *importance* amounts to in historical scholarship thus has a bearing on the much discussed question of historical selection. The question which is the focus of the remainder of the thesis can be put in this way: is it possible that the notion of ‘importance’ be the basis of an objective element of selection decisions?

The notion of *objectivity* in historical accounts is also related to this issue. A realist notion of importance would go a long way towards defusing sceptical arguments against objectivity in history, since many such arguments proceed from the problem of selection. (Though *no* notion of importance would entirely answer worries over objectivity in history: at the very least, the question of whether historical facts can be evaluatively neutral would still remain.) To provide one example: as part of his extended historical survey of the notion of objectivity in the American historical profession, Novick frequently cites concerns over selection. Commenting on the influential historian Charles Beard, Novick writes of the impossibility of having ‘neutral criteria for selecting among the multitudes of facts, or interpreting them, for which one needed an “a priori”, and at least tacitly evaluative, frame of reference’¹.

A realist understanding of ‘importance’ – one which would allow historians to argue fruitfully and with the possibility of reaching a determinate conclusion – would provide a criterion to judge histories, beyond that of simply getting the facts right. For getting the facts right is clearly not sufficient for the production of objective histories; and it has been convincingly argued that it is not even necessary. I believe that Gorman demonstrates this feature of historical writing very succinctly, by juxtaposing three accounts of the life of Lord Haw Haw. The first two are factually

¹ Peter Novick, *That Noble Dream: The “Objectivity Question” and the American Historical Profession* (Cambridge: Cambridge University Press, 1988), 254

correct, yet clearly biased; the third is entirely constituted by inaccurate propositions, yet provides a more ‘acceptable’, or objective story².

The final concept I will mention whose fate is tied to that of ‘importance’ is historical *representation*. A popular argument, given by, for example, Hayden White, is that historical representations are radically underdetermined by the historical evidence. *Even if* the facts are agreed upon (and, we could add, even if the causal relationships are agreed upon), there are many legitimate ways to historically represent those facts. This is so because there are many ways to align these facts in relation to one another; in particular, through exclusion, and through privileging.

On the former, White states that ‘our *explanations* of historical structures and processes are thus determined more by what we leave out of our representations than by what we put in’³. The latter concerns ‘the privileged status given to certain events or sets of events in the series by which they are endowed with explanatory force, either as causes explaining the structure of the whole series or as symbols of the plot structure of the series considered as a story of a specific kind’⁴. If *importance* is a property of events which is determined in a non-pragmatic manner, then such alignments (exclusion and privileging) are restricted; insofar as importance judgements can often form the basis for such alignments.

I hope I have shown that the question of what ‘importance’ amounts to in historical accounts is, in general, absolutely vital. Yet the concept also relates to the overall project of this thesis: to make sense of explanatory competition between historical accounts. The understanding we have of ‘importance’ will affect our understanding of the relationship between historical accounts, in the following ways.

First, consider the consequences of an *irrealist* understanding of ‘importance’. By *irrealist* I mean the thesis that ‘importance’ is a property of entity x solely in virtue of the fact that an observer (or observers) regard(s) x as being important. Such an

² J. L. Gorman, ‘Objectivity and Truth in History’, in *History and Theory*, eds. Fay, Pomper, Vann (Oxford: Blackwell Publishers, 1998), 326-327

³ White, *Metahistory*, 24

⁴ *Ibid.*, 25

analysis of ‘importance’ will emphasise the indexical nature of the term: ‘for an event to be important means (only) that I’m interested in it’. The irrealist attitude might be more precisely specified as a ‘projectivist’ attitude to importance as a property: the view that ‘beauty (or colour, or cause, or virtue) is in the eye of the beholder’⁵. One of the attractions of projectivism is its metaphysical economy: we need no metaphysically mysterious properties to ground the problematic concept, only the idea that our features of our mental nature (pleasure, the desire for regularity) are ‘projected’ upon the world. We must examine whether the realist conception of importance can deliver an equally palatable metaphysical account: this will be a topic covered in chapter eight.

If the irrealist view is correct, then differing importance-claims won’t restrict the consistency of the historical explanations containing those claims. This is so because the differing importance claims do not really assign different properties to the same entity; even though their language may make it *seem* that they do.

Second, consider the consequences of a *realist* understanding. By realist I mean a conception which allows the possibility that an entity’s importance be judged independently of a particular point of view or interest. If this were the case, it would be possible that histories be better or worse with regard to importance-claims, in that such claims could correspond more or less well to the *real* importance of the event.

My claims will be that judgements of importance *can* lead to substantive disagreement between historical accounts, and that they are a legitimate criterion by which to compare historical accounts as better or worse. The precise manner by which this happens will emerge over the course of this chapter and the next; but it is necessary to give a preliminary sketch of the realist position at this point. Danto provides a good starting point; whilst accepting that there are many possible meanings of ‘importance’, he maintains that *historical* importance amounts to tracing the consequences of the effect: ‘to ask for the significance of an event, in the *historical*

⁵ Though, it should be noted that in the field of meta-ethics the alleged inconsistency of projectivism and realism has been staunchly challenged.

sense of the term, is to ask a question which can be answered only in the context of a *story* ... [or] with what different set of later events it may be connected'⁶.

I take Danto to mean that events are important to the extent that they are connected with later events. This notion of importance is best termed *consequentialist*, and is 'realist' to the extent that 'degree of consequence' can be understood in a realist manner. In addition it raises the possibility, which I will argue for, that the term is at least partially *empirically decidable*. In measuring importance by effects, we replace the more amorphous general concept of importance with the somewhat tighter concept of *causal importance*. The central problem for the realist is also thereby defined: *how* are we to use effects to measure importance?

This question will be tackled directly in chapters eight and nine, where I will present a consequentialist model of causal importance in detail. The central feature of this model will be that a cause is important to the extent that, had that cause not taken place, the target-phenomenon would have been more different: the analysis is therefore *counterfactual*. What I wish to examine first, however, are arguments directed not at the detail of any particular realist proposal, but against the very idea that importance can be objectively equated with consequences. These arguments will also serve to set the challenges that the realist model should meet.

Evaluative importance

It is common to maintain that all selection, and judgements of importance, is relative to the historians' values. Notoriously, Weber attempted to argue that objectivity in the social sciences is possible, by recommending a clear separation of factual and evaluative judgements. Yet, if the 'evaluative irrealist' is correct, importance statements are a class of judgements which are inevitably bound up with value

⁶ Arthur C. Danto, *Analytical Philosophy of History* (Cambridge: Cambridge University Press, 1965), 11

judgements. Therefore, presuming that questions of value cannot be settled objectively, disagreement over selection and judgement of importance is likewise not objectively decidable.

The American historian Charles Beard has troubled historians to the present day with his argument that all historical selection is necessarily evaluative, and hence subjective:

‘Into the selection of topics, the choice and arrangement of materials, the specific historian’s “me” will enter.’⁷

‘we clarify the mind by admitting its cultural interests and patterns – interests and patterns that will control, or intrude upon, the selection and organisation of historical materials.’⁸

More recently (in the 1950s), Walsh has agreed that judgements of importance are, at root, evaluative judgements; and Dray has argued that many, if not all, historical debates concerning causal weighting are tacitly evaluative⁹. The general point is easiest made through the provision of examples. Walsh notes that

‘it is not so very long since history books were filled with the doings of kings and queens, warring nobles and turbulent priests; they tended to concentrate on political and military happenings. Since Marx, or rather since the later years of the nineteenth century, the emphasis has shifted to economic and social history, and the main *dramatis personae* are no longer political figures but, for example, scientists and inventors, whilst the place of the successful monarch as hero of the story has been taken by the common people.’¹⁰

Political and military history had given way to economic and social history; history from above has been replaced by history from below. This seems true; the philosophical claim is that these changes reflect the changing value judgements of society in general, and hence of historians.

⁷ Charles Beard, ‘That Noble Dream’, in *The Varieties of History*, ed. Fritz Stern (Cleveland: Meridan Books, 1956), 324

⁸ *Ibid.*, 328

⁹ For example, Dray’s analysis of the debate between AJP Taylor and Trevor-Roper over the importance of Hitler’s role in bringing about the Second World War reduces to evaluative judgements. W. H. Dray, ‘Concepts of Causation in A. J. P. Taylor’s Account of the Origins of the Second World War’ in *History and Theory* 17 (1978): 149-175

¹⁰ Walsh, *An Introduction to the Philosophy of History*, 179

As with any response to an alleged counter-example, three sorts of reply are available to the realist: *denial*, *re-interpretation*, or *acceptance of plurality*. To deny the force of the counter-examples would be to argue that they are not justified examples of importance-judgements. In this case, it could be maintained that whilst importance is always properly decided by the tracing of causal consequences, historians do sometimes allow their values to intrude in an inappropriate way. As a possible example, we could claim that the attention given to social elites in pre-twentieth century historiography was simply unjustified; though we are to hope that the concentration on ‘the common people’ by today’s historian is more justifiable in consequentialist terms. Whilst this response may be appropriate in certain cases, as a general tactic I find it unappealing, since it seems motivated by a purely a priori desire to analyse importance in consequentialist terms. Any example which is shown to rest on an evaluative basis is automatically rejected as unjustified.

More plausible in general is the ‘re-interpretation’ strategy. To follow this would be to claim that the examples are justified claims of importance, but that they can be, after all, justified according a consequentialist approach. For it remains possible to explain the historiographic changes as theory-driven changes, rather than value-driven. If this is possible, then these and similar changes in historiography can be justified according to consequentialist ideas. Perhaps the effect of the common people on the culture and politics of past societies began only recently to be reassessed. Perhaps (more plausibly) advances in socio-economic theory led to a realisation of the importance of these factors. Clearly detailed study would be required to discover whether the theory-driven or the value-driven story is more empirically adequate.

Finally, the pluralist option remains: where the consequentialist accepts that there may be some genuine cases of evaluatively justified importance claims, though there are also non-evaluatively justified claims. Even if the force of Walsh’s and similar counter-examples is accepted, the realist is free to believe in a kernel of objectively decidable importance-judgements, whilst admitting that evaluative opinions affect other such judgements. This conclusion is, however, threatened by a second, more ambitious, evaluative argument of Walsh’s.

Walsh compares the kind of realist-consequentialist strategy which I have outlined to the utilitarian principle of morally judging action. As well as having doubts over the ‘vagueness of the phrases’ used in such a principle (a concern which can only be allayed by working through a realist model in detail), Walsh believes of the ‘realist’ notion of importance that

‘as a formula for choosing between judgements of value, it is not itself a simple judgement of value; but equally, and still more obviously, it is not a straightforward statement of fact. You cannot, that is to say, establish its truth by finding out how things are. To subscribe to it is, in effect, to accept a certain moral outlook, the moral outlook of the Utilitarians.’¹¹

In other words, even if it were possible to render the realist methodology precise enough to be usefully employed – which Walsh doubts – the decision to use that methodology would still be a value laden one.

Whilst the analogy with utilitarianism is intriguing, it misleads. The choice between different moral theories is clearly a moral decision, not a scientific one; no matter how ‘scientific’ are the contents of those theories. In this Walsh is correct: the nature of the choice is not determined by the content of the theories. Yet Walsh goes astray in applying this analogy. With regard to ‘importance’, he maintains that the choice of whether or not to use the consequentialist meaning is not a scientific matter; so must therefore be a moral one.

The problem is that Walsh assumes a false dichotomy: that any choice is either a judgement of value, or a statement of fact. Yet the choice between philosophical analyses is one based neither (entirely) on values, nor (entirely) on facts. The way to arrive at philosophical analyses was the topic of chapter one. I favoured a broadly descriptive approach, relying on internal criticism to arrive at a reflective equilibrium, such that the practice is thereby shown to be rational or justified.

¹¹ *Ibid.*, 182

Intrinsic importance

‘For Meyer, those events are historical which have been efficacious, i.e. have produced consequences. For example, the philosophy of Spinoza was for a long time quite without influence, but later people became interested in it and it began to influence their thought. Hence from being a non-historical fact it became a historical one: it is non-historical for the historian of the seventeenth century but becomes historical for the historian of the eighteenth. This is surely a quite arbitrary and perverse distinction. For the historian of the seventeenth century Spinoza is a highly interesting phenomenon, whether or not he was read and accepted as a leader of thought; because the formation of his philosophy was in itself a noteworthy achievement of the seventeenth-century mind. ...

He [Meyer] sees that a mere past event taken in isolation cannot be an object of historical knowledge, but he thinks that it becomes one in virtue of its connexions with other events, these connexions being conceived by him in the positivistic manner as external causal connexions. This, however, begs the question. If the historical importance of an event is defined as its efficacy in producing further events, what constitutes the historical importance of those others? For he would hardly hold that an event becomes historically important through producing consequences themselves devoid of historical importance. If, however, the historical importance of Spinoza consists in his influencing the German Romantics, wherein consists the historical importance of the German Romantics?’¹²

In this intriguing passage, Collingwood criticises a notion of importance which appears strikingly similar to the realist-consequentialist intuition outlined previously. Collingwood’s criticisms lead him to view events as important *intrinsically*, rather than in virtue of the events’ connections. Further examples seem to support the ‘intrinsic’ notion, in addition to Collingwood’s Spinoza. For Christians, the life of Christ is likely to be the most important set of events in history, and is so no matter what the consequences of those events amount to. The importance of an event, for Collingwood, derives primarily from the ideology of the historian and their contemporary society. He refers approvingly to ‘a new principle of selection based on the interest of the historian and of the present-day life of which the historian is a

¹² R. G. Collingwood, *The Idea of History* (Oxford: Oxford University Press, 1994; originally 1946), 179

representative. ... The subjective element is an essential factor in all historical knowledge.’¹³

The passage drawn from Collingwood suggests three distinct, though related, arguments against the consequentialist view of importance (excluding the general criticism of the ‘positivistic’ [sic] reliance on the reality of causal connections). First, Collingwood charges the ‘connectionist’ with circularity. Events can not be important only in virtue of bringing about other events; they must be important only in virtue of bringing about further important events. But if this is the case, we must have a notion of intrinsic importance with which to ground connectionist considerations, on pain of vicious circularity. Dray also provides the example of the American Civil War, stating that

‘the importance of the selected event – for example, the Civil War – depend[s] on the importance of something else, causally related to it – for example the legacy of bitterness in the South. For it seems odd to suggest that an event could be important by virtue of its relation to things that were not themselves important.’¹⁴

I believe the premise to be challenged is that which denies that connections alone are sufficient to define importance. Causes may be important whether or not their effects are. There is a sense in which Collingwood is correct: in the recognition that since all events have effects, this alone cannot be sufficient to discern importance. The response, which will be incorporated into my subsequent model, is that *magnitude* of effect is the standard by which importance should be measured.

Second, Collingwood points to an apparently unpalatable consequence of the ‘connectionist’ view: that events change in importance over time. Spinoza would be important to a eighteenth-century historian, but not to one of the seventeenth-century. The only connectionist response to this claim can be to simply deny its force. An event’s importance is constituted by its consequences; yet judgements of that importance may well vary dependent upon which point in the chain of effects the

¹³ *Ibid.*, 180

¹⁴ W. H. Dray, ‘On the Idea of Importance in History’ in W. H. Dray, *On History and Philosophers of History* (Leiden: E. J. Brill, 1989), 77

historian finds herself. (This simple denial also serves to demonstrate how radically intuition can differ over the appropriateness of judgements of importance: a lesson we would do well to remember in the construction of 'models' of importance.)

Third, counter-examples of the form of Collingwood's Spinoza challenge at least the universality of a realist-consequentialist analysis. The intuition is that such great thinkers are important in virtue of what they write, rather than in virtue of the effects of that writing. As in the previous section, three sorts of responses are available. The first (direct denial) reply would be to simply reject the intuition that examples like Collingwood's Spinoza are justified claims of importance. A second (pluralist) response would be to affirm the examples, but to claim that the realist-consequentialist view of importance may still be applicable in some, if not most, other situations.

A third (re-interpretation) response would be to argue that the example can, in fact, be justified consequentially. In this case, it may be argued that we need to tweak our idea of 'consequences'. It may be that not many people read Spinoza in the seventeenth century; but this need not count decisively against the 'consequences' of Spinoza's thought. Perhaps we could weight the consequentialist measure so as to favour those ideas which have a really novel, or strong, effect on a few people, rather than a mild effect on many.

More radically, we could include in the measure of consequences those effects upon the world of ideas (the 'third world' of Popper and others); in that case the number of people interested in Spinoza may not count as much as the sheer novelty of Spinoza's thought. Whilst I would be disinclined to adopt the final option on metaphysical grounds, such considerations do demonstrate the radically underdetermined nature of 'consequences'. Without further specification, the realist-consequentialist idea provides little more than the *form* of an answer to requests for importance-judgements. The model developed in chapters eight and nine needs to (and does) restrict the 'consequences' to be considered to manageable respects.

Narrative importance

Dray presents a number of further counter-examples to the idea that importance can always be picked out by considering causal consequences¹⁵. What these examples have in common is that they suggest a connection between the ideas of narrative and importance: that an understanding of the narrative structure of an account may be necessary to properly comprehend historical importance.

The first example suggests that importance may be derived from anticipation, rather than causation. Leonardo's flying machines may not have brought about later developments in flight, but are important nonetheless since they anticipate those developments. The second example suggests that important events are those which come at the beginning or ending of processes, or eras. The importance ascribed to 'turning points' such as the First World War (or, for that matter, the French Revolution) can hardly be denied.

In connection with these example, Dray also notes a potential anomaly in the causal-consequentialist idea of importance. The anomaly results from the fact that causal chains are without end, and are presumably transitive. Hence, in general, the further back in time an event, the more important it is likely to be, since the more causes that event will have. In addition, according to the casual-consequentialist understanding, in all cases where c causes event d, c will be more important than d: since all effects of d are also effects of c, yet c has additional effects (at least, d). Yet these two consequences of the consequentialist idea may well conflict with our usual understanding of historical importance. It is easy to imagine cases where later events in the same sequence are in fact regarded as more important than earlier ones.

One way to interpret the first two types of example, and to avoid the anomaly, is to suppose that history is made up of narratives (or 'processes' or similar). Examples of such narratives or processes might be 'inter-war Europe', 'the long nineteenth

¹⁵ *Ibid.*, 80-81

century’, or, for that matter, ‘the French Revolution’. The importance of elements of a narrative will depend not simply on the temporal position of those elements, as a consequentialist idea would suggest. Consider fictional narratives: we intuitively locate certain points of those narratives as ‘key’ episodes, based on an understanding of the narrative as a whole. We could then use the idea of ‘narratives’ to supplement a consequentialist understanding of importance in such a way as to bring the philosophical model into line with the intuitions implied by Dray’s comments. Whilst it will be unlikely that we are able to formulate strict rules to govern the relation between narrative structure and importance, the following suggestions provide a starting point:

(I) Those events ‘bracketing’ (at the start and end of) narratives will, in general, be weighted as more important than their consequences alone would suggest. This accounts for the seeming importance of ‘turning points’.

(II) Events would not generally be ‘credited’ with the bringing about of events in a different narrative. Between narratives there would be a ‘break’, such that events in one were not regarded as having consequences in another in the sense relative to judgements of importance. This accounts for the anomaly pointed out by Dray.

Whilst I believe that a satisfactory solution could, in principle, be derived in the ways suggested above, such a solution would seem to count against a realist analysis of importance. Presumably, narrative structures do not really exist in the world: they are imposed on a series of events according to our need to find meaning in those events¹⁶.

It is, in any case, better to answer Dray’s comments with the resources provided by the model of historical explanation already developed in this thesis. I have argued that causal *explanation* differs from simply citing causes in two ways. Firstly, causal explanations are always given with relation to a contrast space. The effect of choosing a good contrast space is to render large parts of the causal history of an

¹⁶ For a support of this assertion, see Hayden White, ‘The Historical Text as Literary Artifact’ (originally 1978) in *History and Theory* eds. Fay, Pomper, Vann: 15-33. However, for opposing arguments, see David Carr, *Time, Narrative and History* (Indiana: Indiana University Press, 1986).

event *irrelevant* for the purposes of a causal explanation. This fact means that we don't have to maintain, counter to intuition, that earlier causes are always more important than later causes; simply because, given an appropriate contrast space, the earlier causes are often not properly a part of the causal explanation at all.

Secondly, in a causal explanation causes must be conceived of in a certain way. Viewed purely causally, it can hardly be denied that a consequentialist notion implies that importance constantly increases as more effects are produced. Again, viewed purely causally, there is no place for 'breaks' in the causal sequence. However, remembering that *explanatory* causes must be conceptualised only in certain ways, importance of causal explanations may differ from purely causal importance. Though the causal chain will have no breaks, a causal-nomological one, for example, may do: the 'story' will break at the point where an appropriate law (at the historical level) cannot be cited. Likewise, if the mode of conceptualisation is the rational, then the 'story' will break at the point where there is no rational explanation to link one event to the next.

In general, a purely consequentialist idea of importance is more appropriate for comparing the importance of elements in *different causal chains*. Where we are to compare the importance of elements *within the same causal chain*, we should attend to the appropriate conceptualisation of that chain. The main point is that consequentialist importance should not be understood as a causal phenomenon *per se*, but in terms of causal *explanation*.

Pragmatic importance

Most pressing of all is the challenge to the realist conception of importance presented by a pragmatic understanding. Amongst irrealist approaches to importance, the pragmatic is a perennially popular means of specifying the general position; it is a tradition which stretches back (at least) to Mill. Mill's well known definition of 'the

real Cause' leaves no room for the idea that one cause may be more important than another. I will quote Mill at length, since the views expressed have consistently found approval with a wider audience.

'It is seldom, if ever, between a consequent and a single antecedent, that this invariable sequence subsists. It is usually between a consequent and the sum of several antecedents; the concurrence of all of them being requisite to produce, that is, to be certain of being followed by, the consequent. In such cases it is very common to single out one only of the antecedents under the denomination of Cause, calling the others merely Conditions. Thus, if a person eats of a particular dish, and dies in consequence, that is, would not have died if he had not eaten of it, people would be apt to say that eating of that dish was the cause of his death. ... The real Cause, is the whole of these antecedents; and we have, philosophically speaking, no right to give the name of cause to one of them, exclusively of the others. ... All the conditions were equally indispensable to the production of the consequent; and the statement of the cause is incomplete, unless in some shape or other we introduce them all. ... Nothing can better show the absence of any scientific ground for the distinction between the cause of a phenomenon and its conditions, than the capricious manner in which we select from among the conditions that which we choose to denominate the cause. However numerous the conditions may be, there is hardly any of them which may not, according to the purpose of our immediate discourse, obtain that nominal pre-eminence.'¹⁷

Three of Mill's ideas should be emphasised, with regard to the idea of 'importance':

- 1: The only truly objective manner of selecting causes is to list all the antecedents of a given effect, which together constitute the 'Real cause'
- 2: There can be no 'philosophical' or 'scientific' basis for selecting members of this set as 'cause' or 'condition'
- 3: Indeed, such is the 'capricious' basis for selection, that almost any member of the set constituting the 'Real cause' may be selected as 'the cause'

Of course, in practice we can't usually provide the 'Real cause': the only reason, therefore, for selecting one antecedent over another as 'the cause' can be pragmatic.

¹⁷ Mill, Book 3, Chapter 5, Section 3

A similar conception of importance and selection can be found in Hart and Honoré's influential theory of pragmatic selection¹⁸. Their theory emphasises the fact that the cause which is chosen fulfils a particular pragmatic need. In particular, *the* cause is that which is abnormal, compared to a 'normal' contrast. They make it clear that what is 'abnormal' is dependent upon the investigator's own perspective and experience. David Lewis, likewise, makes it clear that he would analyse selection of '*the* cause' as a purely pragmatic choice¹⁹. It is no surprise to find that Van Fraassen, also, believes that importance is a purely pragmatic feature:

'the salient feature picked out as 'the cause' in that complex process, is salient to a given person because of his orientation, his interests, and various other peculiarities in the way he approaches or comes to know the problem – contextual factors.'²⁰

Mill's view has been a popular choice of historians in their more reflective moments. Stephen Rigby, for example, quotes Mill with approval, and draws the conclusion that 'we, as historians, have no agreed criteria by which to distinguish 'causes' from 'conditions''²¹. Further, Rigby, like Hart and Honoré, lets all the work of defining importance be guided by the pragmatic understanding of contrastive explanation: 'which factor we select as the 'key', differentiating one in any particular situation will depend upon what we contrast that situation with'²². This interpretation seems also to fit Cobban's understanding of judgements of importance, quoted in chapter two.

A related methodological position which has had a large impact on historical writing at certain times maintains that all selection is to be avoided. The thesis is philosophically equivalent to Mill's in that it agrees that all selection is arbitrary. The difference stems from the concern with the negative effects on objective historical writing of 'capricious' selection procedures.

¹⁸ H. L. A. Hart and A. M. Honoré, 'Causal Judgement in History and in the Law', in *Philosophical Analysis and History*, ed. W. H. Dray (New York: Harper and Row, 1966)

¹⁹ Lewis, 'Causal Explanation', 183

²⁰ Van Fraassen, 125

²¹ S. H. Rigby, 'Historical Causation: Is One Thing more Important than Another?' in *History* (1995), 237

²² *Ibid.*, 238

Perhaps the most famous historian who advocated this doctrine is J. B. Bury, writing in the late nineteenth and early twentieth centuries: by ‘patient drudgery’ and ‘microscopic research’ historians should act ‘in the faith that a complete assemblage of the smallest facts of human history will tell in the end’²³. History should record the minutiae of the past, omitting nothing; even if this means that the only narratives which could, in practice, be produced cover an extremely narrow time span. Walsh, in commenting on the argument that selection cannot be realistically defended, and is therefore a ‘capricious’ business, suggests that such concerns are not just of the nineteenth century: ‘I have seen historians so embarrassed by this argument that they have taken the heroic course of maintaining that history proper is not selective at all’²⁴.

Of course, in this thesis I have myself argued that we should not deny that pragmatic features of explanation seeking often guides the selection of causes. Much of chapter five was devoted to the investigation and support of a contrastive theory of explanation, which is above all an attempt to develop a principled theory of pragmatic relevance. But the distinctive claim of the pragmatic irrealist is that this is *all* there is to selection. The pragmatist, inspired by Mill, points to the equal necessity of all causes which together make up the sufficient set (or ‘Real cause’), and claims on that basis that any selection between them is arbitrary. The consequentialist idea is precisely an attempt to locate some basis which can justify claims of greater or lesser importance, equal necessity notwithstanding.

Historical Usage

Before considering the consequentialist model in detail, it is valuable to examine the actual historical usage of ‘importance’ in the histories of the French Revolution. At

²³ Quoted by Walsh, *An Introduction to the Philosophy of History*, 171

²⁴ *Ibid.*, 177

the least, it is to be hoped that such histories demonstrate the application of some of the concepts of ‘importance’ we have outlined. However, we can also hope for more than this from a descriptive survey. We can expect both suggestions of novel situations which may challenge the ideas already introduced, and suggestions of other means to recognise and argue for importance. I will examine five of the eight historians’ explanations; a choice reflecting the interestingly different approaches to importance at work.

Lefebvre

Lefebvre chooses an ‘ultimate cause’, or ‘deeper cause’ of the Revolution (the socio-economic cause): terminology which clearly corresponds to an ascription of importance as we have understood it. The philosophical question to ask of Lefebvre’s work is how this claim can be justified.

As argued in chapter six, the explananda which is explained by Lefebvre’s deep cause is not the French Revolution as such, but a socio-economic change of some sort, of which the French Revolution is one example. The socio-economic discrepancy cited by Lefebvre may be sufficient to explain *a* socio-economic change, though it is not sufficient to explain everything about *this* socio-economic change. The other relevant causes can only be invoked to explain why particular features of the actual change were present: ‘there would have been no French Revolution – such as actually took place – if the king, “handing in his resignation,” had not convoked the Estates-General’²⁵. Thus, whilst all causes are necessary to explain the features of the Revolution, *only the ‘deeper’ cause is necessary to explain why there was a change at all.*

Is this justification equivalent to the consequentialist idea? It seems so. For the deeper cause *is* deeper precisely because without it, there would have been no social revolution at all. Without the other causes cited there would not have been the actual Revolution, but there would have been some sort of similar upheaval, presumably at a similar date. Thus the deeper cause has more of an effect. The idea that important

²⁵ Lefebvre, 2

causes are those which focus on *essential* features of the target event will receive further examination in chapter eight.

Whilst, for Lefebvre, the breakdown of central power (CP) was not as important a cause as socio-economic factors, CP is regarded as pertinent. The reason for this is the ability of CP to explain a range of salient features of the revolution. For Lefebvre, CP explains the revolution's violence, and the fact that the revolution was 'from below'. How are these salient features selected? Partly, this will be due to the focus provided by explicit or implicit contrast cases, as described in chapter five. Yet, the consequentialist approach may be relevant also to the selection of 'pertinent features'. If it is not, then 'pertinent features', and hence 'pertinent causes', will be specified purely pragmatically. Again, we shall need to examine in chapter eight the relation between 'pertinent' features of the target event, and importance of causes.

Schama

Schama's detailed narrativist approach, entailing a reluctance to summarise 'fundamental' causes of the Revolution, might suggest a quite different understanding of importance. Yet, I would argue that such a difference is more superficial than it may at first appear. I will begin by commenting on the typical structure of Schama's narrative.

Schama structures his narrative around exemplary events, which serve as *symbols* of wider themes. The selection of these events is not therefore based on the importance of the event *per se*, but on the importance of what the event can be made to represent. For example, Schama describes in some detail the launch of the first hot air balloon in the grounds of Versailles in 1783²⁶. Quite ingeniously, this one event is used to provide evidence for, and manifestation of, (at least) three major themes.

First, the crowd gathered to watch the spectacle in Versailles symbolises the impossibility of ordering by the traditional principle of rank and seniority. Second, the ballooning is a symbol for the rise of the modern scientist-citizen-hero over the

²⁶ Schama, 123-131

traditional ‘foppish, ornamental courtier’²⁷. Third, the ballooning demonstrates Rousseau’s influence over the emotional life of the French; in particular, the crowd’s mixture of delight and terror in the spectacle. Schama’s narrative moves forward by using such exemplary events as the links between different themes. Clearly, therefore, the selection and privileging of many particular episodes and examples, like the inaugural balloon flight, is not based on a consequentialist method.

However, what of the general themes, of which the particular examples represent? Whilst one of Schama’s most central beliefs is that ‘the Revolution was a much more haphazard and chaotic event and much more the product of human agency than structural conditioning’²⁸, there clearly *are* general themes in his explanation. And Schama does, at times, attempt a defence of the importance of these themes. Defending the importance of the cultural change of the rise of Romantic sensibility, Schama writes that this was

‘of more than literary importance. It meant the creation of a spoken and written manner that would become the standard voice of the Revolution, shared by both its victims and its most implacable prosecutors.’²⁹

This defence proceeds by citing significant effects of the cause in question, the cultural change. In this typical example, then, Schama does implicitly rely on a consequentialist idea in order to defend the importance of a central theme. The difference between Schama and Lefebvre lies perhaps in two respects. Firstly, the sheer number of ‘important’ themes in Schama stands in contrast to the single ‘ultimate cause’ to be found in Lefebvre. Schama shows little inclination to attempt to rank the themes he introduces; his standard is simply either that the theme is important enough to be in the account, or it is not.

Secondly, Schama emphasises the unpredictability and chaotic nature of the events of the narrative. In particular, Schama disagrees that ‘if the whole event was of this epochal significance, then the causes that generated it had necessarily to be of an

²⁷ *Ibid.*, 125

²⁸ *Ibid.*, xv

²⁹ *Ibid.*, 153

equivalent magnitude'³⁰. This belief in unpredictability may not pose a fundamental philosophical problem for the consequentialist idea, but it does raise potential epistemic problems. For Schama's assertion implies that seemingly small causes – those insignificant at the time – may still have large consequences, and may therefore turn out, in the longer run, as more important. The intuition trades on a notion of 'chaos' in its mathematical sense: the principle that minute initial differences in a system's state may lead to large subsequent differences.

Whilst this will not undermine the principles of the consequentialist method, it may make its application more difficult, since in intuitively judging importance we surely must adopt heuristic principles which are broadly linear (non-chaotic). Given the enormous range of possible candidates for 'important causes' of such a complex event as the French Revolution, it is natural, indeed inevitable, that we rely on such initial rules of thumb as assuming that those causes which appear to be 'of epochal significance' are the most important. I shall return to this issue in chapter nine.

Furet

Furet's comments sometimes suggest a historian explicitly sceptical of the possibility of weighting causes. Furet, like Schama, prefers to see the revolution's causal history as fortuitous, and heterogeneous: rather than succumb to 'the powerful attraction of single-cause explanations'³¹, we should see the revolution as

'born of the convergence of very different series of events, since an economic crisis (a complex phenomenon in itself, involving agricultural, 'industrial', meteorological and social factors) took its place alongside the political crisis that had begun in 1787'³²

However, like Schama, Furet can hardly avoid producing statements which, at least implicitly, rely on judgements of importance. I shall examine three quite different examples.

³⁰ *Ibid.*, xiv

³¹ Furet, 21

³² *Ibid.*, 24-25

As we have seen, one of Furet's central themes is that the revolution consisted of a struggle between state and society: 'the Revolution mobilized society and disarmed the State; it was an exceptional situation, which provided society with a space for development to which it does not normally have access.'³³ How does Furet defend the belief that this theme is central to an understanding of the revolution? Furet only provides one sentence as justification:

'This type of analysis has the two-fold advantage of restoring to the French Revolution its most obvious dimension, the political one, and of focusing attention on the true break in continuity it wrought between 'before' and 'after', that is, a change in the ways of legitimating and representing historical action.'³⁴

This justification depends on a prioritisation of certain features of the Revolution: the political aspect of the Revolution, and the fundamental change in political ideas of legitimation. These 'obvious' or 'true' features fulfil a similar role to Lefebvre's 'pertinent features', implying that the causes which explain these features are the most important.

The second example is of a negative claim of importance. Furet suggests that emphasis on socio-economic causes should be minimised. The Marxist socio-economic interpretation of the revolution is mistaken because it

'treats the most radically new and the most mysterious aspect of the French Revolution as no more than the normal result of circumstances', yet 'neither capitalism nor the bourgeoisie needed revolutions to appear in and dominate the history of the major European nations in the nineteenth century'³⁵

The implicit argument is comparative: the socio-economic change (SE) cannot be very important in the case of the French Revolution, since many other countries had SE without a revolution. In other words, SE is not sufficient for the French Revolution, and indeed is far from being sufficient. Such a claim appears ideally suited to interpretation using a consequentialist idea. Furet relies on similar cases to the French Revolution to suggest what might have been the effect of SE without the

³³ *Ibid.*, 24

³⁴ *Ibid.*, 27

³⁵ *Ibid.*, 24

other causes. These similar cases suggest that SE, on its own, wouldn't have had much effect, and therefore wasn't so important. Furet suggests a limited 'counterfactual experiment': a concept which will be central to chapters eight and nine.

The third example is of a rather different sort, since it concerns the importance of time periods, rather than of events or causes. 'If one defines the Revolution as the collective crystallisation of a certain number of cultural traits amounting to a new historical consciousness, the spring of 1789 is indeed the key period.'³⁶ It can be the case that to judge time periods as more or less important is just as natural as to ascribe importance to causes or events. What justifies such judgements of importance of time period? As Furet makes clear, definitions are all important. Important time periods will be those where, according to one's definitions, the events of interest begin or end. 'Events of interest' may, clearly, be specified in many different ways. But one way is according to the idea that the occurrence of causes of importance also signifies an important period of time; another is according to the ideas suggested in the section on 'Narrative importance'.

Skocpol

Skocpol does not, in general, give much attention to relative importance. Instead, she chooses to follow Mill in restricting causal explanation to the practice of listing necessary antecedents of a jointly sufficient set. However, there is one passage where Skocpol attempts to defend the overall 'statist' approach, through a particular type of importance-claim:

'Perhaps especially because the factors that they consider are indeed an important part of the story, Marxists have failed to notice a crucial point: Causal variables referring to the strength and structure of old-regime states and the relations of state organisations to class structures may discriminate between cases of successful revolution and cases of failure or non-

³⁶ *Ibid.*, 45

occurrence far better than do variables referring to class relations and patterns of economic development alone.³⁷

Skocpol's claim is that those causal variables which are correlated with the presence of the effect to a greater degree are the most important variables. This probabilistic formulation is not one which I have so far considered: is it not a candidate for a promising realist understanding of importance?

In general, it is not. The reason that it appears appropriate for Skocpol's needs is because the target of Skocpol's explanation differs in an important way to those of the other explanations considered. Skocpol's main aim is to develop a *general* understanding of social revolutions; the understanding gained of any particular revolution is a secondary goal. And to decide upon the relative importance of causes in bringing about a *type* of event, it does seem entirely appropriate to rely (at least in part) on information concerning the frequency that causes are present. If inadequacy of state structure is a cause of four of five social revolutions, and volatile class relations are a cause of two of five social revolutions, then inadequacy of state structure is the more important cause of social revolutions in general.

However, Raymond Martin has demonstrated that it is an error to suppose that similar considerations can determine the importance of a *particular* event³⁸. Imagine that we have the same information as suggested in the previous paragraph, concerning the frequency of causes in bringing about five social revolutions. Now imagine a sixth revolution, which is accompanied by both inadequacy of state structure and volatile class relations. Given only this information, is it possible to say which of the two causes was the more important cause of this sixth revolution? According to the frequency measurement, we should be able to; yet this seems implausible. We would surely have to know, in particular, how profound the factors relating to state structure and class relations are in *this* particular case.

We should not, therefore, expect a frequency measurement of causes to be capable of allocating causal importance in the case of *particular* events. This is not to say that

³⁷ Skocpol, 34-35

³⁸ Raymond Martin, 'On Weighting Causes' in *American Philosophical Quarterly* 9(4) (1972), 291-292

noticing the frequent correlation of a type of cause with a type of effect may not be of help in spotting which particular cause might be important. However, whilst the frequency measurement may provide heuristic help, it is no analysis.

Hegel

We have already examined Hegel's theory of historical explanation in outline (chapters two and six). What is relevant to note at this point are the consequences of this approach for judgements of historical importance. The concept is not alien to Hegel; he states that judgement in regard to the 'importance and unimportance of facts' is 'the most basic category in historical judgement'³⁹.

Hegel's philosophical system, in particular the emphasis on the centrality of the development of Spirit (hence, freedom), shapes historical importance. Given that Spirit does develop over time, and that, Spirit is the most philosophically central concept, the development of Spirit is also historically most important. In this respect, Hegel appears to be a good demonstration of an 'intrinsic' conception of importance, in the sense explained previously. In a similar way to the importance of 'the life of Christ' to Christians, 'the development of Spirit' is important to Hegel for ideological or philosophical reasons. Crucially, these reasons are not justified empirically; whereas the heart of the consequentialist idea is that importance is an ascription which is capable of empirical support.

However, at other points, Hegel's judgements of importance are more amenable to a consequentialist understanding. In terms of explanatory target, Hegel's approach to the French Revolution and importance is very similar to Lefebvre's. Just as for Lefebvre, for Hegel the central cause of the Revolution is that which guarantees that there will be a revolution of some sort; even though that cause alone is not sufficient to explain the French Revolution as it actually was. Corresponding to Lefebvre's Marxist cause in Hegel's explanation are the philosophical ideas of the Enlightenment – particularly, Kant's – which led to a new conception of freedom.

³⁹ Hegel, *Introduction to the Philosophy of History*, 14

Tasks for chapters eight and nine

The explication of historical practice concerning judgements of importance has raised interesting ideas. In the main, it appears that the consequentialist notion of importance is a good explication of the practice of Lefebvre, Schama and Furet, who represent seemingly widely differing approaches to importance. Skocpol takes a different approach, resulting from general, rather than particular, nature of her explanatory target.

Not all selection and judgement of importance can be defended according to the idea of causal consequence. Nor should we want to: we should be able, at least in principle, to agree with those reviewers of Schama's account who complain that his principles of selection are 'unbalanced'. The important point is really to be able to defend the appropriateness of a standard according to which comments such as these can be judged. However, it would be implausible even to claim that all judgements of importance of historical particulars can be justified consequentially. Hegel's approach shows that the consequentialist idea will not fit all historical judgements of importance: though even in Hegel's case, the idea is not without use.

The major purpose of considering both historical usage, and also the general arguments against a realist-consequentialist notion, has been to raise issues for that notion to content with. These issues will set the framework for chapters eight and nine. In summary, they are:

- (i) To develop an account of importance which is as metaphysically palatable as the irrealist (projectivist) alternative
- (ii) To defend the selection of causes as potentially more important on the basis of empirical evidence; rather than, as Mill et al. assert, purely pragmatically; or as Walsh asserts, purely on the basis of the values of the historian; or as Collingwood asserts, purely on the basis of 'intrinsic' (philosophical or ideological) features

- (iii) To provide some content to the idea of ‘consequences’, so that we know which are to count in measuring importance
- (iv) To understand the relation between the selection of ‘essential’, ‘pertinent’ or ‘obvious’ features (as in Lefebvre’s or Furet’s account) of the event to be explained, and judgements of importance
- (v) To understand whether and how the consequentialist model can justify other potential ways to analyse importance; for example, the approach suggested by an explication of Lefebvre’s judgement
- (vi) To question the practical applicability of the consequentialist idea; particularly in relation to Schama’s assertion that the relevant factors are many, varied, and not necessarily those which initially appear to be of ‘epochal significance’

Chapter Eight: The Counterfactual Model

We have seen how causal importance relates to explanatory competition in general. The crucial philosophical question is to what extent ‘importance’ is an objectively, and preferably empirically, decidable property. In chapter seven I presented objections to a realist notion of importance, and, consequently, problems which a realist model should answer. In addition, I examined the use of importance implicit in the chosen historical explanations. A consequentialist notion of importance well explains much of this usage. In the examination of historical practice, we also find certain broad suggestions as to how such usage can be defended, and, once more, problems which a realist or consequentialist interpretation must face.

Implicit in judgements of the causal importance of particular events is the idea that the event’s importance is proportional to the extent of that event’s effect. Whilst this idea is hardly contentious, and may even appear almost tautological, I believe it becomes more interesting once the implicit concepts in that idea are explicated. In any case, given only this simple core idea, and assuming that we are able to make sense of the phrase ‘the extent of the event’s effect’, much of what has been written on importance in history – specifically the arguments considered in chapter seven – should therefore be rejected.

The Counterfactual Model, or CM in what follows, states that:

Let c = the causal event whose importance is to be judged

e = the effect of c which frames the question of c ’s importance

x = the counterfactual event which results from c ’s absence

Then the principle states that

(CM) c is causally important to the extent that c makes a counterfactual difference; where the extent of a counterfactual difference is proportional to the difference between x and e

There are two major philosophical issues to arise from CM. Firstly, we need to understand what it means to say that a state of affairs would have happened had *c* not taken place. Secondly, we need to understand what it means to say that two states of affairs are more or less different. Before tackling these issues, I shall discuss CM in relation to contrastive explanation; and will compare CM to other potential realist analyses of causal importance.

As a preliminary, I wish to make three clarificatory points. First: what does ‘frame the question of *c*’s importance’ mean? Simply that instead of providing a measure of the overall importance of *c*, CM provides a judgement of *c*’s importance in bringing about *e*. Our discussion of importance is thereby restricted, and made more tractable: we have turned from importance and selection in general (at the start of chapter seven), to causal importance in particular (as chapter seven proceeded), to causal importance *in relation to a given effect*. This is certainly the notion which is required in evaluating the multiple French Revolution histories, since the comparison is between accounts which rate factors as more or less important in leading to the French Revolution. Yet I believe that even this more limited understanding of importance can be of consequence to wider ideas, if only by providing the possibility of a kernel of objective content to the former, broader notions.

Second: on the connection between causal discovery and causal importance. It is natural to defend the citing of a cause, *c*, as bringing about effect *e* by stating that ‘if *c* hadn’t happened, then *e* wouldn’t have happened’. In other words: *c* makes a counterfactual difference to *e*. This was similar to the idea of ‘cause’ defended in chapter six: that *c* was an empirically necessary antecedent, in the circumstances, for *e* (there are, however, differences to these formulations, which I shall not detail here.) CM is an extension of this idea. Indeed this idea of discovering causes can be incorporated as a limiting case of CM; since where *c* is not a cause, *x* and *e* are identical – *c* makes no difference at all. An advantage of showing CM to be an extension of our normal causal-inferential reasoning is to show, at least *prima facie*, that CM presents a plausible mechanism for judging importance¹.

¹ I am careful to say that CM is to extend our normal account of the *discovery* of causes; rather than that CM is to extend our normal account of the *analysis* of cause. For, if we are to analyse causation

Third: CM should be primarily interpreted as providing a *comparative* measure of importance, rather than an absolute one. The notions of ‘making a difference’ and, by implication, ‘being similar to’ are problematic enough when understood comparatively (‘ c_1 is more different to e than c_2 is to e ’). To extend these measurements to an absolute scale would appear to be at least implicitly comparative. An absolute measure of importance seems hard to justify, and in any case is not required in the context of this thesis’ focus on comparison of explanations.

The Counterfactual Model and contrastive explanation

Given that contrast has been shown to be such a significant part of my understanding of historical explanation in general, it is of interest to question the relationship between the notion of importance, understood according to CM, and contrastive explanation.

Contrast can aid the understanding of CM in two ways. In CM, we should interpret e and x (the outcomes) as states of affairs, which are understood with reference to, usually implicit, contrastive alternatives. This has the advantage of dealing with certain cases in which it may seem that CM leads to rather odd judgements of importance. Certain very general causes, such as c_3 ‘the presence of oxygen’ clearly make a large difference (given that there would have been nothing like the French Revolution, had c_3 not existed), yet are not regarded as being historically important. However, as I demonstrated in chapters three and five, it is likely (given a typical set of contrast cases) that such very general causes are excluded as irrelevant, through a failure to meet the Difference Condition.

In addition, the *cause* is best understood according to contrastive specification. For example, there is a clear difference in efficacy between ‘France having Britain as a geo-political competitor (rather than another country like herself)’ and ‘France having

itself as that which makes a counterfactual difference, then the explication of counterfactuals which follows would lead to circularity. I understand that the distinction between ‘causation’ and ‘discovery of causation’ could appear problematic from an empiricist standpoint. I believe the distinction can be drawn successfully, but that this is simply too large an issue to introduce here.

Britain as a geo-political competitor (rather than fragmented states such as the Holy Roman Empire)'.

As well as applying a contrastive understanding to CM, an appreciation of CM can enrich the understanding of contrastive explanation. Though the selection of a similar set of contrasts can narrow down explanatorily relevant factors, it is still likely that there will be more than a single relevant difference between fact and foils. For example, whilst relevant answers to the question 'why was there a revolution in France, rather than in England?' will be significantly less numerous than to 'why was there a revolution in France?', there will still be a number of possible responses. In such a case, CM selects between these differences on the basis of which made the most difference to the explanandum-phenomenon, and so can augment the contrastive explanation.

But there is a more radical way that we can envisage CM affecting contrastive explanation. I have not said much about how contrasts are selected in the interpretation of an explanation seeking question, except that the choice is primarily a pragmatic one, guided by such general principles as that the contrasts should be similar to the explanandum. Yet it is not unreasonable to suppose that such a choice be also partly based on more defensible reasons. CM can help make sense of this choice: the features chosen are those which make more difference to the event being explained. As Lefebvre claimed, for example, it is reasonable to suppose that the fact that the Revolution was a social change is more fundamental to it than the fact that it started in the summer of 1789.

If this is justifiable, then 'features of the event to be explained' may be regarded as amenable to an analysis according to CM, just as 'causes of the event to be explained' are. And this is what we should expect, given that it is usually possible to identify a particular cause with a particular feature, or set of features, of the effect (in other words, causes are generally causes of features of an event, rather than an event *per se*). Allowing the selection of contrasts to guide CM, and considerations akin to CM to guide the selection of contrasts, demonstrates the interplay between *interest* and *importance*. If this is the case, where should the line be drawn: to what extent should the choice of contrast be governed by considerations of importance? I am afraid that

it seems unlikely that a general answer to this question can be provided. The point to be emphasised is simply that *some* choices of contrast, and hence choices concerning which aspect of the event should be explained, are defensible according to CM; and indeed, must be so, if CM is to have much scope.

The Counterfactual Model: comparisons

To demonstrate the flexibility and analytical power of CM, it is of interest to compare the model to five alternative criteria for justifying claims of causal importance. The five criteria considered have been selected for two reasons. They have been proposed in the literature, and they (like CM) refer to features of the cause which are plausibly *objectively decidable*: meaning not purely pragmatic nor evaluative, but at least partly empirically justifiable. I do not claim that it is impossible to reduce certain criteria to others, and indeed we shall see that common issues are raised. However, they are considered as they are because they have been suggested in quite different contexts.

The first alternative means of assessing importance is via judgements of *probability*. Martin develops two complementary definitions of importance. The first, 'D1', is similar to my model, CM. The core of Martin's second definition, 'D2', is that

(D2) a factor A is more important than a factor B in bringing about P, if the probability of a factor of type-P is greater given A than it is given B.²

Both D2 and CM admit of degree: respectively, the probability may be raised to a greater or lesser extent; and there may be more or less of a difference. It might be thought that 'c made more of a difference to e' and 'c raised the probability of e more' are equivalent. Martin believes the two ideas to be different: he speaks of 'the two senses "of more important" that we set out to analyse'³ (although he doesn't say why

² Martin, 'On Weighting Causes', 298

³ *Ibid.*, 299

he believes this). I agree with Martin that D2 and CM are not equivalent; I disagree that there are two senses of importance at work here. CM can do the whole job.

Imagine asking whether the most important cause of the President's election (e_1) is that she had the support of the centrist party (c_6), or that she had the support of the leftist party (c_7). Let us suppose that the support given by these parties was equal, in terms of delivering the same number of votes; but that both were necessary for the President's election (both are causally efficacious to some degree). Given just this information, c_6 and c_7 equally raised the probability of e_1 , and c_6 and c_7 made the same degree of difference to e_1 . In this case, therefore, D2 and CM coincide in their judgements of causal importance.

Now, however, add the additional fact that if the President had not had the support of the centrist party (c_6), not only would she have lost, but she would have also suffered further negative consequences; which might include not being nominated for the next election. However, had the President not had the support of the leftist party (c_7), she would have lost, but suffered no further consequences. In this case, D2 would still regard c_6 and c_7 as equally important, on the grounds that they equally raise the probability of e_1 .

Yet, given the additional information, CM would regard c_6 as more important than c_7 , on the grounds that the situation where c_6 didn't occur would have been more different than the situation where c_7 didn't occur. The reason for the difference is that CM takes into account not only the causal effect of c_6 and c_7 , but also the implications of the differing situations where the effect fails to occur. And it seems clear that CM gives the answer which matches intuition: that in the modified scenario, the support of the leftist party was more important than the support of the centrist party.

A measure of importance can be derived from a concept which has already been discussed in chapters four and five, that of a 'complete explanation'. We might think that causal histories which are more complete thereby point to causes which are correspondingly more important. Interpreting 'complete' as 'sufficient' (which I have suggested is part of a reasonably plausible explication of 'complete explanation') would imply that causes which are more nearly sufficient are thereby more important. There is, however, no need to conduct a comparison of CM and a 'sufficiency'

analysis of importance, since the latter is equivalent to Martin's D2. This can be appreciated through the realisation that a cause is sufficient for an effect to the extent that it raises the probability of that effect. A fully sufficient cause increases the probability of the effect happening to 1.

However, given that Martin's D2 and CM are very similar, this raises the interesting possibility of analysing completeness in terms of CM. For a start, CM will be able to provide an insight not only into what a complete explanation amounts to, but also what is meant by a more or less complete explanation. This would be an advantage in understanding relations of exclusion between actual historical explanations, since a role for partial completeness seems necessary in interpreting KPEE (just as I argued, in chapter five, that we should make space for 'partial overlap of explanatory target'). In that case, the analysis of CM will have a bearing on explanatory exclusion in addition to explanatory competition more generally; but this is not an aspect of CM which I will examine in any detail.

A somewhat different attempt to connect importance to features of probability can be derived from Hitchcock's 'probabilistic relevance' test (PR)^{4,5}, which itself owes a large debt to Van Fraassen's contrastive treatment of explanatory virtue⁶. The key idea is that an explanation is *salient* (which I take to be equivalent to the *importance*) to the extent that it limits a set of contrastive possibilities. Specifically, the explanatory salience of *c* is proportional to

$$\frac{\text{(number of contrastive possibilities remaining once } c \text{ is taken into account)}}{\text{(number of contrastive possibilities before } c \text{ is taken into account)}}$$

For example, I and my friend wish to go out to dinner. There are three contrasting options: French, Italian, and Thai. We end up eating Italian. My preference was to eat something European, my friend's preference was to eat pasta. In this case, according to PR, my friend's preference is explanatorily more relevant than mine,

⁴ Hitchcock, (especially) 603-604

⁵ This idea was discussed as a possible replacement of Lipton's Difference Condition for explanatory *adequacy* in chapter three; there I argued that PR should not replace the Difference Condition. But, given the emphasis on explanatory *salience*, perhaps PR is better interpreted as a test of causal importance, as I have done at this point.

⁶ Van Fraassen, 146-151

since theirs limits the possibilities from three to one, whilst mine limits the possibilities from three to two.

There is no doubt that the idea of limiting possibilities can be a guide to importance; yet we should not treat PR as an analysis of importance. Consider a case in which my preference for European food is very strong, whereas my friend's preference for pasta is a mere whim, easily overridden. In that case, it might well be appropriate to cite my preference as more important than my friend's, despite the relative probabilistic relevance of each.

A further problem with PR as an analysis of causal importance, or explanatory salience, is that if the contrastive possibilities are not 'correctly' enumerated, then the criterion will lead to unintuitive results. To change the example: imagine assessing the importance of the assassination of Archduke Ferdinand in bringing about the First World War according to PR. If we were to regard the set of possible outcomes (the contrast set) as

{(nothing happens), (local conflict takes place), (war breaks out between Austria and Russia), (world war breaks out)}

then the assassination restricts the possibilities from four to (the latter) three. This would accord the assassination with some salience or importance, but not much. However, it is easy to see that the judgement of importance might be skewed through selection of different possibilities: if the set was

{(no Austrian reaction at all), (Austria is upset but takes no action), (a strongly worded Austrian letter is sent to Serbia), (world war breaks out)}

then the assassination would restrict the possibilities from four to (the last) one, and would therefore count as extremely salient; just as salient as, for example, the fact that there was a long term arms race between Austria and Serbia.

The fourth alternative realist means of analysing historical importance is via the contrasting notions of *precipitant* and *pre-condition*. The former refers to causes which occur a relatively small time before the effect; loosely, those which 'trigger' the effect. The latter refers to causes which began a relatively long time before the effect; those which provide the necessary structural background for the occurrence of

the effect. A popular opinion is that pre-conditions are causally important, precipitants less so. Morton White considers the view that this principle is axiomatic⁷, and Martin defends the idea through connection with his previously cited principle D2: his argument for D2 is that this principle is required to elucidate the intuition that pre-conditions are more important than precipitants.

I agree that we usually regard pre-conditions as more important than precipitants. It is only natural to place, for example, long term socio-economic or intellectual transitions at the heart of an account of the origins of an event such as the French Revolution, rather than the particular accidents and decisions immediately preceding the event⁸. However, I disagree with M. White and Martin that such a preference for pre-conditions is always correct. Consider, for example, Skocpol's claim, quoted in chapter two, that the key fact requiring explanation is 'the specific political crisis that launched the French Revolution.'⁹ And, in any case, I disagree that we need a separate motivation to CM to account for this typical preference.

What we mean by 'pre-condition' and 'precipitant' could be specified according to simply the length of time that the causal feature had been present, but such a definition does not appear to capture the reason why we would be inclined to regard pre-conditions as generally more important than precipitants. Instead, I suggest that the terms be specified more precisely according to the following schema. Effect *e* is brought about by a combination of precondition *S* (*S* for 'structural') and precipitant *P*₁. However, it is the case that alternative precipitants, *P*₂, *P*₃ ..., could have combined with *S* to produce an outcome not too dissimilar to *e*; for example, they might lead to something very like *e* but dated a few months earlier or later. Crucially, the situation for *S* and for *P* is not symmetrical: there are no alternatives to *S* which could easily be imagined, that would have brought about an effect similar to *e*. This seems a reasonable way of explicating the difference between pre-conditions and precipitants, and demonstrates the situation to be a special case of comparing causes having readily available *functional equivalents* with those which do not.

⁷ Morton White, *Foundations of Historical Knowledge* (New York: Harper and Row, 1965), 134-141

⁸ That is, as long as we can argue both that such long term changes took place and that they were causally relevant at all; facts which are disputed by some, most notably Schama.

⁹ Skocpol, 52

If this is an accurate explication of what is meant by pre-condition and precipitant, then CM is able to subsume the above case; so long as certain terms in CM – most notably ‘absence’ – are understood correctly. The wrong way to interpret ‘absence’ would be to suppose that had P_1 not taken place, there would have been no precipitant at all. For if this is the case, the importance of P_1 and of S will be equal. The correct way to interpret ‘absence’ is to use the idea that, had P_1 not taken place, it is likely that one of $P_2, P_3 \dots$, would have occurred instead. If this is so, then the absence of P_1 would make relatively little difference; therefore, the judgement of importance suggested by CM would concur with the intuition that the pre-condition was more important than the precipitant. An advantage of this explication in terms of CM is that it does not guarantee, a priori, that pre-conditions *are* more important than precipitants. Whether this is so depends on the relevant, empirically informed, counter-factual judgements. For this explication, more needs to be set about the notion of ‘absence’; this will be a major element of chapter nine.

The final realist analysis of importance to be considered is that which claims that important causes are those which are *robust*. Whilst the socio-economic conditions of pre-Revolutionary France may be considered robust causes, the decisions of the King may not be: for that reason, the former may be considered the more important cause. Such an analysis is best treated in a similar manner to the treatment of pre-conditions vs. precipitants.

A robust cause is one which would remain causal in similar contexts. A non-robust (fragile) cause would be unlikely to remain causal even if the situation were only a little different. Let us imagine that socio-economic changes, occurring over the course of roughly a century, are deemed more important than the foolish equivocation of the crown during the years 1788 and 1789. We might defend this judgement by pointing to the *stability* of the former cause. It was more likely that the former, rather than the latter, cause would exist in a range of close, or relevant, possible situations.

Defined in the above way, robust causes correspond in general to pre-conditioning causes (and non-robust to precipitating). This is so because there are usually fewer plausible functional alternatives to robust causes, just as there are few or no plausible functional alternatives to pre-conditioning causes. To the extent that an analysis of

importance in terms of robustness corresponds with this idea, it can be subsumed by CM in the manner demonstrated in the previous section.

I have aimed to demonstrate that CM can account for, and subsume, alternative ‘realist’ analyses of causal importance. These alternatives may provide heuristic benefit in the search for important causes (in a similar manner to the frequency analysis as a guide to particular importance, discussed in chapter seven). I do not want to suggest that there are not some cases which are very neatly accounted for by one or other of the alternative suggestions considered. However, the generality of CM demonstrated in this section provides an argument for its *analytical* role, in addition to its potential use as a methodology guiding justifications of causal importance.

Counterfactuals as implied experiments

The task for the remainder of this chapter will be to outline an account of counterfactuals which is both philosophically and practically appropriate, and to show how this account renders CM plausible. In this section I will first suggest why we should be concerned with an analysis of counterfactuals at all. I will then give my ‘implied experiment’ account of counterfactuals, and the reasons why this account should be adopted.

We might think that an *analysis* of counterfactuals is not what is needed for our purposes, since our interest is in how counterfactuals could be used, rather than how they should be analysed. I have some sympathy with this protest, especially in the context of the present study which primarily aims to use philosophical concepts and resources to tackle a pressing historiographical problem, rather than to enter analytical debates over the nature of these concepts. Indeed, the account which I develop will, it must be admitted, fall short of being a complete analysis of counterfactuals in general. Yet I will reject the overtly pragmatic emphasis on *method* alone, for three reasons.

Firstly, as I argued in chapter one (with respect to Martin's adoption of the Deductive-Nomological analysis), analyses are not normatively neutral. It is a mistake to suppose that the analysis of a concept can be entirely divorced from the methodology which utilises that concept. Rather, analyses should be developed within a framework set by methodological understanding. It is certainly the case that methodology is constrained by the analyses at work.

Secondly, and as a corollary of the first point, it is likely that an analytical investigation will produce issues which have an important bearing of the use of that concept. At the very least, this possibility should not be ruled out a priori. In what follows, I shall argue that there are points of debate in the literature dealing with analyses of counterfactuals which finds resonance in issues relevant to the practical use of counterfactuals. In particular, the traditional analytic problem concerning cotenability, and the doctrine of 'No Gratuitous Difference', are both relevant to the applicability and use of counterfactuals (as shown in chapter nine).

Thirdly and most importantly, it is a fact that analytical worries can interrupt practice. Scepticism over the possibility of counterfactuals being anything more than 'a web of hazy imaginings, fantasy and the subjunctive mood'¹⁰ leads, quite naturally, to the exclusion of counterfactual reasoning from 'rigorous' historical debate. The textbook *Historians' Fallacies* is particularly critical of the use of counterfactuals in history:

'The *fallacy of fictional questions* ... consists in an attempt to demonstrate by an empirical method what might have happened in history, as if in fact it actually had ... There is nothing necessarily fallacious in fictional constructs, as long as they are properly recognized for what they are and are clearly distinguished from empirical problems. ... But they *prove* nothing and can never be proved by an empirical method.'¹¹

Even a defender of the use of counterfactuals in history, Raymond Martin, worries that 'whatever the problems with counterfactuals, we seem to be stuck with them'¹².

We need to be able to show that it is reasonable to suppose that counterfactuals can be

¹⁰ Robert Musil; quoted in the introduction to *Virtual History: Alternatives and Counterfactuals*, ed. Niall Ferguson (London: Macmillan Publishers, 1988), 4

¹¹ D. H. Fischer, *Historians' Fallacies: Towards a Logic of Historical Thought* (London: Routledge, 1971), 15-16

¹² Raymond Martin, 'Causes, Conditions and Causal Importance' in *History and Theory* 21 (1982), 70

both decidable in principle, and that they can actually be known, over at least a large range of historical cases.

My main concern in the following investigation of counterfactuals is the question of how far counterfactual decidability can be grounded in an objectively defensible manner. A secondary task will be to connect analytic and practical concerns. I will, throughout, limit my attention to the sorts of counterfactuals which might be required in CM; I will not, for example, focus upon ‘counter-legals’ (counterfactuals whose antecedent supposes an alteration of physical law).

My starting point is to focus on the relationship between antecedent and consequent. I shall suggest that that relationship is causal. From this intuition can be developed an attractive way to understand counterfactuals: as imaginary, or ‘implied’, experiments. The same facts that make ordinary experiments true or false give counterfactuals their truth values – namely, causal laws and causal regularities. This causal knowledge needed to assess implied experiments is arrived at by actual experiment. Thus, actual experiments produce knowledge of causal generalities, whilst counterfactuals (implied or imaginary experiments) are parasitic upon such knowledge: in much the same way that a prediction is parasitic upon experimental knowledge¹³.

More precisely: a counterfactual asserts that an implied experiment leads from antecedent to consequent, where the antecedent ‘sketches’ the initial conditions of the experiment, and the consequent ‘sketches’ the outcome of that experiment. This implied experiment involves, paradigmatically, the move from initial conditions (which include, but are not exhausted by, those specified in the antecedent), via relevant causal laws, to an outcome. The truth or falsity of a counterfactual statement is determined by the match between the implied experiment outcome and the counterfactual consequent.

¹³ Indeed, IE implies that counterfactuals and predictions are symmetrical: that ‘e will happen if c does’ (asserted before c took place) and ‘e would have happened if c had’ (asserted after c failed to take place) are equivalent. This is an idea which Edgington challenges (Dorothy Edgington, ‘Counterfactuals and the Benefit of Hindsight’, 2002, unpublished). She uses ‘switch’ counterfactuals, such as the lottery example: ‘you would have won had you selected ticket 938’ and ‘if you select ticket 938 you will win’. In ‘switch’ cases, there seems to be an asymmetry: whilst the counterfactual is appropriate, the prediction is not. Nevertheless, Edgington’s challenge is not irrefutable; the intuition can plausibly be denied (by maintaining the prediction to be correct, though strange).

Symbolically, the *Implied Experiment account of counterfactuals* (IE) states that:

Counterfactual S (if A had been, then C would have been) is true if and only if:

- (i) Implied experiment IE (experimental initial condition IC to experimental outcome O) is correct *and*
- (ii) IC is a satisfactory development of A *and*
- (iii) O is a satisfactory development of C

This account can be seen to work well when considering simple counterfactual statements. Consider the assertion that ‘if I dropped my glass on the hard floor, then it would break’. We would go about deciding upon the truth of this statement by first imagining the antecedent situation of the glass being dropped, then combining this fact with relevant laws (laws concerning gravity, properties of solids, perhaps chemical bonds), and properties of the items involved (that the glass is a fragile solid and so forth). These facts lead (deductively or probabilistically) to a situation predicted by the consequent – that the glass breaks.

What are the conditions for the correctness IE, the implied experiment? As I have suggested, the paradigm case of correctness is when IC can be conjoined with relevant laws, and can be shown to deductively lead to O. This is the case in the above example. That is certainly sufficient for the correctness of IE, yet I do not wish to claim that it is necessary.

As the counterfactual situations considered become more complex, and the laws involved less easy to guess, the notion of determining the implied experiment according to deductively applied laws must be relaxed. This is on pain of restricting the account to only very simple counterfactuals, such as the dropped glass. The most obvious way to make the assessment of counterfactuals more tractable is to replace (knowledge of) ‘relevant causal laws’ with (knowledge of) ‘approximate empirical regularities’. It does not matter, for the purposes of IE, whether or not such regularities are *really* laws: in other words, whether they are, to use Mill’s terminology, general or derivative laws.

Whilst we must be mindful of such potential errors as mistaking correlation for genuine causation, the need to supplement strict laws with regularities is nonetheless necessary; both in common practice of deciding counterfactuals, and for the implied experiment model. It is important to notice, however, that the desired objectivity of the approach remains. According to IE, our assessment of a counterfactual can be no better than the match between (i) the regularities *we* use to assess the counterfactual situation, and (ii) the actual regularities governing the appropriate implied experiment. And both the degree of fit between (i) and (ii), and the nature of (ii) itself, are mind-independent facts.

Indeed, I think that the implied experiment account should be properly interpreted in a still broader manner. In certain situations it is permissible to assess an implied experiment on the basis of *simulation*. In gaining scientific knowledge there are (at least) two ways to proceed: we can apply to our topic in question a *theory* which has already acquired plausibility in other contexts; or we can *model* the topic in question. For example, both a developed theory of aeronautics and a wind-tunnel can be useful in attempting to understand the mechanics of a particular aircraft¹⁴. Similarly, I suggest that implied experiments may be understood and assessed by modelling or simulation, just as they are by theory. (To appropriate Putnam's well known question: how do we know that square pegs can't fit into round holes? Certainly not through the application of a complete, low level theory!)

This extension is, I believe, necessary in order to appreciate the full extent of counterfactual reasoning, and in order to distinguish the implied experiment account from other analyses, particularly the meta-linguistic (discussed in the next section). Such an extension does, however, make it difficult to provide strict, and specific, conditions for when an implied experiment be considered correct. The only fully general criterion that can be given is that *an implied experiment is correct where the corresponding actual experiment is correct*.

¹⁴ This example is drawn from Stich and Nichols (Stephen Stich and Shaun Nichols, 'Folk Psychology: Simulation or Tacit Theory' in *Mind and Language* 7 (1992): 35-71). Their topic is the question of whether our folk 'theory of mind' is really a tacit theory, or is rather a method of simulation.

The implied experiment model of counterfactuals has a number of attractions, particularly insofar as CM is concerned. Most importantly, scepticism concerning the decidability of counterfactuals can be allayed. Even amongst those theorists who have attempted to employ the notion of counterfactuals to ground historical practice, such scepticism has been evident (in, for example, the work of Hawthorn¹⁵, Martin¹⁶, and Hammond¹⁷). Using the implied experiment model, we can say that as long as the experimental initial condition is properly specified, and we have the correct kind of experimental knowledge (theoretical or stimulatory), then counterfactuals are philosophically and practically decidable.

A second advantage to the implied experiment account is, therefore, that the sources of methodological difficulties concerning counterfactual decidability can be clearly pinned down. Difficulties may be due to the specification of initial facts, or to a lack of knowledge concerning relevant laws or regularities. Chapter nine explores each of these issues. Thus the implied experiment account also helps us to understand under what conditions counterfactual reasoning *does* become dubious. Hammond, for example, is forced to maintain that ‘beyond a certain point it doesn’t make historical sense, and perhaps doesn’t make philosophical sense, to persist in weighting causes’¹⁸; yet this point is not located. I shall argue in what follows that certain counterfactuals do indeed lack epistemological (historical) and, less often, philosophical sense; but using the implied experiment model we can give a good explanation as to when this should be so.

Thirdly, the implied experiment account is psychologically compelling - it serves as a general account of counterfactual decidability, but also (in a very general sense) as a phenomenologically realistic methodology. There are clear similarities with certain psychological accounts of counterfactual reasoning; for example, Nichols/ Stich/ Leslie/ Klein’s ‘off-line simulation’ model¹⁹. This treats counterfactual reasoning as a

¹⁵ Geoffrey Hawthorn, *Plausible Worlds: Possibility and Understanding in History and the Social Sciences* (Cambridge: Cambridge University Press, 1991)

¹⁶ Martin, ‘On Weighting Causes’ and ‘Causes, Conditions and Causal Importance’

¹⁷ Michael Hammond, ‘Weighting Causes in Historical Explanation’ in *Theoria* 43 (1977): 103-128

¹⁸ *Ibid.*, 118

¹⁹ Shaun Nichols, Stephen Stich, Alan Leslie, David Klein, ‘Varieties of Off-Line Simulation’ in *Theories of Theories of Mind*, eds. P. Carruthers and P. K. Smith (Cambridge: Cambridge University Press, 1996), 53-59

mental process beginning with a *pretend* belief X, which is then processed using our normal mechanisms of inferential reasoning. These mechanisms are conducted ‘off-line’ (in other words, outside the agent’s actual situation), in order to find out what would follow from X. This topic has intriguing parallels with the question of counterfactuals, particularly in the discussion of ‘off-line simulation’ as a means of explaining and predicting others; such an idea seems close to my notion of an implied experiment.

A final advantage is that the implied experiment account provides an description of counterfactuals in terms of what are (or at least, would seem to be) markedly less mysterious entities: matters of fact, and causal laws, regularities or properties. In comparison to the ‘Possible Worlds’ theory of counterfactuals, considered in the next section, counterfactuals are kept firmly in this world. To Lewis’ question of what we refer to when we talk about ‘ways things could have been’²⁰, we reply that we refer to facts about this world (events, laws, and so forth), mentally *arranged* in a manner different to that in which the facts were actually arranged.

Alternative analyses of counterfactuals

Whilst the term ‘implied experiment’ is my own, as is to be expected the general approach of the account bears some similarity to other analyses of the last 50 years. Mackie makes use of the notion of *supposition*; “‘if P, Q’ is to assert Q within the scope of the supposition that P”²¹. Indeed, in spelling out what he means by ‘supposition’, in a non-trivial way, Mackie takes the term to mean ‘envisaging the possibilities and consequences in question’²².

The implied experiment account is also similar in certain respects to Goodman’s analysis, which has since been known as a Meta-Linguistic analysis (henceforth

²⁰ David Lewis, *Counterfactuals* (Oxford: Blackwells, 1973), 84

²¹ J. L. Mackie, *Truth, Probability and Paradox* (Oxford: Clarendon Press, 1973), 93

²² Mackie, 100

ML)²³. ML maintains that a counterfactual states that an argument of a certain kind exists, or perhaps that a counterfactual *is* such an argument, in outline. The relevant argument is one whose premises include the antecedent proposition, in addition to other initial conditions and laws of nature; and whose conclusion is the consequent. True counterfactuals refer to a valid argument of this kind: this is why they are *meta-linguistic* entities. Note that the same sort of items appear in both IE and ML: in particular, initial conditions and laws of nature.

However, the implied experiment account does disagree with ML. Most fundamentally, they disagree over the reference of counterfactual statements. In IE, counterfactuals refer to implied experiments – models, simulations; in ML, counterfactuals refer to arguments. Under the implied experiment model, arguments are merely the clearest means of assessing implied experiments. To focus entirely on arguments is to illegitimately restrict the scope of counterfactual reasoning. As I have argued, it is often the case that counterfactuals are decidable, even where we do not have precise knowledge of the laws driving the relevant system.

In addition, ML faces a number of serious problems due to its focus on arguments. The notorious problem of cotenability is one: discussion of which must wait until chapter nine. Related is the fact that ML has great difficulties in properly accounting for the asymmetry in counterfactuals. Why is it that ‘if the glass is dropped, then the glass breaks’ is a sensible counterfactual, whilst ‘if the glass breaks, then the glass is dropped’ is not? For, given a sufficiently detailed account of the glass breaking, and relevant laws, it is not implausible to suppose that an argument could be constructed leading to the fact that the glass was dropped. A further problem is that ML encourages a view of counterfactuals which is insufficiently empirically sensitive. Counterfactuals’ decidability should rest on the same entities as actual empirical investigation; and such entities can, and do, alter as science progresses.

The latter two comments suggest a parallel between ML and the Deductive-Nomological (D-N) analysis of explanation. I suggest that the difference between ML and IE is analogous to that between a D-N analysis of explanation and an ontic (causal) conception of explanation. The D-N analysis explicates explanations in

²³ Nelson Goodman, *Fact, Fiction and Forecast*, 7-8

terms of arguments, propositions, and deductive relationships; ML explicates counterfactuals in the same way. The major problems of the D-N model concern its difficulties in accounting for explanatory relevance, and for explanatory asymmetry. This is so since the provision of a D-N argument is no guarantee for explanatory relevance, as we saw in chapter three. The problems of relevance and asymmetry remain even if the D-N model were to be relaxed so as to require high probability rather than certainty. Analogously, the problems with ML remain whether deduction, or just probabilistic likelihood, is required in order to secure the truth of the associated counterfactual.

The ontic conception of explanation proposed a very simple solution: that explanations trace whatever entities are actually involved in bringing about the event to be explained (usually, and loosely, causes). In the same way, the implied experiment analysis of counterfactuals does not place any particular a priori restriction upon what items should be included in a counterfactual, for these will depend upon our actual inferential mechanisms (as the implied experiment mirrors actual experiments).

The major alternative to ML is the Possible Worlds (hereafter PW) analysis of counterfactuals, initially proposed by Stalnaker²⁴ and Lewis²⁵. Although there are variations between different PW models, in particular between Stalnaker's and Lewis' analyses, the points I wish to make concern the general features of a PW account. For this reason, it is sufficient here to specify the PW analysis as follows: 'if it had been the case that a, then c would have obtained' is non-trivially true if in the most similar world to ours in which a is true, c is also true²⁶.

The first point I wish to make is that, simply as stated, PW is free of content: in the sense that it cannot actually deliver a judgement concerning which counterfactuals should and should not be asserted. This claim is not contentious – Lewis is quite in

²⁴ Robert Stalnaker, *Inquiry* (Cambridge, Mass: Bradford Books, 1984)

²⁵ Lewis, *Counterfactuals*

²⁶ Two of the more salient differences between Stalnaker's and Lewis' versions of PW are: (i) that Lewis admits of ties between close worlds, whilst Stalnaker does not; and (ii) that whilst Lewis describes PW in terms of measures of similarity between worlds (as I have done), Stalnaker speaks of a selection function between worlds.

agreement²⁷ – yet is of consequence. For there to be content to PW, we must supply content to the notion of ‘similarity’. Indeed, depending on the similarity function, PW could give results entirely in accordance with ML: if the function is defined such that changes of law always produce a less similar situation than changes of fact.

Because the similarity function is unconstrained, the use to which counterfactuals are put in CM is problematic. Recall that the aim of CM was to demonstrate that historians can disagree over causal importance in a manner where argument is appropriate; where the differing ascriptions of importance are not *just* ‘different’. Yet PW provides no guidance as to which counterfactuals are properly assertable. To do this, the similarity function must be constrained.

If the required similarity function was closely related to our intuitive judgements of similarity, then we could at least demonstrate a conceptual advance: from counterfactual judgements to similarity judgements. But this is simply not the case: ‘similarity’ turns out to be largely a term of art. This is best seen by considering the notorious ‘Nixon’ example²⁸: ‘if Nixon had pressed the button there would have been a nuclear holocaust’ should come out true. Yet, is it not the case that a world which suffers a holocaust is *less* similar than one where Nixon presses the button, but the mechanism fails? Under the PW account, this would lead to the counterfactual being deemed false.

In response, Lewis suggested²⁹ that the following priorities be adopted when judging ‘similarity’. The least similar worlds are those which differ widely in terms of law. Next most important in judging difference are large differences in particular matters of fact; followed by small violations of law. The least important aspect of similarity lies in small differences of particular fact. It is clear that using these vague principles charitably could lead to judgements which match intuitions concerning such examples as ‘Nixon’. But the important point is that any content given to the similarity function is derived from our pre-existing intuitions concerning which counterfactuals are acceptable, and which not. Of course, this intuitive basis should play a role in the analysis of any concept; yet we need more for the application of CM.

²⁷ For example, David Lewis, ‘Counterfactual Dependence and Time’s Arrow’ (originally 1979) in David Lewis, *Philosophical Papers Volume II* (Oxford: Oxford University Press, 1986), 41

²⁸ First raised by Kit Fine, ‘Review of David Lewis, Counterfactuals’ in *Mind* 84 (1975), 452

²⁹ Lewis, ‘Counterfactual Dependence and Time’s Arrow’, 47-48

In short, we need the possibility of normativity in our judgements of causal importance; a normativity that is absent from the PW analysis of counterfactuals. Any content PW can provide is based either on intuitive notions of 'similarity', which leads to incorrect judgements; or entirely on pre-existing counterfactual judgements. It does nothing to normatively ground such judgements. We need an account which delivers plausible *general* rules for the decidability of counterfactuals, such that specific counterfactual statements can be judged according to their fit with those rules. Further, I think that it is not enough for those rules to be constructed only in such a way that they fit with existing judgements of counterfactuals; we should expect those rules to suggest *why* those judgements are the way that they are.

I would argue that the implied experiment account should be relied upon as the means to justify judgements of counterfactual correctness. This model fits with the wide range of historical counterfactuals which we intuitively wish to assert, delivers tangible results, and can be seen as normatively binding in the right way.

Chapter Nine: The Counterfactual Model: Defence

In chapter eight I presented the Counterfactual Model (CM) of causal importance. CM is an attempt to specify the intuition that important causes ‘produce more of an effect’. I have argued that CM is able to subsume alternative conceptions of causal explanation. However, to establish its plausibility, it is necessary to investigate the analytical components of CM: *counterfactual* and *degree of difference*. The latter notion will be the subject of the penultimate section of this chapter. The task of examining historical counterfactuals was begun in chapter eight, in which I made a case for the Implied Experiment (IE) account of counterfactuals. In the first three sections of this chapter, I continue the investigation into IE, through a critical consideration of key aspects of that model.

Is the Implied Experiment model an analysis?

I have attempted to defend IE over the more traditional analytical approaches to counterfactuals, at least with regard to their use in grounding historical counterfactuals. The major advantages of IE are, first, that it is a model which can be normatively binding: thus it can account for the fact that counterfactuals can be right or wrong; second, that it is broad enough to be able to potentially justify the way that counterfactuals are actually asserted. Further advantages will be discovered through consideration of the topics of the present chapter.

However, we must also question whether IE is really in direct competition with ML (the metalinguistic analysis) and PW (the possible worlds analysis). In other words, to what extent is IE an *analysis*? Regarded as an analysis, a first concern with IE is that it doesn’t cover all counterfactuals. In particular, counter-legals are impossible to

resolve, due to the fact that IE requires the use of actual laws and causal regularities in deciding upon counterfactual outcomes. It may be possible to develop an approach, broadly in keeping with IE, such that counter-legals can be understood: by, for example, using actual laws (etc.) except for those explicitly regarded in the antecedent as differing from actual laws. However, whether or not this is possible, there is a more serious difficulty for IE.

It is awkward to specify both what precisely an Implied Experiment *is*, and under what conditions it is correct, without using counterfactual terminology. When is an implied experiment correct? My answer, in chapter eight, was that an implied experiment is correct where ‘the corresponding actual experiment is correct’. But, surely, the point is that there is no corresponding actual experiment: should we not therefore be more forthcoming in rephrasing the above to ‘...the corresponding actual experiment *would* be correct’? And: what do we refer to when we refer to counterfactuals? I answered, ‘to facts about this world (events, laws, and so forth), mentally *arranged* in a manner different to that in which the facts were actually arranged’. Again, it is plausible to suppose that the phrase ‘actually arranged’ suggests a pre-existing implicit understanding of the notion of a counterfactual alternative.

I won’t attempt to refute the thrust of this criticism; I think it is true that, analytically, IE must be supplemented by a more general notion of counterfactuals. Therefore, IE should be regarded as a normative explication of our counterfactual practices, rather than a fully fledged analysis. However, two caveats. First, I take it that, analytical shortcomings notwithstanding, IE performs a vital role within the model of historical importance given by CM. Second, the analysis which ‘frames’ IE can be specified quite minimally.

As in PW, I believe that it is necessary to suppose that counterfactuals are, in the most general terms, ways of selecting between possibilities. However, these possibilities need not be regarded as fully fledged ‘possible worlds’. My preferred approach will be to treat counterfactuals as propositions which select an appropriate world-line from those which diverge from the actual world at a point just prior to the time referred to by the antecedent. ‘Appropriate’ is – in the case of historical counterfactuals, at any rate – to be specified by IE. IE therefore retains the starring role in the analysis of

counterfactuals, even if it is not the whole story. The role of ‘world-lines’ will receive further attention in the following section.

The Initial Condition, or experimental situation

There are two main elements of the Implied Experiment account of counterfactuals which require further investigation. First, the ‘initial condition’ (IC) of the experiment; second, the extent to which the ‘experiment’, and hence the counterfactual, yields definite and knowable outcomes. The former is the subject of this section.

I will approach the role of IC through the parallel issue of *cotenability*. This issue was raised by Goodman, in discussion of his own analysis of counterfactuals: ML. The issue of cotenability is, in Goodman’s terms, ‘the problem of relevant conditions’; what premises should be included in a counterfactual argument? Consider the plausible counterfactual claim:

DEM: ‘If Germany had won the Second World War, Britain would not now be a democracy’

Considering that ML requires the *deduction* of the consequent, clearly more premises than ‘Germany won the Second World War’ are required. We also need such statements as ‘Germany was a dictatorship’, and ‘The German war aims included alteration of governments’, if the deduction is to have any chance of being valid.

Yet not all facts can be used as premises in a counterfactual-argument. For a start, ‘Germany did not win the Second World War’ should be excluded, even though it *is* a fact. For if premises of this sort were admitted, then any counterfactual whatsoever should be true: since the premise, together with the antecedent, forms a contradiction; and a contradiction logically implies any conclusion we would care to name. But there are also other facts which, though not in logical contradiction with the antecedent, should not be asserted as premises in the counterfactual argument. If facts of this sort were admitted, we would again be in the unfortunate position of having to

affirm that many intuitively false counterfactuals were, according to ML, true. For example:

DEM₂: 'if Germany had won the Second World War, and her war aims included alteration of governments, then Germany would not have been a dictatorship'

looks odd, to say the least. Yet it is perfectly acceptable, according to ML, *as long as* the fact that 'Britain is now a democracy' is included as a premise. This can be justified if we suppose that that 'Germany won the Second World War', 'The German war aims included alteration of governments', and 'Germany was a dictatorship' are jointly sufficient to imply that Britain is *not* now a democracy. Given that we have assumed that it is, one of the other premises must be denied.

Cotenable premises are those which can, and should, be properly asserted in a fully specified counterfactual-argument, alongside the explicit antecedent. DEM and DEM₂ illustrate the bad counterfactuals that can result from adopting premises which are not cotenable. There are three related questions which fall under the heading of 'the problem of cotenability'. First, under what precise conditions are premises cotenable? Second, is it possible to state these conditions without using counterfactual terminology? Third, there is the question of *back-tracking* (raised by DEM₂). Whilst counterfactuals usually assert the dependence of future events upon past events, under what conditions, if ever, should the past be said to be counterfactually dependent upon the future?

Goodman, notoriously, failed to answer these questions in a satisfactory manner. For this reason, he doubted whether counterfactuals would feature as part of a mature scientific world-view. I will leave the metalinguistic analysis in trouble at this point, and instead apply the analogous questions to IE. IE doesn't treat counterfactuals as implied arguments, but implied experiments; so it doesn't talk in terms of 'providing a complete set of premises'. However, it does require IC: 'the experimental initial condition', which, in many cases, will appear to be quite similar. Just as a full set of cotenable premises is required by ML, a properly specified initial condition is required by IE. And, just as a counterfactual antecedent does not state all the premises required for an assessment of counterfactuals according to ML, I made it

clear in chapter eight that, according to IE, the initial experimental condition is only ‘sketched’ by the antecedent.

In terms of IE, the key point to recognise is that *implied experiments can only be assessed to the extent that we are provided with a tractable experimental situation*. In other words, if we are asked ‘what happens if...’, and that ‘if’ is not properly specified, then no answer is possible. The tractable experimental situation is rarely provided explicitly. In the case of DEM, for example, ‘If Germany had won the war’ is ambiguous; it could be read as ‘Germany defeated Britain, Russia, and the US’, or it could be read as implying only a European victory. It is highly likely that which scenario is envisaged will affect the counterfactual outcome.

The potential for ambiguity goes to the heart of the Counterfactual Model. The initial experimental situation is referred to in CM as an *absence*: ‘the counterfactual situation that results from c’s absence’. But there are many possible ways for something to be absent: for an absence leaves a further, positive situation. By imagining the absence of ‘the success of the D-Day invasions’, we might imply that the invasion succeeded one day later, one year later, one decade later: or perhaps not at all.

I suggest that there are three core ideas which relate to construction of the initial condition of the implied experiment, which together answer the analogous ‘problem of cotenability’ as it applies to IE. First, for the construction of a reasonable implied experimental situation, much needs to be *stipulated*. Whilst we often aren’t very explicit about what precisely we wish to be imagined upon asserting a counterfactual, we should be able to fill in the relevant gaps if called upon to do so.

In chapter eight, I argued that not only is the *effect* to be understood contrastively, but often the *cause* should be, too. To recognise the contrastive nature of the cause under examination will often resolve the ambiguity at a stroke: for we do not deal with the cause ‘Germany winning the war’, but ‘Germany winning the war (rather than just defeating European countries/ rather than failing completely)’. Contrastive choice is primarily governed by pragmatic considerations; this is therefore an example of *stipulation*.

Of course, many aspects of the initial experimental condition will not affect the truth of the counterfactual (whether Germany won the war on a Tuesday or a Wednesday, for example), so there is no need for the initial condition of an implied experiment to be infinitely detailed. Still, it is fair to say that the use of historical counterfactuals will generally require careful specification of the initial situation, above and beyond the bare outline sketched by the antecedent in question.

The second and third features of the construction of IC serve to limit the pragmatic emphasis of the first. Second, whilst we are free to stipulate what the ‘absence’ of *c* amounts to, *certain interpretations are more natural than others*. In particular, recall from the discussion of precipitants and preconditions that ‘absence’ had to be interpreted in an appropriate way, if CM was to be able to subsume the insights from that alternative analysis of causal importance. In that case, it was envisaged that if cause *c* was not the case, another event *d* would most likely have performed as a near functionally equivalent cause. The ‘absence of *c*’ should, in that case, be interpreted as the positive instantiation of *d*.

Third, Lewis has argued that there are premises (or features, in terms of IE) which we are *not* free to simply stipulate away. Such features are empirical facts grounding the truth of the counterfactual, and so must be included in the specification of the initial conditions. Lewis’ example¹ ‘if I had looked in my pocket, I would have found a penny’ should not be resolved by stipulating whether the counterfactual contains the fact that ‘there was (or was not) a penny in my pocket’. Rather, whether or not there was a penny is an empirical fact which should be incorporated into the counterfactual, and which should therefore ground the truth or falsity of that counterfactual.

The correct response is that counterfactuals *can* be stipulated in any way, but that excessive manipulation produces odd counterfactuals, and negates the purpose of counterfactuals such as Lewis’. How is ‘excessive manipulation’ to be understood? Primarily, I suggest, according to the idea that, *implied experiments should not, in general, incorporate those facts which are in logical or causal contradiction with the explicit antecedent*. This principle is hereafter referred to as ‘no causal contradiction’. This idea provides a possible definition of cotenability: cotenable features are those

¹ Lewis, *Counterfactuals*, 68

which are not in logical or causal contradiction with the antecedent. This is why DEM₂ should be rejected: ‘Britain is now a democracy’ is an effect of ‘Germany won the war’ and the other two causes mentioned, so should be rejected. (It is likely that Goodman could not accept this rather simple answer due to prior philosophical commitments, and due to the emphasis on linguistic, rather than ontic, items.)

The ‘no causal contradiction’ rule leads to two difficulties. First, whilst ‘no causal contradiction’ clearly establishes the outlines for formulating a satisfactory experimental situation, the problem of analysis – that counterfactual terminology is required in the explication of the condition – reoccurs at this point. One of Goodman’s problems with cotenability was that ‘to establish any counterfactual, it seems that we first have to determine the truth of another’², since his definition of cotenability was that cotenable premises ‘would not be true if A [the antecedent] were’³.

‘No causal contradiction’ is not so blatant; it defines cotenable features as those which are not in contradiction with the antecedent. Yet, once we note that the antecedent doesn’t *actually* have any effects, it is clear that the circularity in the definition remains. To properly specify ‘no causal contradiction’, we must state that premises which fail the test are *those which are in contradiction with the effects the antecedent would have had*. A non-circular definition of cotenability in terms of PW can be provided. However, I will argue that the *formal* nature of PW counts against the *substance* of such an answer (just as, in chapter eight, I argued that the formal nature of PW counted against its ability to ground counterfactual decidability).

Lewis defines cotenability of x with antecedent Φ at a world i ‘if and only if x holds throughout some Φ -permitting sphere around i ; say also that x is cotenable with Φ at i if x holds throughout every sphere around i , whether or not Φ is entertainable.’⁴ The latter clause refers to the cotenability of a necessary truth with any antecedent. The former, of more interest to present purposes, states (roughly) that x is cotenable iff x is

² Goodman, 16

³ *Ibid.*, 18

⁴ Lewis, *Counterfactuals*, 69

present in the nearest Φ world. However, whilst PW provides the form of an answer to the problem of cotenability, it does not provide any content: despite Lewis' claim that 'a metalinguistic theorist who rejects the foundations of my definition of cotenability may yet find that the definition offers some guidance about which premises are cotenable with which antecedents'⁵. To demonstrate this inability, it is sufficient to show that where we do supply some minimal content to 'similarity', it leads us astray in judgements of cotenability.

Heller⁶ argues that it is not the *most similar* worlds, but the *relevant* ones which should be used to judge cotenability (and whilst relevance is certainly not unrelated to similarity, it cannot be reduced to it). His example is 'if quail had not made tracks there, there would not have been any tracks of that kind there'. The counterfactual is surely false: for maybe a pheasant would have produced tracks of that kind. Yet a world where the quail is absent requires less of a 'miracle' (and is therefore more similar) than a world where the quail is absent *and* a pheasant appears. Heller's example shows that the antecedent world is not so much that which is closest to our own, as simply that which is *relevant*. And, given that such a notion of 'relevance' is a primitive notion, PW can provide no answers to any particular question of cotenability. This emphasis upon relevance fits well with the emphasis I have placed upon pragmatic *stipulation* of an explicit experimental condition as part of IE.

A further example is provided by Kazez⁷: Stressa gives birth to an intellectually disabled child. Stressa is under a great deal of stress, she drinks, she smokes, and she is poor. The only possible medical reason for the birth of an intellectually disabled child is that Stressa drank. Yet, using the PW account of counterfactuals, it would turn out that stress is a relevant cause. This is so because it is plausible to suppose that the closest (most similar) worlds in which Stressa's life is not stressful are also those in which Stressa did not drink. Therefore, according to PW, the counterfactual 'if Stressa's pregnancy had not been stressful, it would not have produced the birth of an intellectually disabled child' is true. Once again, the lesson is that the closest

⁵ *Ibid.*, 70

⁶ Mark Heller, 'Might Counterfactuals and Gratuitous Differences' in *Australasian Journal of Philosophy* 73(1) (1995), 93-94

⁷ Jean Kazez, 'Can Counterfactuals Save Mental Causation' in *Australasian Journal of Philosophy* 73(1) (1995), 84

worlds are not necessarily those worlds which should be used to ground the counterfactual. The relevant world would have been one wherein Stressa was stressed, but did not drink. ‘Stressa’ also demonstrates the relevance of a notion of ‘no causal contradiction’, which provides the intuitively correct response in this case. The initial condition should include all facts which might affect the outcome, so long as they are not in causal contradiction with the antecedent: we should, therefore, include Stressa’s drinking in the experimental situation.

The ‘no causal contradiction’ principle leads to a further problem. I noted that the principle holds only *in general*. Consider again DEM: should we not, therefore, include in the experimental situation the fact that the Russians had captured of the Reichstag in 1945, whose effects do, after all, contradict the antecedent? Or Germany’s failure in the Battle of Britain? Certainly, such scenarios seem, at best, implausible: to imagine Germany’s ultimate victory is not to also imagine the actual defeats that Germany suffered. Yet, there must be a limit to the ‘no causal contradiction’ rule, or we are faced with a problem which we might term the ‘problem of unravelling’ – that once one fact is supposed to be otherwise, all causes of that fact are also supposed to be otherwise. If this is the case, then past events are counterfactually dependent upon future events: thus counterfactuals always ‘back-track’. This contravenes our usual insistence that, whilst ‘back-tracking’ is possible, it is not the norm.

It may be fair to assume that, for most counterfactuals, it is not relevant whether or not back-tracking occurs. For the purposes of working through a counterfactual outcome using IE we might as well assume that the counterfactual in question hasn’t entirely ‘un-ravelled’, but it doesn’t really matter. However, there are counterfactuals where it does matter, such as:

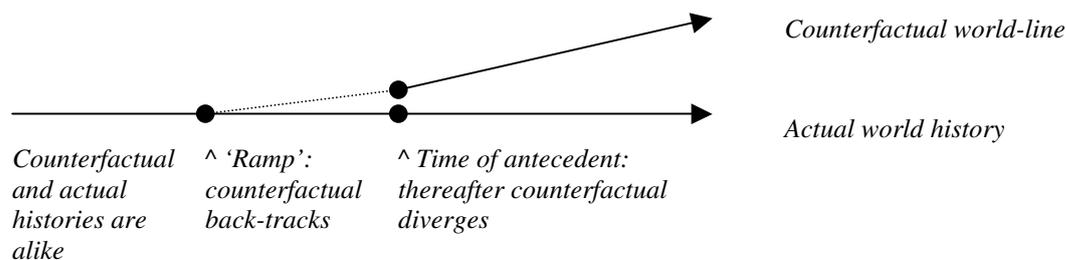
DEM₃: ‘if Germany had won the Second World War, there would have been more radiation present today’

I imagine that such a counterfactual (if true) is made true by the fact that, had Germany won the War, they would have developed the atomic bomb, and it is likely that a limited nuclear war would have taken place. In this scenario, Germany’s

winning the war ‘back-tracks’: the causal history of Germany’s victory counterfactually depends on that victory.

The primary response to the issues of back-tracking and unravelling should be to restate the need for explicit stipulation when providing initial experimental situations; where back-tracking is a possibility, the context should be made clear. This insistence on the need for a pragmatic limitation to back-tracking counterfactuals can be defended, if we recall a feature of contrastive explanations: that a well chosen contrast limits relevant causal history. Those parts of the history of the target-fact which corresponds to causes in the history of the contrast cases is irrelevant, and presupposed. Hence, in a given contrast case, much causal history will be presupposed, and hence not amenable to alteration by back-tracking.

Using the notion of ‘world-lines’, suggested in the previous section, allows a general conception of the role of back-tracking in specification of the counterfactual initial condition. The general intuition to be followed, whose details should be provided by stipulation, is that counterfactual scenarios are supposed to diverge from the actual world at some point shortly before the time of the antecedent. There is a ‘ramp’ leading to the fork pointed to by the antecedent:



In summary, an implied experiment is only possible where a tractable experimental situation is provided. Primarily, this is achieved according to pragmatic stipulation of the necessary details, much of which can be achieved according to a precise contrastive focus. However, two non-pragmatic rules limit the application of pure *stipulation*. First, the ‘absence’ of a cause should be understood such that we give consideration to what is most likely to have taken the place of that cause. Second, in general, facts which are not in logical or causal contradiction to the antecedent should

feature in the experimental situation. To provide a non-circular definition of the initial situation, we must say that an appropriate *world-line* is selected: where ‘appropriate’ is to be understood with reference to the above three features.

Decidability

A major requirement in explicating CM is that counterfactuals be *decidable*. By this I mean that, where there are different opinions about the truth value of a counterfactual, one opinion will be right, and the other opinions will be wrong. Moreover, there will be available, in principle at least, evidence or reasoning which should persuade those others, if they are rational, that their opinion was wrong. The primary reason for adopting the implied experiment account of counterfactuals – as opposed to, particularly, the possible worlds analysis – was to ensure that historical counterfactuals are indeed decidable in this sense. (Note that other conceptions of ‘decidable’ are possible, notably those which are ‘realist’ in a more extreme sense. For example, we might claim that there exists a *truth-maker* for every counterfactual, whether or not anyone can obtain evidence to support it, or reason convincingly about it. There seems no need to attempt to justify such a strong conception as this.)

The core claim is that, where counterfactuals are decidable, IE provides the correct result. IE suggests that to the extent that we can approximately model a historical situation (perhaps using laws, perhaps not), counterfactuals can be correctly framed about that situation. Given the background of the implied experiment account, it is possible to examine in more detail possible reasons why it might *not* be possible to correctly formulate definite counterfactuals. Such reasons can be broadly termed epistemic – limitations resulting from our lack of the relevant kinds of knowledge – pragmatic – where we don’t *want* the counterfactual to be definitely decidable – or ontic – limitations resulting from, in some sense, the counterfactual situation itself. In order to defend IE, for each of the cases below I shall need to either explain how IE is

the best analysis to account for the decidability, or argue that the counterfactual should not be regarded as decidable at all.

““You talk about our position, the left flank weak and the right flank too extended,” he went on. “That’s all nonsense, there’s nothing of the kind. But what awaits us tomorrow? A hundred million most diverse chances which will be decided on the instant by the fact that our men or theirs run or do not run, and that this man or that man is killed.””⁸ – Tolstoy

Epistemic deficiencies may lead to limitations in our ability to decide whether counterfactuals are true or false. It may be that there are *too many variables* to consider in modelling the historical situation. What often amounts to the same thing, it may be that the counterfactual situation heralds *too great a departure* from actuality to be resolvable. In part, problems may be limited by proper specification of the experimental situation, considered previously. However, it is reasonable to suppose that problems may remain. The following example demonstrates both of these epistemic difficulties concerning decidability:

‘If there had been no nation states, there would not have been a revolution’

In addition, we may simply *lack the relevant experimental knowledge* to understand the implied experiment. For example, if we know nothing about the King’s psychology, then the following cannot be judged:

‘If the King hadn’t had a bad day hunting, then he would not have summoned the Estates-General’

Epistemic limitations will affect the decidability of counterfactuals. Under IE, counterfactuals mirror both predictions and experimental simulations, so to the extent that these activities are not possible due to a lack of empirical knowledge, neither will be our judgement of the corresponding counterfactuals. It is the case that we will often claim that ‘it would have been likely that e’, rather than ‘it would have been the case that e’: the likelihood in the former locution is a measurement of subjective probability. However, we should note that IE is more permissive than, in particular,

⁸ Leo Tolstoy, *War and Peace*, 828

ML, in that knowledge of strict historical laws is not necessary to decide upon counterfactuals. IE at least allows counterfactuals to be as decidable as we intuitively understand them to be.

Various heuristics are available in judging epistemically complicated counterfactuals. These heuristics will be similar to the usual experimental heuristics which deliver empirical knowledge. For example, Pork argues that comparable cases are often of use in helping to judge complicated historical counterfactuals:

‘confronted with the problem of justification of the counterfactual statement, historians usually intuitively try to find for a comparison some other *real* situation ... which in some important respect is similar to the *possible* situation reflected in the counterfactual claim ...’⁹

Pork’s strategy bears similarity to Mill’s Method of Agreement, in which the causal factor is discovered by comparing two examples which exhibit the same outcome, and treating the common antecedents as potential causes.

In chapter seven, we saw that Schama believed that events of ‘epochal significance’ need not follow from causes of the same significance: hence such a dramatic event as the French Revolution might follow from seemingly insignificant causes. As I suggested, this idea implies that history might be *chaotic*, in the sense of the mathematical theory of chaos¹⁰.

Chaotic systems are those which require only a small alteration in initial conditions to produce a large change in the end state. The most well known example is, of course, the butterfly’s wings: in a chaotic system, whether or not the butterfly flaps its wings on one side of the world may affect whether or not there is a tornado on the other. It is believed that the earth’s meteorological system *is* a chaotic system. Whether or not history is a chaotic in the precise sense, it seems undeniable that small causes may have large effects. This might be thought to raise difficulties for IE, since it implies that we should have less confidence in our initial assessments of which cause made a greater difference than we actually do.

⁹ A. Pork, ‘Assessing Causal Importance in History’ in *History and Theory* 24 (1985), 66

¹⁰ For an excellent introduction of chaos theory and its philosophical implications, see Peter Smith, *Explaining Chaos* (Cambridge: Cambridge University Press, 1998)

Two remarks are appropriate. Firstly, this is an epistemic difficulty for IE (and hence for CM), in the same sense as the previous examples, and hence raises similar issues. Given that chaotic systems are understood primarily by *modelling*, not by the application of strict laws, IE is a good position to account for any possible counterfactual judgements we *can* make about chaotic systems. Secondly, we can be reasonably confident that most historical systems do not exhibit the sensitive dependence on initial conditions apparent in chaotic systems; if only because when we do highlight such dependence it is as the exception, not the norm. If historical systems were chaotic in a thorough-going way, then the practical applicability of judging historical counterfactuals would indeed be rather drastically restricted.

Weaknesses in the specification of the experimental situation can affect decidability. Whether or not a counterfactual back-tracks affects the outcome of that counterfactual: in the case of DEM₃, for example. Hence, the issue of back-tracking is also relevant to decidability, and so the extent of back-tracking must be made clear as part of the specification of the initial condition. In general, I have claimed that, to the extent that the initial situation is not fully specified, the resulting counterfactual may not be decidable: between certain outcomes, if not completely. As I have argued, the experiment cannot have a determinate outcome if the *experimental scenario is not properly specified*.

Being under-specified in this way is not always a failing of counterfactuals, however. For example, recall the example discussed in chapter seven. The question was ‘why did my friend and I eat Italian (rather than French or Thai)?’. ‘My preference for European food’ limited the possible outcomes from three to two; whilst the competing cause ‘my friend’s preference for pasta’ limited the possible outcomes from three to one. The latter cause in this case would, if judged according to CM, not yield a single outcome. The antecedent ‘if my friend hadn’t wanted to eat pasta then ...’ (supposing that this absence amounted to my friend having no preference) could be completed either by the consequent ‘we would have eaten Thai’, or ‘we would have eaten French’. It must therefore be recognised that certain counterfactuals are framed which can only be properly completed by disjunctive states of affairs. This is not a

limitation of IE, but a recognition that we don't always want counterfactuals to lead to a single outcome.

There are also reasons that counterfactuals may not be decidable which are not due to epistemic shortcomings, nor due to the formulation of the initial counterfactual situation. The relevant causal processes of the implied experiment may be *indeterministic*. If the implied experimental situation involves genuinely indeterministic causal processes, then there will be no definite fact of the matter as to the consequent which should follow the antecedent. At best, such consequents must state that 'there was a x% chance of situation S'. The probability referred to in such counterfactuals will be objective probabilities, rather than the subjective probabilities referred to in cases of epistemic limitation. Just as it is hard to know whether history is really chaotic, it will not always be an easy matter to distinguish empirical deficiencies from genuine indeterminacy of causation.

In short, there are three ways that counterfactuals may not imply a single, definite outcome. In each we would say 'it might have been the case that ...', though 'might' performs a different role in each. Counterfactuals may be hampered by epistemological shortcomings, in which case the 'might' is subjective. Counterfactuals may be hampered by indeterminacy of causal mechanism, in which case the 'might' is objective. And counterfactuals might be framed with the intention of leading to a disjunction of possible outcomes, in which case the 'might' derives from pragmatic features of stipulation of the initial conditions. It is an advantage that IE allows us to recognise these three types of 'might' counterfactuals which do occur in our subjunctive reasoning.

Similarity

‘Similarity, ever ready to solve philosophical problems and overcome obstacles, is a pretender, an impostor, a quack’¹¹

After a long examination of the *counterfactual* element of CM, the remainder of the chapter is devoted to the second potentially problematic concept in CM: *degree of difference*. CM requires a comparative judgement of difference between states of affairs, of the type ‘x is more different to e than y is to e’. It seems philosophically permissible to assume that ‘similarity’ is the antonym of ‘difference’, and therefore to focus analytical efforts on the former notion. Despite my reservations concerning the use of similarity between worlds to ground counterfactual truth values, it is necessary that we consider the concept at this point in CM; even if similarity is to play no substantive role in the counterfactual element. I shall consider two criticisms: first, that similarity should be abandoned in favour of ‘identity in certain respects’; second, that similarity is too subjective a notion to ground the normative judgements required by CM.

We need to know whether similarity is a concept which can be relied upon to do the philosophical work required by CM. Some examples are certainly easier to accept than others; notably, situations where quantitative measurement is available and appropriate. We can count numbers of votes, numbers of people born and who have died, and government finances: it seems an easy matter to state that the situation where the President gains 45 votes is more similar to the situation where she gains 50 votes than it is to the situation where she gains 35. If it is possible to discern that cause A contributes x to the effect, and cause B contributes y, and if we can quantitatively measure x and y, then the similarity judgement may seem easy to come by.

¹¹ Goodman, 437

However, this is not always (or even often) possible; if CM were limited to quantitative cases it would be a disappointing theory indeed. Of more importance to note, though, is that even where quantitative features are available, it may not be appropriate to rely upon them. To use Martin's example¹²: Senator X controls five votes, Senator Y only his own vote. The counterfactual resulting from Senator X not voting is more different to what actually happened than is the counterfactual resulting from Senator Y not voting to what actually happened. Or is it? What if Senator Y's vote has other sorts of consequences, say that it carries a great deal of respect from the press and public? If that were so, whether the vote(s) delivered by Senator Y or Senator X was more important would depend on which features were deemed to be relevant to the similarity judgement. There is no reason to suppose that quantitative features should automatically take precedence.

The example suggests that, to be able to judge similarity, we need to know which features of the comparison situations are relevant. It might be thought that the logical end-product of such an idea would be the conclusion that similarity does *no* work, relying entirely on prior judgements of relevance. This is the conclusion of Goodman's reductive analysis of similarity¹³.

Goodman provides seven 'strictures' on similarity, but there is a common theme of reductive analysis to his critique. This is that similarity cannot ground philosophical analyses, or solve philosophical problems; rather, 'similarity' itself depends on other philosophical concepts to the extent that it is philosophically superfluous. In particular, Goodman argues that *similarity between tokens x and y is only applicable once we also state the respect in which x and y are alike*. But, to state the common attribute is to do all the descriptive work; to state in addition that the tokens are similar is unnecessary.

Goodman's example uses tokens of the letter 'a': letters which are of different sizes, fonts, upper or lower case, and so forth (A, a, a, **A**). We wish to say that these tokens are all similar. Yet without specifying the sense in which the tokens are similar, the claim is empty. What is comprehensible (is not empty of content) is the statement

¹² Martin 'Causes, Conditions and Causal Importance', 71

¹³ Goodman, 437-446

that ‘all the tokens are similar in being ‘a’s’. However, this is simply a long winded (and philosophically misleading) way of saying ‘all the tokens *are* ‘a’s’.

This has a clear application to the material in case studies concerning ‘important properties’. Recall that Lefebvre and others based, to some extent, their judgement of causal importance upon the relative importance of the different properties of the event to be explained. For example, Lefebvre maintained that it was pertinent that the revolution was instigated ‘from below’, and that it was violent. If Goodman is correct, specification of ‘pertinent’ features entirely sets the respect by which the historical (and counterfactual) situations are to be compared. Similarity, and hence CM, can do no work in establishing anything about the importance of a cause; since once the respect has been specified, the situations are either alike or not alike.

We need to recognise that judgements of similarity must be, and are, constrained by specification of the ‘respect’ according to which the items are to be compared. Indeed, CM incorporates this idea directly, through the necessity for an event, *e*, which ‘frames the question of *c*’s importance’. Given that ‘*e*’ is to be understood contrastively, the respect according to which the similarity judgement is to be made can be quite precise. The relevant dimension of similarity is that by which the target fact and the contrasts differ.

For example, suppose that we wish to assess the importance of *c_e*: ‘France not having an entrepreneurial culture in the late eighteenth century’. How similar is the outcome resulting from the absence of this cause to the actual outcome? It is almost impossible to say, and by extension, almost impossible to compare to another judgement of similarity resulting from the absence of a competing cause. However, we are given more than this to work with according to CM. If the contrastive effect which frames the judgement of *c_e*’s importance is ‘why did France, rather than Britain, not go through an industrial revolution in the eighteenth century?’, then the relevant aspects for the similarity judgement will be specified: they are those by which France and Britain differ. It is likely that the resulting similarity judgement will judge *c_e* to be highly important, since the counterfactual situation draws close to that of the foil: Britain’s being industrialised. However, if the contrastive effect

which frames the judgement is ‘why did France, rather than Britain, go through a revolution?’, then it is unlikely that the counterfactual difference will be so great.

What I will dispute, with respect to Goodman’s argument, is the contention that similarity does no work at all, and should be replaced by ‘identity in certain respects’. Whilst examples may seem to support the claim that similarity reduces entirely to identity of type, I believe that this is generally not the case, and in particular, not the case when comparing complex states of affairs of the sort required by CM. Consider two examples: (1) a similarity comparison between two objects; the Goodman-like claim being that the objects are similar only to the extent that they share some same property, say, colour. (2) Comparison of two historical situations (such as the French nation with and without a pre-existing entrepreneurial culture); the Goodman-like claim being that the situations are similar only in respect of, say, their economic aspects.

The claim I wish to make is that *even granting ‘Goodman-like’ restrictions*, similarity is still doing real work. This is because, whilst the ‘respect’ in which the items are similar has been stated, these ‘respects’ are still too broad to eliminate genuine use of similarity. Shared colour, in the case of (1), can be thought to refer to many more detailed properties; thus to claim that object x is more similar to z than y is to z in respect of colour is not to analyse away ‘similarity’. Likewise, in the case of (2), ‘economic aspects’ is a broad term, encompassing trade, fiscal policy, industry, income, and more.¹⁴

One option for a reductive critic of the use of similarity judgements would be to move to a more extreme position. A defender of Goodman could allow that the above examples (colour and economic aspects) indeed do not serve to analyse away ‘similarity’, but are therefore incomplete as they stand. What would be needed are examples which specify the exact properties which the tokens share, rather than such general aspects as ‘colour’. I grant that this is a possible response; however, such a

¹⁴ It is interesting to note that parallel arguments can be constructed for concepts other than ‘similarity’. The question of whether it is possible to point to ‘progress in history’ is often raised; a common reply is to note that to judge progress we have to specify the kind of ‘progress’ we mean. This is surely correct – we do require some sort of idea as to what our attention should be focused upon. But, as I have argued above, this is not to analyse away ‘progress’; for no matter what aspect is selected, there will still be a number of more detailed aspects which still require weighting according to a more detailed, but still irreducible, standard of ‘progress’.

position is far less intuitively appealing than Goodman's original position, and it might be doubted whether it could be carried through in practice. It seems that even if we are to accept Goodman's request that similarity claims be qualified, a substantive role for similarity judgements remains.

There is a further reason why we should be dubious of Goodman's reductive argument. Goodman's attempt to analyse 'similarity' in terms of 'type' appears to mirror Wittgenstein's famous 'rule-following' argument, wherein the status of rules (including those constituting what we mean by 'types') are challenged. Many have interpreted Wittgenstein's argument as an attempt to demonstrate that the successful application of a rule requires an understanding of similarity. This point has been raised by Gentner and Medina:

'[Wittgenstein] argues that following a rule cannot consist simply in being guided by a representation of the rule (e.g. '+2'). Such a representation *by itself* does not determine unequivocally how to continue the number series, for it can be interpreted and applied in various ways (e.g. as 'x + 2 if x < 1000, otherwise x + 4'). Wittgenstein emphasises that rules are not self-interpreting: a rule does not contain within itself what counts as a correct application to each new case. ... Wittgenstein's point is that, in order to be cognitively useful, rules have to be supplemented with standards of similarity for their application.'¹⁵

If Wittgenstein's argument is successful, interpretation of 'type' relies on judgements of 'similarity' just as much as vice versa.

Whilst the notion of similarity has not received a great deal of attention in the philosophical literature, it has attracted a good degree of interest from psychologists. Debate here has focused on whether similarity, or rule-following, is the cognitively prior activity; and (more relevant to CM) whether similarity is a concept that can be used in explanation of cognition and behaviour. Much of the literature has emphasised the fact that similarity judgements are too flexible and unconstrained to ground any other cognitive activity: a claim which, if true, would be problematic in relation to the task which similarity is intended to perform in CM.

¹⁵ D. Gentner and J. Medina, 'Similarity and the Development of Rules' in *Cognition* 65 (1998), 288

In many ways, similarity judgements do appear subject to great variation. After all, there is very little that can be stated axiomatically concerning similarity judgements; Hahn and Chater¹⁶ suggest (the only) three plausible candidates. First, similarity is maximal for identity. Second, similarity is graded. Third, similarity is some function of common properties (the precise function cannot be specified). Within this framework there is room for much flexibility.

The way that similarity judgements change according to an array of pragmatic factors has often been noted in the psychological literature. To summarise: Tversky¹⁷ shows that judgements vary according to what else is in the comparison set. The similarity of a to b can differ when asking ‘how similar is a to b, compared to c’ to ‘how similar is a to b, compared to d’. Tversky also argues that similarity is not transitive, and thus depends on the order the tokens are presented. Sloman and Rips¹⁸ show that judgements vary according to the type of task being performed. Goldstone¹⁹ shows that they vary according to what ‘heading’ the tokens are put under. His example is of ‘children’ and ‘jewellery’; which were thought to be not very similar until placed jointly under the heading ‘things to be rescued from a burning house’. Goldstone also reports that a judge’s age affects similarity, as does their technical expertise.

This could be worrying with regard to CM, due to the possibility that similarity judgements are subject to at least the same variation as are judgements of causal importance. However, we can mitigate the impact of experimental evidence that similarity is context dependent, in two ways.

Firstly, there is opposite empirical evidence: that similarity judgements are – to some extent – ‘hard-wired’. As Goldstone states, ‘it is difficult *not* to notice the similarity between a 400 Hz tone and a 402 Hz tone, or two shades of red’²⁰. It is an oft noted fact in the psychological literature that judgements of categorisation and judgements

¹⁶ Ulrike Hahn and Nick Chater, ‘Similarity and Rules: Distinct? Exhaustive? Empirically Distinguishable?’ in *Cognition* 65 (1998), 206

¹⁷ A. Tversky, ‘Features of Similarity’ in *Psychological Review* 84 (1977): 327-352

¹⁸ S. A. Sloman and L. J. Rips, ‘Similarity as an Explanatory Construct’ in *Cognition* 65 (1998): 87-101

¹⁹ R. L. Goldstone, ‘The Role of Similarity in Categorization: Providing a Groundwork’ in *Cognition* 52 (1994): 125-157

²⁰ Goldstone, 136

of similarity diverge. The best evidence is derived from experiments where the subjects were asked to group objects according to specific criteria. The similarities between the objects which were irrelevant to the subjects' task were shown to still significantly slow such groupings. The conclusion to draw is that similarity judgements are, to some extent, primitive and inescapable.

Secondly, much of the results reported above pose no concerns for similarity judgements, *in the context of CM*. The fact that we have accepted that similarity judgements in CM are conducted 'according to certain respects' makes some of the psychological evidence irrelevant in assessment of CM. We can, and should, take on board both Tversky's claim that similarity judgements vary according to what else is being compared, and Goldstone's claim that similarity judgements vary according to the 'heading' that the items are placed under.

Likewise, Sloman and Rips' claim that similarity judgements vary according to the type of task can be discounted with respect to CM: for, in the context of CM, all similarity judgements *are* performances of the same task. Finally, to the extent that similarity is *theory dependent*, we should not be concerned. For CM aims to show that judgements of importance *are*, in a sense, theory dependent: that our ascriptions of importance can be better or worse depending on our grasp of the system under examination. For the purposes of CM, if the similarity judgements of the expert and the novice diverge, we should trust those of the expert.

I conclude that similarity is a cognitive strategy which cannot be analytically reduced to 'identity in certain respects', and which can be shown through empirical experiment to be pragmatically variable, yet also sufficiently stable. This is especially so given the use to which 'similarity' is put in CM: for the similarity judgements required by CM are constrained to a large degree. There is no one rule to decide similarity; it is precisely this flexibility which allows similarity to be applicable to complicated situations as much as to perceptual aspects of objects.

Conclusions

I have examined a realist method for determining importance; to the extent that this has been successful, I have also described a common way for historical explanations to conflict in a substantive manner going beyond a simple disagreement in explanans. I believe I have demonstrated that it can be successful, even though we must subject that claim to limitations and provisos. However, even modest success is sufficient to refute the sceptical and relativist claims of Mill, Walsh et al concerning importance.

The method has focused on degree of effect as the judgement which underpins claims of importance. 'Degree of effect' has in turn been analysed counterfactually. The counterfactual model of importance (CM) has been shown to subsume competing accounts, and to provide results which generally agree with (and explain) both intuitive and historical claims of importance.

CM required examination of two key terms: 'counterfactual' and 'difference' (or its converse, 'similarity'); the presence of either of these could potentially be regarded as a source of scepticism over CM. I have defended a new explication of counterfactuals, the Implied Experiment account. Using IE, we should regard counterfactuals as referring to an initial experimental situation (specified, in part, by the antecedent), and its subsequent development (where this development is understood by analogy with actual experiments). I have argued that this yields a notion of counterfactuals such that they admit of substantive and rational defence; historical competition over causal importance may also, thereby, be considered in terms of substantive and reasoned debate. I have defended a flexible, yet constrained understanding of similarity.

It can hardly be denied, however, that 'importance' can be used in ways not covered by CM. *Sometimes* morality plays a part: we may select those causes which were most blameworthy, for example. Sometimes, 'intrinsic', non-empirical, reasons are relevant: as in the case of Hegel's historical account. And there will doubtless be other reasons for ascribing importance in particular cases. There is clearly a place for 'personal' importance, in addition to the 'objective' importance explicated above: what is important in an individual's life, for example, is not a matter of consequences,

but of simply what is *felt* to be important. The significance of CM is in demonstrating that there is *an* objective notion of importance, which is used frequently by historians, and which, in a public discipline such as history, can be recommended as an appropriate way for historians to decide upon and to justify judgements of importance.

A recurrent, underlying, issue of Part III has been the nature of philosophical analysis. In particular, there have been two key analytical steps in the preceding argument, each of which could be challenged on the basis of the type of analysis at work. (i) Importance in general was analysed in terms of the degree of effect brought about; and hence in terms of counterfactual difference. (ii) Counterfactual decidability was analysed in terms of an implied experiment. The common question faced has been to what extent we should replace rough and ready judgements with precise rules. I have attempted to do this to some extent, though not to claim that *all* practice must fall under those rules. To this extent, I have attempted to follow the broad recommendations of a ‘family resemblance’ method of analysis.

The thesis as a whole has, I hope, added to the understanding of competition between historical explanations. Treating historical explanation as the activity of tracing *causes* of an aspect of an event, which are *conceptualised* in certain ways, has allowed my treatment to be flexible, yet able to deliver definite results. I have paid particular attention to the role of *causal* claims in leading to explanatory exclusion and, through the Counterfactual Method, to competing claims of historical importance. To be better able to comprehend these two fundamental relations between historical explanations represents advance enough for now. Further investigation could, and should, focus attention on the role of *conceptual categorisation* in competition between historical explanations. But that task will have to wait.

Bibliography

- Achinstein, Peter: 'The Pragmatic Character of Explanation' (originally 1984) in *Explanation*, ed. David-Hillel Ruben (Oxford: Oxford University Press, 1993): 326-344
- Anderson, R. J. and Sharrock, W. W.: 'Understanding Peter Winch' in *Inquiry* 28 (1985): 119-122
- Aristotle: *Metaphysics*, trans. John Warrington (London: Dent, 1956)
- Barnes, Eric: 'Inference to the Loveliest Explanation' in *Synthese* 103 (1995): 251-277
- Beard, Charles: 'That Noble Dream' (originally 1935) in *The Varieties of History*, ed. Fritz Stern (Cleveland: Meridan Books, 1956): 314-328
- Bellaimey, James E.: 'Family Resemblances and the Problem of the Under-determination of Extension' in *Philosophical Investigations* 13(1) (1990): 31-43
- Bromberger, Sylvain: 'An Approach to Explanation' in *Analytical Philosophy, Second Series*, ed. R. J. Butler (Oxford: Oxford University Press, 1965): 72-105
- Carr, David: *Time, Narrative and History* (Indiana: Indiana University Press, 1986)
- Chartier, Roger: *The Cultural Origins of the French Revolution*, trans. Lydia G. Cochrane (Durham, N.C., London: Duke University Press, 1991)
- Cobban, Alfred: *The Social Interpretation of the French Revolution* (Cambridge: Cambridge University Press, 1964)
- Collingwood, R. G.: *The Idea of History* (Oxford: Oxford University Press, 1994; originally 1946)
- Craig, Edward: *Knowledge and the State of Nature* (Oxford: Clarendon Press, 1990)
- Danto, Arthur C.: *Analytical Philosophy of History* (Cambridge: Cambridge University Press, 1965)
- Davidson, Donald 'Actions, Reasons and Causes' (originally 1963) in Donald Davidson, *Essays on Actions and Events* (Oxford: Clarendon Press, 1980): 3-19
- Davidson, Donald: 'Freedom to Act' (originally 1973) in Donald Davidson, *Essays on Actions and Events* (Oxford: Oxford University Press, 1980): 63-81
- Davidson, Donald 'Hempel on Explaining Actions', (originally 1976) in Donald Davidson, *Essays on Actions and Events* (Oxford: Oxford University Press, 1980): 261-275
- Dray, W. H.: 'Concepts of Causation in A. J. P. Taylor's Account of the Origins of the Second World War' in *History and Theory* 17 (1978): 149-175
- Dray, W. H.: 'On the Idea of Importance in History' in W. H. Dray, *On History and Philosophers of History* (Leiden: E. J. Brill, 1989): 73-91

- Dray, W. H.: “‘Explaining What’ in History’ in *Theories of History*, ed. Patrick Gardiner (New York: Free Press, 1959): 403-408
- Dretske, Fred I.: ‘Contrastive Statements’ in *Philosophical Review* 81 (1972): 411-437
- Edgington, Dorothy: ‘Counterfactuals and the Benefit of Hindsight’, *unpublished* (2002)
- Eisenstein, Elizabeth: ‘Who Intervened in 1788? A Commentary on “The Coming of the French Revolution”’ in *American Historical Review* 71(1) (1965): 77-103
- Ferguson, Niall (ed.): *Virtual History: Alternatives and Counterfactuals* (London: Macmillan Publishers, 1988)
- Feyerabend, Paul: ‘On the ‘meaning’ of scientific terms’ in *Journal of Philosophy* 61 (1964): 266-273
- Feyerabend, Paul: ‘Consolations for the Specialist’ in *Criticism and the Growth of Knowledge*, eds. Imre Lakatos and Alan Musgrave (London: Cambridge University Press, 1970): 197-200
- Feyerabend, Paul: ‘On the Critique of Scientific Reason’, in *Method and Appraisal in the Physical Sciences*, ed. Colin Howson (London: Cambridge University Press, 1976): 309-339
- Fine, Kit: ‘Review of David Lewis, Counterfactuals’ in *Mind* 84 (1975): 451-458
- Fischer, D. H.: *Historians’ Fallacies: Towards a Logic of Historical Thought* (London: Routledge, 1971)
- Foucault, Michel: ‘Nietzsche, Genealogy, History’ in Michel Foucault, *Language, Counter-Memory, Practice*, ed. Donald Bouchard (Oxford: Blackwell, 1977): 139-164
- Furet, François: *Interpreting the French Revolution*, trans. Elborg Forster (Cambridge: Cambridge University Press, 1981; originally 1978)
- Garfinkel, Alan: *Forms of Explanation: Rethinking the Questions in Social Theory* (New Haven, London: Yale University Press, 1981)
- Gentner, D. and Medina, J.: ‘Similarity and the Development of Rules’ in *Cognition* 65 (1998): 263-297
- Goldstone, Robert L.: ‘The Role of Similarity in Categorisation: Providing a Groundwork’ in *Cognition* 52 (1994): 125-157
- Goodman, Nelson: *Fact, Fiction and Forecast* (Cambridge, Mass: Harvard University Press, 1983; originally 1954)
- Gorman, J. L.: ‘Objectivity and Truth in History’, in *History and Theory*, eds. Brian Fay, Philip Pomper, Richard T. Vann (Oxford: Blackwell Publishers, 1998): 320-340
- Graham, Gordon: *Historical Explanation Reconsidered* (Aberdeen: Aberdeen University Press, 1983)
- Hahn, Ulrike and Chater, Nick: ‘Similarity and Rules: Distinct? Exhaustive? Empirically Distinguishable?’ in *Cognition* 65 (1998): 197-230

- Hammond, Michael: 'Weighting Causes in Historical Explanation' in *Theoria* 43 (1977): 103-128
- Hart, H. L. A. and Honoré, A. M.: 'Causal Judgement in History and in the Law', in *Philosophical Analysis and History*, ed. W. H. Dray (New York: Harper and Row, 1966): 213-237
- Hawthorn, Geoffrey: *Plausible Worlds: Possibility and Understanding in History and the Social Sciences* (Cambridge: Cambridge University Press, 1991)
- Hegel, Georg Wilhelm Friedrich: *The Philosophy of History*, trans. J. Sibree (New York: Dover Publications, 1956; originally 1840)
- Hegel, Georg Wilhelm Friedrich: *Encyclopedia of the Philosophical Sciences*, trans. W. Wallace (Oxford: Oxford University Press, 1975; originally 1817)
- Hegel, Georg Wilhelm Friedrich: *Elements of the Philosophy of Right*, trans. H. B. Nisbet, ed. Allen Wood (Cambridge: Cambridge University Press, 1991; originally 1820)
- Hegel, Georg Wilhelm Friedrich: *Introduction to the Philosophy of History*, trans. Leo Rauch (Hackett, 1995; originally 1840)
- Heller, Mark: 'Might Counterfactuals and Gratuitous Differences' in *Australasian Journal of Philosophy* 73(1) (1995): 91-101
- Hempel, Carl G.: 'The Function of General Laws in History' in *Journal of Philosophy* 39 (1942): 35-47
- Hempel, Carl G.: 'Aspects of Scientific Explanation', in Carl G. Hempel, *Aspects of Scientific Explanation, and Other Essays in the Philosophy of Science* (New York: Free Press, 1965): 331-496
- Hilton, Denis and McClure, John: 'Are Goals or Preconditions Better Explanations? It depends on the question' in *European Journal of Social Psychology* 28 (1998): 897-911
- Hitchcock, Christopher: 'Contrastive Explanation and the Demons of Determinism' in *British Journal for the Philosophy of Science* 50(4) (1999): 585-612
- Hooper, John: 'Changing Perceptions of Jules Michelet as Historian: History Between Literature and Science, 1831-1874' in *Journal of European Studies* 23(3) (1993): 283-298
- Katz, Bernard: 'Kim on Events' in *Philosophical Review* 87 (1978): 427-441
- Kazez, Jean: 'Can Counterfactuals Save Mental Causation' in *Australasian Journal of Philosophy* 73(1) (1995): 71-90
- Kellner, Hans: 'Beautifying the Nightmare: The Aesthetics of Postmodern History' in *Strategies: A Journal of Theory, Culture and Politics* 4 (1991): 289-313
- Kim, Jaegwon: 'Noncausal Connections' in *Nous* 8 (1974): 41-52
- Kim, Jaegwon: 'Explanatory Realism, Causal Realism and Explanatory Exclusion' in *Midwest Studies in Philosophy* 12 (1988): 225-239
- Kim, Jaegwon: 'Causation, Nomic Subsumption and the Concept of an Event' (originally 1973) in Jaegwon Kim, *Supervenience and Mind: Selected*

- Philosophical Essays*, ed. Ernest Sosa (Cambridge: Cambridge University Press, 1993): 3-21
- Kim, Jaegwon: 'Mechanism, Purpose and Explanatory Exclusion' (originally 1989) in Jaegwon Kim, *Supervenience and Mind: Selected Philosophical Essays*, ed. Ernest Sosa (Cambridge: Cambridge University Press, 1993): 237-264
- Kitcher, Philip and Salmon, Wesley: 'Van Fraassen on Explanation' (originally 1987) in Wesley Salmon, *Causality and Explanation* (Oxford: Oxford University Press, 1998): 178-189
- Kuhn, Thomas: *The Structure of Scientific Revolutions* (Chicago: Chicago University Press, 1962)
- Lakatos, Imre: 'Falsification and the Methodology of Scientific Research Programmes' in *Criticism and the Growth of Knowledge*, eds. Imre Lakatos and Alan Musgrave (London: Cambridge University Press, 1970): 91-195
- Lakatos, Imre: 'History of Science and its Rational Reconstructions' in *Method and Appraisal in the Physical Sciences*, ed. Colin Howson (London: Cambridge University Press, 1976): 1-39
- Laudan, Larry: 'Progress or Rationality? The Prospects for Normative Naturalism' (originally 1987) in *Philosophy of Science*, ed. David Papineau (Oxford: Oxford University Press, 1996): 194-214
- Lefebvre, Georges: *The Coming of the French Revolution*, trans. R. R. Palmer (Princeton: Princeton University Press, 1947)
- Lewis, David: *Counterfactuals* (Oxford: Blackwells, 1973)
- Lewis, David: 'Counterfactual Dependence and Time's Arrow' (originally 1979) in David Lewis, *Philosophical Papers Volume II* (Oxford: Oxford University Press, 1986): 32-66
- Lewis, David: 'Causal Explanation' (originally 1986) in *Explanation*, ed. David-Hillel Ruben (Oxford: Oxford University Press, 1993): 182-206
- Lipton, Peter: *Inference to the Best Explanation* (London: Routledge, 1991)
- Lipton, Peter: 'Contrastive Explanation' (originally 1990) in *Explanation*, ed. David-Hillel Ruben (Oxford: Oxford University Press, 1993): 207-227
- Lipton, Peter: 'Making a Difference' in *Philosophica* 51(1) (1993): 39-54
- Mackie, J. L.: *Truth, Probability and Paradox* (Oxford: Clarendon Press, 1973)
- Manicas, Peter: 'Review Article: Theda Skocpol, States and Social Revolutions' in *History and Theory* 20(2) (1981): 204-218
- Markwick, P.: 'Interrogatives and Contrasts in Explanation Theory' in *Philosophical Studies* 96 (1999): 183-204
- Marras, Ausonio: 'Kim's Principle of Explanatory Exclusion' in *Australasian Journal of Philosophy* 76(3) (1998): 439-451
- Martin, Raymond: 'On Weighting Causes' in *American Philosophical Quarterly* 9(4) (1972): 291-299
- Martin, Raymond: 'Causes, Conditions and Causal Importance' in *History and Theory* 21 (1982): 53-74

- Martin, Raymond: *The Past Within Us* (Princeton: Princeton University Press, 1989)
- McCarney, Joseph: *Hegel on History* (London: Routledge, 2000)
- Michelet, Jules: *History of the French Revolution*, trans. Charles Cocks, ed. Gordon Wright (Chicago: University of Chicago Press, 1967; originally 1853)
- Mill, John Stuart: *A System of Logic, Ratiocinative and Inductive* (Charlottesville: Lincoln-Rembrandt, 1996; originally 1843)
- Murray, Bill: 'Review Article' [of Schama's *Citizens*] in *Australian Journal of French Studies* 26(1) (1989): 100-107
- Nichols, Shaun; Stich, Stephen; Leslie, Alan; Klein, David: 'Varieties of Off-Line Simulation' in *Theories of Theories of Mind*, eds. P. Carruthers and P. K. Smith (Cambridge: Cambridge University Press, 1996): 39-74
- Novick, Peter: *That Noble Dream: The "Objectivity Question" and the American Historical Profession* (Cambridge: Cambridge University Press, 1988)
- Pelczar, Michael: 'Wittgensteinian Semantics' in *Nous* 34 (4) (2000): 483-516
- Pork, A.: 'Assessing Causal Importance in History' in *History and Theory* 24 (1985): 62-69
- Rappaport, Steven: 'Inference to the Best Explanation: is it Really Different from Mill's Methods?' in *Philosophy of Science* 63(1) (1996): 65-80
- Rawls, John: *A Theory of Justice* (Oxford: Oxford University Press, 1971)
- Rigby, S. H.: 'Historical Causation: Is One Thing more Important than Another?' in *History* (1995): 227-242
- Risjord, Mark: 'No Strings Attached: Functional and Intentional Action Explanations' in *Philosophy of Science* 66(3 Supplement) (1999): S299-S313
- Ruben, David-Hillel: 'Explaining Contrastive Facts' in *Analysis* 47 (1987): 35-37
- Ruben, David-Hillel: 'Singular Explanation and the Social Sciences' in *Philosophy* 65 (1990): 95-117
- Ruben, David-Hillel: *Explaining Explanation* (London: Routledge, 1990)
- Salmon, Wesley C.: 'A Third Dogma of Empiricism' in *Basic Problems in Methodology and Linguistics*, eds. R. Butts and J. Hintikka (Dordrecht: D. Reidel, 1977): 149-166
- Salmon, Wesley C.: 'Scientific Explanation: Causation and Unification' in *Critica* 22 (1990): 3-21
- Salmon, Wesley C.: 'The Importance of Scientific Understanding' in Wesley Salmon, *Causality and Explanation* (Oxford: Oxford University Press, 1998): 79-91
- Schama, Simon: *Citizens: a Chronicle of the French Revolution* (London: Viking, 1989)
- Sehon, Scott: 'Deviant Causal Chains and the Irreducibility of Teleological Explanation' in *Pacific Philosophical Quarterly* 78(2) 1997: 195-213
- Sintonen, Matti: 'Why Questions and Why Just Why-questions?' in *Synthese* 120 (1999): 125-135

- Skocpol, Theda: *States and Social Revolutions: A Comparative Analysis of France, Russia and China* (Cambridge: Cambridge University Press, 1979)
- Sloman, S. A. and Rips, L. J.: 'Similarity as an Explanatory Construct' in *Cognition* 65 (1998): 87-101
- Smith, Peter: *Explaining Chaos* (Cambridge: Cambridge University Press, 1998)
- Spitzer, Alan B.: 'Narrative Problems: the Case of Simon Schama' in *Journal of Modern History* 65(1) (1993): 176-192
- Stalnaker, Robert: *Inquiry* (Cambridge, Mass: Bradford Books, 1984)
- Stich, Stephen and Nichols, Shaun: 'Folk Psychology: Simulation or Tacit Theory' in *Mind and Language* 7 (1992): 35-71
- Stich, Stephen and Weinberg, Jonathan M.: 'Jackson's Empirical Assumptions' in *Philosophy and Phenomenological Research* 62 (3) (2001): 637-643
- Stich, Stephen: 'Normativity and Epistemic Intuitions', *unpublished* (2002)
- Swinburne, Richard: 'Are Mental Events Identical with Brain Events' in *American Philosophical Quarterly* 19 (1982): 173-182
- Temple, Dennis: 'The Contrast Theory of Why-Questions' in *Philosophy of Science* 55 (1988): 141-151
- Thagard, Paul: 'Probabilistic Networks and Explanatory Coherence' in *Cognitive Science Quarterly* 1 (2000): 91-114
- Tolstoy, Leo: *War and Peace*, trans. Louise and Aylmer Maude (Oxford: Oxford University Press, 1991; originally 1868-9)
- Tversky, A.: 'Features of Similarity' in *Psychological Review* 84 (1977): 327-352
- Van Fraassen, Bas C.: *The Scientific Image* (Oxford: Clarendon, 1980)
- Walsh, W. H.: *An Introduction to the Philosophy of History* (London: Hutchinson, 1951)
- Walsh, W. H.: 'Colligatory Concepts in History' (originally 1967) in *The Philosophy of History*, ed. Patrick Gardiner (Oxford: Oxford University Press, 1974): 127-144
- White, Hayden: *Metahistory: the Historical Imagination in Nineteenth-Century Europe* (Baltimore, London: Johns Hopkins University Press, 1973)
- White, Hayden: 'The Value of Narrativity in the Representation of Reality' (originally 1980) in Hayden White, *The Content of the Form: Narrative Discourse and Historical Representation* (Baltimore: John Hopkins University Press, 1987): 1-10
- White, Hayden: 'The Historical Text as Literary Artifact' (originally 1978) in *History and Theory*, eds. Brian Fay, Philip Pomper, Richard T. Vann (Oxford: Blackwell Publishers, 1998): 15-33
- White, Morton: *Foundations of Historical Knowledge* (New York: Harper and Row, 1965)
- Winch, Peter: *The Idea of a Social Science* (London: Routledge, 1958)

Wittgenstein, Ludwig: *Philosophical Investigations*, trans. G. E. M. Anscombe
(Oxford: Blackwell, 1953)

Woodward, James: 'A Theory of Singular Causal Explanation' (originally 1984) in
Explanation, ed. David-Hillel Ruben (Oxford: Oxford University Press, 1993):
246-274