

O N S O C I A L F A C T SDETAILED TABLE OF CONTENTS

	PAGE
<u>CHAPTER ONE: INTRODUCTION</u>	1
1. <u>The topic and approach of this study</u>	1
2. <u>On saying what social sciences are about:</u> <u>some preliminary notions</u>	7
i. Global versus piecemeal theory.....	7
ii. Some criteria for judging global theories.....	8
iii. Some nontrivial forms of global theory	9a
iv. Criteria for judging piecemeal theories.....	11
 <u>CHAPTER TWO: MAX WEBER AND 'SOCIAL ACTION'</u>	13
1. <u>On 'social action' in general</u>	13
i. Motives for speaking of 'social action'.....	13
ii. The proliferation of 'social action' talk.....	16
2. <u>Weber as global theorist and analytic individualist</u>	19
3. <u>What Weber's 'social action' is</u>	21
i. Action	22
ii. 'Social action'.....	26
4. <u>What Weber's 'social action' is not</u>	28
5. <u>Possible cases of social action</u>	32
6. <u>Weberian social action and collectivity concepts</u>	34
i. A note on collectivity concepts	34
ii. A solitary social actor.....	35
iii. The Weberian social action thesis.....	36
iv. A formulation rejected.....	38
v. Another formulation rejected.....	39
vi. Necessary conditions and states of nature.....	44
vii. Conclusions.....	46
7. <u>Social action and the subject matter of sociology</u>	48
i. Introduction	48
ii. The definition of 'suicide'	50
iii. 'De facto' social action	51
iv. The Durkheimian approach to suicide	53
v. Approaches to suicide	55
8. <u>Weber's 'social action' as a scientific concept</u>	58

9. <u>Conclusions</u>	61
i. Weber's notion of social action.....	61
ii. A result concerning global theories.....	63
iii. An outstanding query.....	64
 <u>CHAPTER THREE: ACTION AND THE SOCIAL</u>	66
1. <u>Introduction: two claims about action</u>	66
i. Winch <u>contra</u> Weber	66
ii. Two claims about action	66
iii. Some initial defenses of Weber	69
2. <u>Elements in the argument from 'meaningful behaviour'</u> <u>to social settings</u>	71
i. 'Meaningful behaviour' and meaningful words.....	71
ii. Meaningful words and rules: a crucial assumption.....	73
iii. Concepts and rules.....	74
iv. The definition of 'meaningful behaviour'.....	75
3. <u>Rule-following as problematic</u>	79
i. Introduction	79
ii. 'Going on the same'	79
iii. Stipulating one's future path	80
iv. Rules and rule-formulations.....	81
v. Meaning and doing	88
vi. Paradoxes of conformity	96
4. <u>Winch's positive account of rule-following</u>	100
i. Introduction: Winch's example.....	100
ii. 'In principle discoverable'.....	103
iii. Social settings and rule-following.....	107
iv. Established rules and rule-following.....	110
v. The content of meaningful behaviour.....	114
5. <u>Further considerations on the social nature of action</u> ...	117
i. An argument for the social nature of all action.....	117
ii. Are all actions social in Weber's sense?.....	119
iii. Conclusions: two further concepts of 'social action'.....	120

CHAPTER FOUR: <u>A DURKHEIMIAN CONCEPTION OF A SOCIAL PHENOMENON</u>	123
1. <u>Introduction</u>	123
i. Contents of this chapter.....	123
ii. Durkheim's project	125
2. <u>Durkheim's basic conception of a social fact</u>	123
i. The basic conception versus the 'definitions'.....	123
ii. Towards an example.....	132
3. <u>Collective practices</u>	138
i. Introductory: examples	138
ii. On the logic of ascriptions of collective practice....	145
iii. Elements for a definition of 'collective practice'....	148
iv. A working definition of 'collective practice'.....	157
4. <u>In defense of common knowledge</u>	167
i. Is common knowledge possible?	167
ii. Knowledge	168
iii. Knowledge and belief	170
iv. An example.....	174
v. Objections reconsidered	175
vi. On some alternative accounts of common knowledge phenomena.....	177
vii. Conclusions: on common knowledge and collective practices.....	181
5. <u>A social group produces social phenomena</u>	182
i. Introduction	182
ii. A general account of production by society.....	183
iii. A 'special' account of production by a social group...	187
iv. A contrast case.....	189
v. Reasons as productive forces.....	191
vi. 'Basic reasons'.....	193
vii. The reality of our account.....	199
6. <u>A Durkheimian notion of social phenomenon</u>	201
i. The definition	201
ii. Some examples	202
iii. Conclusions.....	204
CHAPTER FIVE: <u>ABOUT CONVENTIONS</u>	206
1. <u>Introduction: A suggestion about conventions</u>	206
2. <u>David Lewis on convention: overview</u>	208
i. Conventions as (L-) regularities.....	208
ii. The basic preference structure underlying conventions: Lewis's final account.....	211
iii. Lewis's argument as a whole.....	213
iv. A note on game-theory.....	216

3. <u>Lewis on co-ordination problems: a critique</u>	222
i. Introductory	222
ii. Lewis's definition of 'co-ordination problem'....	223
iii. Shifts in Lewis's thinking, and some queries....	231
iv. Lewis's examples: two distinguishable structures.....	233
v. Lewis's arguments for his definition of 'co-ordination problem'	237
vi. Conclusions on the structure of co-ordination problems.....	259
4. <u>Lewis on 'convention': a critical commentary</u>	259
i. Conventions as 'solutions' to co-ordination problems	259
ii. On conventions started by agreement: a possible amendment.....	264
iii. Lewis's 'common knowledge' condition.....	267
iv. The arbitrariness of conventions.....	272
v. Conventions as norms.....	274
vi. Conformism	275
vii. Conventions and language	277
viii. Conclusions on Lewis's concept.....	279
5. <u>Other notions of convention</u>	281
i. Critics of Lewis	281
ii. Conventions and conformity	286
iii. A nonconformist sketch of conventions.....	289
6. <u>Conventions, social facts, and the subject matter of social sciences: some claims in favour of conventions</u>	290
7. <u>Social groups revisited</u>	
i. Conventions and social groups.....	296
ii. The concept of a social group.....	298
 <u>CHAPTER SIX: ON SOCIAL FACTS</u>	303
1. <u>A Theory of Socialness</u>	303
2. <u>'Mental connectedness'</u>	312
3. <u>Causal connections between persons</u>	314
4. <u>On characterizing social sciences in general</u>	316
 <u>B I B L I O G R A P H Y</u>	317

ON SOCIAL FACTS

BY

MARGARET P. GILBERT

ST. HILDA'S COLLEGE

Submitted for the degree of Doctor of Philosophy

Trinity Term, 1978

a b s t r a c t

Four concepts are considered in relation to the question: can an illuminating characterization of the social sciences be given in terms of one concept of a relatively natural kind of thing? Weber's concept of 'social action' provides neither a general characterization, nor an important partial account, or so I argue after examining its relation to collectivity concepts, to suicide studies, and to standard desiderata for scientific concepts. I next assess the notion of 'meaningful' action. Peter Winch claims that such action is always 'social' in some sense, because it involves rule-following and rules 'presuppose' a social setting. I consider the nature of Winch's Wittgensteinian arguments about rules; two senses in which all action might be 'social' emerge; however, were 'social actions' in either sense the focus of a science, it would not therefore aptly be called a social science, the senses of 'social' here being too weak.

I turn next to what I allege is Durkheim's basic notion of a 'social fact', roughly, that of a way of acting which 'inheres in' and is 'produced by' a social group. I present a highly articulated reconstruction of this notion: a 'collective practice', Pr, of a social group, G, will be a 'Durkheimian social phenomenon', according to this revised conception, if and only if either Pr or another collective practice of G provides members of G who conform to Pr with a 'basic' reason for so conforming. A central element in my account of collective practices is a notion of 'group common knowledge' derived from David Lewis.

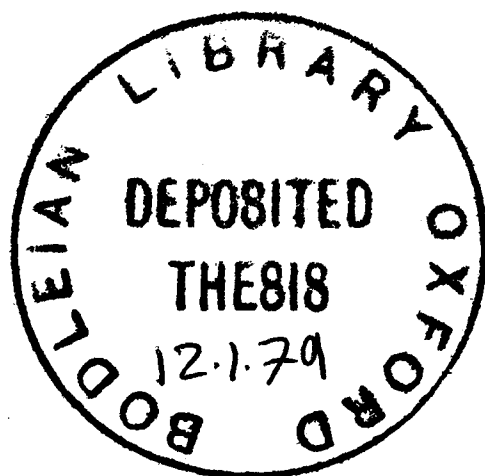
*

I finally undertake a detailed critique of David Lewis's account of conventions and of the 'co-ordination problems' Lewis claims underlie conventions; I argue for a kind of account different in form from Lewis's, in which conventions are not, and do not necessarily involve, 'regularities' in behaviour.

The Durkheimian notion is judged the best partial characterization of a social science considered. Its presupposition of the notion of a social group is, I argue, no flaw. I conclude with a general theory of 'socialness', and hence of social science, based on my judgements about the four concepts considered.

* [Because of the official word limit on the length of theses in philosophy, the D.Phil. examiners were asked to omit this section]

O N S O C I A L F A C T S



P r e f a c e

I have been developing the ideas embodied in this thesis since 1971. Papers on Weber's notion of 'social action' have been presented at a number of Universities, and at Oxford seminars, since 1972. In June 1977 a paper on Durkheim's notion of a social fact was presented to a philosophy workshop in Canada. I have also developed my ideas in the course of giving one series of lectures at Reading University and another at Princeton. I am grateful to colleagues and students for stimulating discussions and useful questions. Steven Lukes and John Mackie presided over the writing up of my ideas into their present form. Two major intellectual debts should be acknowledged, first, to David Lewis for his book Convention, and for helpful discussions in Oxford and Princeton. David Lewis very generously encouraged me to make public the criticisms of his account of conventions embodied in Chapter Five of this thesis. Second, my husband Saul Kripke has spent more time than he could afford, if less than I could have used, discussing aspects of my work. His public lectures on Wittgenstein have also been extraordinarily rewarding. All specific debts are, I trust, acknowledged in the text.

Finally, I should like to thank Mrs. Joan Oliver for her efforts in preparing the typescript.

T A B L E O F C O N T E N T S

Chapter One: Introduction.....1

Chapter Two: Max Weber and 'Social Action'.....13

Chapter Three: Action and the Social.....66

Chapter Four: A Durkheimian Conception of
a Social Phenomenon.....123

Chapter Five: About Conventions.....206 *

Chapter Six: On Social Facts.....303

Bibliography.....317

* Because of the word limit on theses in philosophy, the D.Phil. examiners were asked not to read Chapter 5, pp. 206-298.

CHAPTER ONE

INTRODUCTION

1. The topic and approach of this study

If I had to place this essay in a particular established academic field I would say that it was an essay in the philosophy of social science. However, some caveats would immediately be necessary.

First, I concentrate hardly at all on the nature and method of 'science' or on the nature of 'scientific' explanation, or on the relation of these to social subject matters.¹ My topic is, rather the concept of a 'social' phenomenon. This brings me to my second caveat. It is not common for philosophers of social science to bring up this topic, let alone to dwell on it at any length.² Sociologists, on the other hand - in particular the great sociologists Max Weber and Emile Durkheim - have given it a harder look.

1 Such an emphasis is found in most works in the field. See for instance Karl Popper's The Poverty of Historicism, and Peter Winch's The Idea of A Social Science, two well-established recent works.
^{For detailed references to all works cited, see the Bibliography.}

2 Winch says it is an important topic, but does not seem to approach it directly or critically. See op.cit. pp 41,42.

Perhaps philosophers of social science have neglected to examine the concept of the social because, knowing what so-called 'social scientists' actually deal with, they see no further reason to inquire into what the socialness of a phenomenon is.¹ Whereas there clearly remains the question: what are the conditions under which any scientist or investigator is aptly called a 'social' scientist?

The question 'what makes a science a social science?' may not grip one. One might argue that the philosophy of science in general need only interest itself in the nature and method of actual, developed or developing sciences, no matter what they are called. So, the argument might run, the very idea of the philosophy of social science is just the idea of the philosophy of science applied to sociology, economics, social anthropology, and perhaps psychology, and perhaps history too, if it is a 'science'.

I would agree that if we have a body of knowledge which we consider scientific, then well and good, and interesting, and it probably does not matter much what the science is called, or in particular, whether it is categorised as 'social' or not 'social'. Yet I think that some questions which naturally arise about the category of social science express genuine, and interesting problems. I have in mind questions like: is psychology a social science? Is history a social science or study? I think that the occurrence of such questions may often indicate confusion about things social.

People may sometimes shy away from thinking about the range and nature of social phenomena, for the following reason. It would appear that societies and societal features like institutions and conventions are social phenomena if anything is: they might even seem, at first sight, to be the only plausible focus for something aptly called a social science.

1 For brevity's sake I shall often use the phrase 'the concept of the social' as short for the 'concept of a social phenomenon in general'. Similarly I shall use the term 'socialness' as short for 'that property which entitles a phenomenon to be called a 'social phenomenon'. It should be clear that I am not interested in every single sense of the term 'social'(my interest is not, for instance, in 'social' as a synonym for 'sociable')but is rather in that sense-or those senses-it is given in the phrase 'a social phenomenon'.

However, at this point people may become ensnared by certain ontological questions: are societies 'real'? or are just their members real? They may feel that the obvious answer is that the people who are members of so-called 'social groups' provide what these groups possess of reality. So they could eventually reach this view: social sciences are just about people doing things. This claim, however, hardly illuminates the nature of social phenomena in general, or of collectivities in particular. As we shall see in the next chapter, those taking this line may feel moved to evolve special concepts of 'social action' which they hope will do this, but such concepts have a tendency not to fulfil this aim.

In my view, as regards the general nature of social groups and their properties, one will do better not to start at the individual level, with individualistic assumptions, however obvious and appealing these appear, but rather at least to begin by looking as directly as possible at collectivity concepts themselves, and to note their special features.¹ A grasp of the nature of at least some such concepts should, prima facie, play a vital role in one's appreciation of the range and relationship of social phenomena.

Someone might now argue that it is not a fear of ontological mistakes that has led to a neglect of my topic. It is rather - it might be suggested - that everyone already knows what 'socialness' is. And in a sense I assume that this is true. That is they - and I - have the concept of a 'social' phenomenon, have a grip at least on when this or some synonymous characterization is, and is not, appropriately applied, at least if the phenomenon it is applied to is specified with sufficient clarity.

¹ This will be my own approach in Chapter 4.

Without a grip on the concept of the 'social' in this sense one could not embark on my topic at all. But this of course does not mean that one needs in advance any precise analysis of this concept, or any appreciation of the range of phenomena all of which come within its scope, or of the kinds or degrees of socialness that there are. None of these things is essential to a grasp of the concept that I assume that I and other philosophers of social science, and others, possess. Nonetheless, the closer we get to all these things, the clearer we would be to understanding both our own ideas and the ways in which reality may have a 'social' character.

As the following chapters will show, social things are a motley. Our concept of the social is such that things of many different types can aptly be characterized as social phenomena. That is, while they are obviously of the same 'type' in being social phenomena, they are also obviously of importantly different types in other ways. Thus as we shall see, some, if not all, human actions may be intuitively social, and there are also, for instance, 'social' conventions, and even if it is not clear exactly what conventions are, it is pretty clear that one man's doing such-and-such at time t is never going to be identical with the occurrence or existence of a social convention.

In the final chapter I shall turn to the general concept of the 'social'. In the main body of the work I shall be more concrete, isolating a variety of types of social phenomena with reference to a more specific, but closely related, topic.

The particular question I focus on does bring in the idea of a science. It is this:

Can one give an interesting, plausible characterization of the subject matter of sciences that would aptly be called 'social' in terms of one precisely defined notion, the notion perhaps of one particular type of social thing?

Both Durkheim and Weber seem to have thought this was possible, though the notions they proposed were dramatically different.

In the next section of this chapter I shall say something more about the nature of our 'guiding' question, making some needed distinctions. I conclude this section with a survey of the way I go about discussing it.

A careful look at the proposals of Weber and Durkheim will, in effect, put two very different kinds of 'social phenomena before us: a kind of feature of social groups, in Durkheim's case, and a kind of action, in the case of Weber. I myself find Durkheim's proposal the more intriguing and plausible of the two. In senses to be explained below I shall argue, in fact, that neither Durkheim nor Weber provides an adequate 'global' theory, but that considered as a 'piecemeal' theory Durkheim's proposal is superior to Weber's. My preference is in fact, to be precise, for a radically reconstructed version of Durkheim's notion of a social fact.

~~Along with Weber and Durkheim, two other authors, and two other phenomena, will be discussed in detail.~~ I discuss Weber's notion of 'social action' in Chapter 2; following naturally on this, in Chapter 3,

I investigate the claim that all (meaningful) action is social in some way, scrutinizing the argument for this in Peter Winch's book The Idea of A Social Science. Although philosophers of social science have paid Winch a great deal of attention, this particular aspect of his work has not been adequately dealt with.

I turn to Durkheim in Chapter 4, presenting my reconstruction of his notion of a social fact. [The notion of a social convention might be proposed as a rival to this in characterizing the social sciences. In Chapter 5, before comparing the two notions, I examine the notion of convention, partly through the eyes of David K. Lewis, whose book Convention (published in 1969) has become something of a classic on the subject. But since I cannot accept Lewis's analysis, nor variants on it recently put forward, I also branch out on my own.] *

Armed with an appreciation of Weber's 'social action', Winch's 'meaningful behaviour', Durkheim's 'social facts' [and social conventions according to Lewis, and according to me,] and having discussed their individual merits as accounts of social science, I shall, at the end of this essay, hazard some brief but I hope useful and systematizing remarks on the intuitive concept of a social phenomenon, and then conclude my cautionary tale of the opaque notion of a social science.

At certain points in this work I have felt obliged to discuss in some detail what amount to central questions in areas of philosophy 'outside' philosophy of social science. Issues like 'What is knowledge?', 'What is language and knowledge of meaning?'. In my experience however, philosophy really is a seamless web. To the extent that I dwell on some such issues, it may look as if I am no longer doing philosophy of social science at all. However, in my view, whatever I have tackled should be tackled at some point if we are to have the strongest possible grasp of the range and nature of social realities. Needless to say concentration on some things has entailed skimming on others; and for some questions, no time spent on them seems to be near enough.

* See note on table of contents page.

2. On Saying What Social Sciences are About: Some Preliminary Notions.

i. Global versus Piecemeal Theory

If someone, say Weber, tells us that he is going to say what a social science is about, we may think that he is about to propose what I shall call a global theory of social scientific subject matters, and it sometimes looks as if there are a number of global theorists around. That is, it often seems to be assumed that it is possible to say in one word or phrase what a social science (any social science) is about (or is fundamentally, or basically, or ultimately, and so on, about). In other words it is assumed that one can, in a word or phrase, correctly say what a science or scientific theory will be about if it is a social science, or theory. In effect a global theorist proposes a notion, say the notion of a ϕ , which, he claims, is such that if T is a social theory, then T is about ϕ s. A global theory of what social sciences are about will, then, make a statement of the form: social theories are (all) theories about ϕ s.

Alternatively, perhaps in more Anglo-Saxon fashion, one might expect that one who tells us what social sciences are about will tell us, of certain subject matters, that they at least would be social scientific subject matters - without prejudice to the claim that other things may also be the focus of a social science. The piecemeal theorist, as we may call him here, will in effect propose a notion, say the notion of a ψ , which, he claims, is such that, if T is a theory about ψ s, then T is a social theory.

At this point consider the phrases 'What theory T is about', and 'The subject matter of theory T'. It is perhaps most natural to assume that, if T is a theory in which xs are explained by reference to the behaviour of ys, then T is about xs, rather than ys. However, things are probably not so clearcut. One might incline to the view that an explanatory theory is about its explananda because of a sense that to be about something a theory must be 'focussed on' or 'essentially concerned with' that thing, and so on. Yet these last phrases are quite vague and it is not clear that either they or 'about' point to explananda exclusively.

The most liberal line in this area is, presumably, that if a theory explains xs in terms of ys, then it is about both xs and ys. In favour of this line is that such a theory evidently takes account of both types of phenomena, reference to both types of phenomena are part of it. Moreover, such a theory is to be distinguished, along with one which explains xs in terms of xs, or ys in terms of xs, from theories with no concern at all for either xs or ys.

Similar views might, evidently, be held about a 'theory' which presented the analysis of a concept, say the concept of a social convention in terms of coordination problems in a population¹. For our purposes, there is no great need to choose between these views. We may, without argument, opt for the 'liberal' one.

Returning to our distinction between global and piecemeal theories of the social sciences, we may now put the distinction thus: a global theory will claim that social scientific theories are all theories which, given a notion of a ϕ , either explain ϕ s or explain something in terms of ϕ s. Meanwhile a piecemeal theory claims that if a theory explains ϕ s, or couches an explanation in terms of ϕ s, then it is a social theory.

ii. Some criteria for judging global theories of social science

It may seem that the one correct global theory of social scientific subject matters has to be this:

a social scientific theory is any scientific theory
about phenomena which are aptly called 'social phenomena',
according to an intuitive notion of a social phenomenon.

1. See David Lewis (1969).

As it stands, I do not think this will do, though it may be roughly right. The fault I have in mind is not triviality but this: I believe that in order to qualify as the subject matter of something worth calling a social science a subject matter must, indeed, be intuitively 'social' in some way, but also its socialness must be of a sufficiently strong kind. I shall try to justify this claim by reference to examples in the course of my discussion, when what I mean by 'strength of socialness' should also become clearer. For now we shall simply note the following as two criteria for judging global characterizations of social science: first, the putative subject matter must be intuitively social; second, the socialness must be of (intuitively) a sufficiently strong kind.

We might now revise the proposed global 'theory' thus:

a social scientific theory is any scientific theory
about phenomena which are, intuitively, (a) social
phenomena, (b) phenomena which are social to a sufficiently
high degree.

Unfortunately this characterization is still relatively trivial and, consequently, distinctly unilluminating. We should obviously prefer an informative characterization.

At this point, more restrictive characterizations of social science are worth considering. Any attempt to make a more specific global characterization clearly risks being false. Such a characterization could, however, turn out to provide at least an illuminating piecemeal theory, or (less exciting but still useful) an aid to saying more about the concept of social phenomena in general, and hence to making more of the trivial account just considered. I originally approached the concept of a social phenomenon and saw that there were 'degrees' of socialness, by assessing quite restrictive accounts as either global or at least piecemeal theories of social science. I shall here let the order of exposition follow the 'order of discovery'. We shall return to the trivial theory just noted, after assessing some less trivial ones.

iii. Some non-trivial forms of global theory

The claim that ^arestricted but correct global account of social scientific subject matters can be given should not be dismissed out of hand, even if it would be surprising prima facie should the trivial theory just adduced turn out to be wrong. In this section I indicate some relatively plausible forms such a global theory might take. All these involve their proponent in isolating one particular type of social thing, and going on from there.

One such claim might be the following. Though there are distinguishable types of social phenomena (like actions, on the one hand, and social conventions, on the other) one type of social phenomenon is basic to all others, so that if something is a theory of As (where As are some other social kind of thing) it is ultimately or basically or fundamentally a theory of these basic things, since As just are, say, aggregations or 'plural' versions of the basic 'atoms'. Social sciences would then be either sciences dealing with atoms-collected-into As in various ways, or about individual atoms. Weber's claim that 'social action' is the subject matter of social science could be construed to be of this type.¹ I discuss Weber's claim in detail in the next chapter.

Another attempt at a non-trivial global theory might argue that one particular type of social phenomenon is pre-eminently social, social perhaps to the highest degree. In order to make this plausible as a global theory, one could try one of two lines.

1. Of course Weber himself did not use the phrase 'social action' but rather the German 'soziales Handeln'. However this has come into the literature in English as 'social action', and I shall in what follows write freely of 'Weber's definition of "social action"' and so on.

(1) One could argue that these highly social things are the primary social things, in such a way that other social things all have a derivative or borrowed status. As they would, for instance, if their socialness consisted entirely in their standing in some relationship to these primary social things.¹

It is not clear that anyone actually holds a global theory of this form. We come fairly close to one, however, in Peter Winch's claims that (a) all (meaningful) action is 'social' and (b) this socialness is a function of each action's 'presupposing' socially established rules. An acceptance of such claims might lead one to think that 'socially established rules' were the primary social phenomena insofar as they were 'presupposed by' every other kind. However, I am not sure that Winch himself does think this, nor indeed whether this notion of the primacy of socially established rules is ultimately coherent. Clearly the first thing to do in this case is to establish (or reject) Winch's expressed claims.

(2) Another form of global theory, starting from the idea of the pre-eminent socialness of some particular type of phenomenon, might run thus: only things with the type of socialness in question - socialness to the highest possible degree - can be the focus of something properly called a social science. I have already said that I believe a sufficient degree of socialness is necessary, but this line is surely going to lead to implausible results if pressed too far. Nonetheless it is certainly worth considering what the most highly social phenomena might be - we consider some candidates in Chapters 4 [and 5.]

¹ Cf. How someone might say they were French because their mother was French; although they themselves were born in China and speak only Chinese.

The most promising non-trivial form of global theory, then, appears to be the 'atomistic' one. We shall keep them all in mind, however, to see if anything further can be said for or against them. As we noted, any concept a global theorist proposes may be considered for its use in a piecemeal theory. Let us briefly consider what we should be looking for as far as these go.

iv. Criteria for judging piecemeal theories.

What might one expect of a good piecemeal theory? In addition to the fulfilment of the criteria of socialness and strength of socialness we would hope it is not too piecemeal. The more it is subsumed under it the better: thus breadth is a desideratum. To cite a simple kind of example: if you think Jones's saluting his commanding officer is clearly something social, then you could say that any theory which explains or perspicuously describes this very phenomenon would be a social theory, but it would be more helpful to work out what general kind of phenomenon this is, such that all phenomena of that kind are equally social.

Another desideratum, which should rein in one's desire for breadth, is that whatever one stresses as a social kind of thing be a relatively natural kind of phenomenon, as far at least as one can tell a priori. For otherwise we have here one obvious problem with the original 'trivial' global theory, writ small. For whatever the analysis of the concept of a social phenomenon may turn out to be, social things are surely a motley, still, and so saying that social science will deal with them will never be completely illuminating: we need some appreciation of their kinds. It is this kind of thought which could reasonably lead to a search for good piecemeal theory in the first place. So a piecemeal theory had better not be strikingly open to this criticism itself. We shall see this criterion's cutting edge in the next chapter particularly.

So much, then, for our highest generalities. We now turn to a breed of theories, theories of 'social action', whose most eminent proponent is Weber.

C H A P T E R T W O

M A X W E B E R A N D ' S O C I A L A C T I O N '

1. On 'Social Action' in general.

i. Motives for speaking of 'Social Action'.

There is much use in social scientific and related philosophical literature of the phrase 'social action'. As I shall shortly argue we can assume no standard meaning for the phrase, though use of it is doubtless encouraged by Weber's strong emphasis on his technical concept of social action. I think it is important to note that there are some quite pressing motives for using this phrase even without a precise understanding of it; these arise out of a number of distinguishable kinds of 'individualism' with respect to social scientific subject matters.

First, someone might agree with Durkheim that the subject matter of sociology is (whole) societies or social groups, their properties, and the inter-relations of these properties. Let us call such a person a 'topic holist'. That is, he holds that social science in general has as its topic whole societies, etc. Such a person may still be in sympathy with a certain individualistic analytic line, a certain view of the analysis

of collectivity concepts and the meaning of claims about collectivities. For instance, he may argue, as Weber does, that to say that a society exists is just to say or imply something about the actions of individuals. Perhaps he will say that there is some analytically true statement of the form 'a society exists if and only if individual people act thus and so'. Perhaps he does not think there is a tidy analysis in terms of a set of individually necessary and jointly sufficient conditions, but does assume that, if there is a society S, then there being such a society is, necessarily, only a matter of certain individuals acting in some way sufficient for a society to exist.

Let us call those who believe, roughly, that claims about collectivity concepts can be analysed in terms of claims about the actions and states of individuals 'analytic individualists'. We have already noted (in Chapter 1) a general feeling one might have, which may not be expressed in such logical terms, that the only thing real about a society is the individual members of the society. We might call this view 'ontological individualism'. Even topic holists, then, may be individualists of one kind or another. Let us leave the topic holist for a moment.

More radically individualist lines about social scientific topics are possible. Thus one may be a 'topic individualist' of some kind.¹ One might, for instance believe that a social science is simply a science that deals with human actions. As Ernest Nagel observes:

'The subject matter of the social sciences is frequently identified as purposive human action, directed to attaining various ends or 'values' whether with conscious intent, by force of acquired habit, or because of unwitting involvement'.²

¹ Ontological individualism with regard to societies could itself lead one to such a view, as we indicated in the introduction.

² E.Nagel - The Structure of Science, p. 473

Examples of social scientific concern that might be cited here are: capitalistic activity, religious behaviour. Let us call this line the unrestricted action thesis. This line is obviously, it seems to me, open to the question: why call all studies of action (or all and only studies of action) social studies? After all, what is so social about an action as such?

One way of answering this might be, indeed, to show somehow that by their nature human actions are social, that is all human actions are, by their nature, social actions in some appropriate sense. That this is so has in fact been argued by Peter Winch.¹

Another option that a topic individualist might take is this. Say that social sciences are not, indeed, those sciences which deal with any possible type of action, but are rather those which are concerned with a special subclass of human actions, which can plausibly be called 'social actions'. Thus the 'socialness' of the subject matter of social sciences could one again be preserved. We might call this line the restricted action thesis.²

Topic individualism then is faced with a clear motive for claiming that all, or some important class, of actions are social phenomena. Those who hold the unrestricted action thesis need to say all actions are social. Those unwilling to say this might shift to a restricted action thesis: at least some actions are social, and it is these the social scientist is concerned with.

¹ See Chapter 3 below.

² Cf. Nagel (1961), pp. 473-4. 'A somewhat more restrictive characterization limits [the subject matter of the social sciences] to the responses men make to the actions of other men, in the light of expectations and "evaluations" concerning how these others will respond in turn'. Other restrictive characterizations are possible, as we have already seen, and as will be further borne out.

Let us now go back for a moment to analytic and ontological individualism. It might seem that the problem of 'disappearing socialness' that we have noted does not afflict a topic holist, even if he also is an individualist. However, an individualist might well begin to wonder what it was that made individuals into a social group or collectivity of some sort. Again, what is so social about individuals performing actions? Now one does not after all want to appeal to something other than individuals and their properties. That was one's motive, presumably, for being an individualist in the first place. So the analytic individualist, also, may be inclined to look for a special kind of action, aptly called 'social' action, which is what makes individual agents into a collectivity. The socialness of social groups and their properties could then be seen to be a matter of the socialness of certain actions, albeit actions individuals perform.

In conclusion, we can see a number of possible motives for developing a notion of social action in social theory or metatheory. Moreover, as a corollary of this, we can see how even if a satisfactory notion or set of notions has not actually been developed, or is even known not to have been developed, there are motives for using the phrase 'social action' all the same. It at least gives the impression that one's particular problem of disappearing socialness has been solved.

ii. The proliferation of 'social action' talk

Since Weber, many writers in the social sciences and in philosophy have indeed given the impression that social science is a science of something called 'social action', Talcott Parsons' massive general work on sociology (published in 1937) is entitled The Structure of Social Action. One could go on citing references almost indefinitely.¹

¹See for instance W.J.Sprott, Science and Social Action (1954), J.D.Douglas The Social Meanings of Suicide-chapter 13-'Social actions as meaningful actions' (1967), Quentin Skinner 'Social Meaning' and the Explanation of Social Action and now Martin Hollis, Models of Man: Philosophical thoughts on social action (1977).

It is generally not explicitly stated whether or not a global or a piecemeal theory is at issue, and the implicit claim about social action can be considered from both points of view. So, these questions arise: What is 'social action'? Can we usefully characterize the subject matter of social sciences or studies in terms of it, and, if so, how can this be done?.

The first question - what is social action? - is not easy to answer. First, it is fairly clear that there is no universally accepted sense of the phrase among these writers.¹ An explicit definition is rarely given, and sometimes, one suspects, the phrase may be used without any special understanding, that is, we have mere mouthing of a convenient phrase. As far as I can ascertain the 'everyday' uses of the phrase 'social action' (there seems to be at least one, connoting something like 'action aimed at improving the structure of society'), these do not determine the sense of the phrase as used by most writers in the fields concerned. Authors may, of course, be following Weber in his use of the phrase, but this does not mean that they use it in his sense. This sense, for one, is often misreported.² In other cases the context suggests a different sense.³ To compound confusion, 'action' and 'social action' are often used as if they are interchangeable.⁴

¹ A similar phrase 'social behaviour' also lacks any definitive meaning. I am certain that not everyone who uses this phrase agrees, for instance, with that of G.C.Homans in Social Behaviour: its elementary forms, p.2. 'the behaviour must be social, which means that when a person acts in a certain way, he is at least rewarded or punished by the behaviour of another person'.

² See A.Inkeles, (1964), p.7, for example.

³ Cf. for instance, J.D.Douglas, op.cit.

⁴ See, for example, Parsons, op.cit., and in the article 'The role of ideas in social action' (1938). Weber himself does this also, which may account for some of the misreporting of his notion just noted. He, at least, does give an explicit definition of 'social action' so we know what he is supposed to mean by it.

Apart from Weber's notion (which is defined roughly as action 'oriented' towards another person)¹ and the everyday notion noted above, a large number of definitions of 'social action' are possible, given the meaning of the qualifier 'social'. While some definitions are indeed unnatural, many natural interpretations of the phrase 'social action' remain available.

A definition which would be ruled out is, say, 'action performed by a person with brown hair'. On the other hand the following all appear to be natural definitions of the phrase, whatever their own obscurities: 'action aimed at compliance with someone else's wishes', 'action of a society', 'action done in order to conform to the conventions of one's society', and so on. This compounds the difficulty of interpretation. Of course we could just 'interpret' the phrase thus: 'action which is social (in some way or other)'. However, given the plethora of possible ways, this leaves the term supremely vague, not a virtue of a technical term.

Different more specific notions of 'social action' will serve different purposes. Our initial question here may now be amalgamated with the rest:

is there a natural sense of 'social action' such that we can give a global or piecemeal theory of social science in terms of social action of the relevant kind?

Regarding global theories, the answer may seem obviously: 'no'. What, after all, of societies, social conventions, and so on, and so forth?. However we have seen how a number of lines of thought--all 'individualisms' of some kind, provide a motive at least for answering 'yes'. We now proceed to assess the importance of Max Weber's carefully defined notion of 'social action' for our purposes.

¹ I give a detailed exposition shortly.

2. Weber as global theorist and analytic individualist.

Weber begins his massive work Economy and Society with a section entitled 'The Fundamental Concepts of Sociology'.¹ Pride of place goes to the notion of 'social action' whose precise definition we shall shortly consider. It looks as if, in our terms, Weber is putting forward a global theory of sociological subject matters in terms of social action.

Thus consider the following quotations:

- (1) 'Sociology (in the sense in which this highly ambiguous word is used here) is a science which attempts the interpretive understanding of social action...' (p.88).
- (2) 'Sociology... is by no means confined to the study of 'social action'; this is only, at least for the kind of sociology being developed here, its central subject matter, that which may be said to be decisive for its status as a science' (p.114-5).

The second quotation may seem to express less of a 'global' theory than the first, yet the claim it makes is quite strong. Social action, in Weber's sense, is claimed to be the central subject matter of sociology. From now on we shall take this as a theory of social scientific subject matters in general, that is, we shall assess it as such. In fact, everything we say will be equally relevant to the perhaps more limited claim about sociology that Weber explicitly makes.

¹ Unless otherwise noted, all quotations from this work are from T.Parsons' edition and translation (with A.M.Henderson) of Part I, published under the title The Theory of Social and Economic Organization (1947), and all page references to Weber's work are to this volume.

Weber's global theory may result from his analytic individualism. The sociologist 'cannot afford to ignore' 'collective concepts' such as the concept of a state, association, or business corporation.¹ On the other hand 'it is the task of sociology to reduce these concepts to 'understandable' action, that is, without exception, to the actions of participating individual men'.² This proposed sociological task looks like a purely analytic one: a task of showing how certain concepts 'reduce' to certain others, or may be analysed in terms of them. Weber's individualism here, then, lies in his claim about analysis: the analysis of collectivity concepts will be in terms of the actions of individual men.

In the following brief statement Weber informally sketches out an analysis of the desired kind; it is also in terms of 'social action'.

'When reference is made in a sociological context to a 'state', a 'nation', a 'corporation', a 'family', or an 'army corps' or to similar collectivities, what is meant is... only a certain kind of sequence of actual or possible social actions of individual persons'. (p.102)³.

¹ p.101

² Gerth and Mills, eds. p.55

³ In this quotation I have replaced Parsons' 'development' with 'sequence'. I thank F.H.Stewart for suggesting this.

In this chapter I shall ask whether Weber's concept of 'social action' has a serious contribution to make to the analysis of collectivity concepts. I shall also ask whether it can play a useful role in general in demarcating the province or part of the province of the social sciences.

I should stress that I shall not be attempting exegesis of Weber's text except in order to clarify the content of his concept of social action. My critique of this notion will be very much in my own terms.¹

In the next two sections I give a careful presentation of Weber's relatively well-articulated notion of a social action, before attempting to assess it.

3. What Weber's 'social action' is.

What exactly is Weber's notion of a 'social action'? At the beginning of Economy and Society Weber gives an explicit definition for 'social action' and discusses cases in an attempt to clarify its range of application. In spite of the care Weber takes, there remain some obscurities in the notion, as in Weber's more basic notion of 'action'. I hope that the following account will be reasonably clear intuitively, reasonably faithful to Weber's gist, and adequate to my purposes here.

¹ I shall formulate my criticisms as boldly as possible; I mean to imply no doubts about Weber's genius.

We should start, as Weber does, with his definition of 'action' simpliciter.

i. 'Action'

Weber says:

'In "action" is included all human behaviour when and insofar as the acting individual attaches a subjective meaning to it' (p.88).

He adds, in clarification, that, in effect, action in his sense need not involve any obvious physical movement; it may consist, for instance, in one's deliberately refraining from 'positive intervention' in a situation.

The full subjective meaning of a piece of behaviour may, as Weber understands it, be quite complex. First, there is the 'subjective meaning of a given act as such' (see p.94). It appears that this would be captured by the answer to the question, asked of a particular 'beholder', 'What does he mean to be doing?'. To cite two of Weber's examples, a man wielding a chopper in a certain way may be chopping wood - the stressed words may capture the subjective meaning of his given act 'as such' (see p.95)¹; a man moving a gun in a certain way may be aiming the gun at an animal, and so on.

I think that we can reasonably claim that the core of Weber's notion of behaviour with a 'subjective meaning' lies here. That is, I suggest that we can give a rough account of it thus: when a person A behaves in a certain way he attaches a subjective meaning to his behaviour if and only if something of the form 'A means to be doing so-and-so' is true of A.

¹ They may not, even if he is in fact moving in such a way that wood is being chopped by his chopper. He may be in some state of automatism such that he 'attaches no meaning' to his behaviour.

This does not in fact capture the concept of intentional action, as one assumes Weber intended it to. In the next chapter, when we discuss Winch's idea of 'meaningful behaviour', we shall say more about this.¹ For now we may proceed without further discussion.

In order to arrive at the full subjective meaning of a given act, further questions are relevant.² In particular 'What is his motive in e.g. chopping wood?', and perhaps also 'Why does he have this motive?' Weber gives examples of motives: the wood chopper may be working for a wage, or chopping a supply of firewood - chopping in order to gain his wages, or to obtain a supply of firewood. The gun aimer may have been 'commanded to shoot as a member of a firing squad', aiming in order to shoot, in obedience to commands made to him as a member of a firing squad. We might also describe this last case, more naturally perhaps, as the man's aiming in order to shoot, and purposing to shoot, or shooting, because he has been ordered to do so by the leader of the squad of which he is a member.

One's motives in doing something, e.g. chopping wood, may of course be manifold. In particular, it is worth mentioning a motive which may frequently be a 'subsidiary' motive. This is a desire to do what one does for other reasons in a socially acceptable way.

¹ See Chapter 3, introductory section.

² Only, I assume, on the assumption that the first question is relevant.

For example, one may wish to rebuke a colleague, but only in a socially acceptable way; thus one's saying 'I'm awfully sorry old chap, but...' to X could be motivated by the desire to rebuke X and to do so in a socially acceptable manner.¹ It is clear that such an aim, even though subsidiary to the main one, should be included in an action's subjective meaning.

We will not go far wrong, then, if we take it that the full subjective meaning of a piece of behaviour will be given by the answer to the question: 'What does the person concerned mean to be doing, what does he hope to achieve by doing it, and what is his reason for seeking such an achievement?' One modification to this conception should, however, be made.

Weber mentions the case of one who is 'working off a fit of rage' by his woodcutting. He calls this 'an irrational case' (pp.8-9), and distinguishes between these and other cases. Let us consider the locution 'He's working off a fit of rage' as an answer to the question 'Why is he chopping wood?'. This could be interpreted in terms of what have been called 'in order to' motives. That is, it could be interpreted as telling us what the agent hopes to achieve by chopping wood, as in "he's doing that in order to get rid of his rage". Here there is the connotation of some deliberate attempt on the agent's part to rid himself of his anger, but such a locution will also cover a case where the above interpretation is not appropriate.

¹ Cf. Weber, op.cit.pp.124-130, on action oriented to a belief in the existence of a 'legitimate order'.

It is rather that he is angry, and in his anger, he is chopping up the wood. He is expressing his anger, rather than deliberately doing something to get rid of it. Insofar as we place the act in its 'context of meaning' by stating that it is an expression of anger, this case has not been covered by our account of subjective meaning so far.

We may take it then, that the subjective meaning of a piece of behaviour is given, roughly, by the answers to the questions: 'What does he hope to achieve by doing it (e.g., to get rid of his anger), and what is his reason for seeking such an achievement (e.g., to be nice to his wife), (and so on)' where these are appropriate; and in answer to the questions: 'What does he mean to be doing (e.g., chopping wood), what emotion is being expressed (e.g., anger) and what or whom is this emotion directed at (e.g., his wife), (and so on)' where these are appropriate.

This takes account of the main aspects of 'subjective meaning' that Weber mentions,¹ and hence of the main aspects of his notion of 'action'.

Regarding the thoughts of the agent, we may suppose that in cases of 'in order to' motives there will be a belief linking intended act and aim, the belief that chopping wood is the, or at least, a, means of gaining a wage. Or, more generally, a belief of the form ' ϕ -ing' is a or the means of achieving goal G, (and similarly for 'motives for motives' and so on.)

¹ Weber's notion of 'Verstehen' (usually translated 'understanding') is closely tied to his notion of 'meaning'. Understanding an action is, basically, grasping the action's meaning. (see pp.94-96). I shall not discuss this notion further here.

Citation of such a belief of the agent will, moreover, contribute to an explanation of the action. A question which may be raised, and which is not as far as I know clearly answered by Weber, is whether such 'means-end' beliefs are part of the meaning of the action in question.

I think this issue need not trouble us here. We are concerned with social actions, and whether an act is 'social' or not, in Weber's sense, is indeed, as must be stressed, a matter of its meaning's content, but the 'means-end' beliefs just referred to may be assumed to have a virtually unchanging content, except insofar as the other aspects of meaning (e.g. aim, act intended, 'subsidiary' motives) which we have noted, vary. What is more, this virtually unvarying content will not by itself make an action social. So whether we allow these beliefs to be part of an action's meaning or not, will not alter the 'socialness' or otherwise of any action. At this point let us turn to Weber's definition of 'social action'.

ii. 'Social action'

According to Weber, action is social when:

"by virtue of the subjective meaning attached to it by the acting individual(s) it takes account of the behaviour of others and is thereby oriented in its course"(p.88)

The reference to the behaviour (Verhalten) of others does not seem to be crucial: one's production of something would, according to Weber, be a social action if it were 'oriented to the future wants of others'(my stress)¹. What seems to be most important is that the 'others' involved in the subjective meaning of social actions be animate.

¹ (Note that Parsons translates "Verhalten" here as "attitudes and actions" in his Structure of Social Action.)

Paradigm 'others' will be human persons, but 'others' is to be interpreted at least widely enough to apply in principle to non-human animals.¹

What role must the thought of others play in order that the action be social? Note that it is not enough that the agent merely have someone else 'in mind' while he acts; this should be fairly clear from the account of 'subjective meaning'. Thus someone cutting wood might be working for a wage, and happen to be thinking meanwhile about a friend of his. This thought would not contribute to the subjective meaning of the act, nor therefore would it make it a social action. On the other hand, were he doing overtime in order to help his mother financially, this would bring his mother into the subjective meaning of his action, and so make his action social.

Weber says that the reference ~~of~~ another in the subjective meaning of a social action must 'orient' the action 'in its course'. I take it that he wants to draw a distinction illustrated by the following case. Suppose we ask someone what she is doing, and she says 'I'm going to look at the clock my father gave me'. Is the action she describes social? In order to ascertain this, we need a more precise account of the subjective meaning of her action. Perhaps her intentions are most aptly described as follows: she intends to look at a certain clock (which, she is aware, was given to her by her father) in order to tell what time it is. Thus we might well say that what she really means to do is to look at the clock in question, in order to tell the time - that is, we might well drop out of our description the reference to her father and his action, but perhaps she wants to look at the clock in question only insofar as it is the clock given to her by her father (perhaps because she is intent on reflecting on his generosity).

¹ (Cf. p.104. - Here Weber refers to the 'theoretical possibility' of 'a sociology of the relations of men to animals'.)

Then it seems that the reference to her father has an essential role to play in capturing the subjective meaning of her act. We might say, then, that an action will be 'Weberian-social' if and only if reference to another has an essential role to play in capturing the subjective meaning of the act. More briefly, given the notion of subjective meaning, all we really need to say is that an action is Weberian-social if and only if, its subjective meaning includes reference to another.

4. What Weber's social action is not.

It is worth pointing out, explicitly, as Weber does, some particular kinds of action which are not Weberian-social, though intuitively one might be inclined to think of them as 'social' actions, either in thinking their motives through, or 'in the field', in the course of observing a certain tribe, say, and applying some intuitive concept.

Action which is merely influenced by the presence of other people is not ipso facto social in Weber's sense. Thus an action does not become social just because the agent is somehow influenced by being a member of a crowd. Weber in effect argues that studies of 'crowd psychology' like those of Le Bon, are not sociological.¹

¹ See pp.113-4. Cf. Durkheim's, whose initial instinct, in the first edition of the Rules of Sociological Method, was to treat crowd phenomena as social phenomena. See our Chapter 4: last section.

A second point Weber stresses is this:

the imitation of others' actions is not necessarily a social action, though it may be.¹ It depends on the precise nature of the imitative action. Thus compare seeing the Queen in a certain hat, and buying a similar looking hat because 'that was a nice hat I saw the other day', and buying a hat of a certain kind 'because the Queen has one like it'. In the former, not the latter, case, 'one's action is causally determined by the action of others, but not meaningfully'.

Action similar to that of many other persons is not thereby social action (p.113). An example (from Weber): We all put up our umbrellas at the same time - rain has started to fall and we are responding to it. An example involving a longer time and space span would be the following: All over the country, every morning, people wake up, stretch, and yawn. Given that they mean to stretch, they may well have no further 'motive' than that they feel like stretching. They need have no thought of what others are doing, have done, or will do, in any case. Thus though an observer would see everyone in the country conforming with regularity to this pattern of behaviour, there might well be nothing Weberian-social in it. Thus not all regularities conformed to throughout a society involve the participants in social action. Weber, is, in fact, particularly interested in the different types of 'uniformities' in behaviour found within societies.²

¹ The reference to imitation occurs in reply to sociologist Gabriel Tarde's claim that imitation was the fundamental social phenomenon (p.114). Durkheim stresses the difference between himself and Tarde in the Rules, p.10-11 (English ed.)

² Cf. pp.120-124. We discuss in detail the nature (and intuitive socialness) of some such regularities in Chapter 4.

Evidently, there are cases to contrast with the foregoing: cases where a regularity in behaviour is conformed to throughout a society and does involve social action. We might, for instance, all drive on the right because of 'a mutual declaration to the effect that a certain kind of action will be undertaken or is to be expected'.¹ Again, if everyone wears a flower in his or her hat because their chief has commanded this, there will be a widespread regularity in behaviour involving social action. The fact that we obey the (same) chief results in homogeneity. It is true as usual that the homogeneity in no way makes our action social. On the other hand, it is not the mere socialness (in Weber's sense) of the action here, or in the previous example, that results in the homogeneity. Here it is rather the (extra) fact that we have the same chief (and are following the same order of his). In general, just any plethora of social actions does not produce any (further) homogeneity in behaviour. I think Weber would agree with this, but we might briefly consider a passage which could give a different impression.

In the essay from which the last quotation comes, (written before the part of Economy and Society on which our discussion and interpretation rely) Weber remarks:

'it is by no means only social action (here: Gemeinschaftshandeln) which makes it appear "as if" action was determined by a consensual order; rather such an effect can be produced equally and even more drastically by the various forms of "homogeneous" and "collective" behaviour (Massenhandeln) which are not part of social action',²

He uses his example of the umbrella-raisers here too. This is an example of 'homogeneous mass behaviour'.

¹ See Weber, 'Some Categories of Interpretive Sociology', Logos IV. reprinted in the Appendix to Economy and Society eds. G. Roth and C. Wittich. (1977).

² Ibid.

How are we to understand this passage? Weber's remark could, I think, be taken as advice to a sociological observer. Mere homogeneous mass behaviour may involve 'patterns' on the ground that are not different in outward form from patterns resulting from, say, reference to past agreement, so the observer should not assume all regularities involve social action, let alone such specific references. Whether behaviour is social depends on its subjective meaning, and here only in the latter case is the subjective meaning of the right type. Weber at least does not make it clear that there can be a great deal of social action going on without there being the appearance of 'determination by a consensual order', and one hardly feels clear, in reading these early passages, that it really is the notion of social action itself that is felt to be central, as opposed to behaviour crucially like or actually involving 'determination by a consensual order', whatever precisely that is.¹ However, the 'official doctrine' is certainly that the broader category of social action is the most fundamental.

Finally, we should be quite clear that according to Weber's notion:

(a) a social action can be performed when an agent is alone. The idea of social action here is not the idea of action 'directed at' someone in the agent's present environment. Moreover, (b) Weber would agree that the other or others to which social action is oriented need not be personally known to the agent in question.

¹ Adherence to social conventions might naturally be said to involve 'determination to a consensual order'.

What he says is:

'The 'others' may be individual persons, and may be known to the actor as such, or may constitute an indefinite plurality and be entirely unknown as individuals. This money is a means of exchange which the actor accepts in payment because he orients his action to the expectation that a large but unknown number of individuals he is unacquainted with will be ready to accept it in exchange on some future occasion' (my emphasis) (p.112)

Certain queries arise in relation to Weber's intentions here.

Must one's orientation be to others who one believes to exist (i.e. others who one believes are real people or animate beings?). A further query is - must one be right in this belief, that is, must one's orientation be to others who are indeed real people or animate creatures? Let us at this point consider some cases.

5. Possible cases of social action

(1) Must one believe that the others one has in mind when one acts actually exist, in order to perform a social action? An example of someone not fulfilling this condition would be a solitary islander who believed he was the sole survivor of an atomic war. He might nonetheless light a fire in order that if, by some miracle, someone else both survives and passes by, he will approach the island. He believes then, that there may be others in existence, and they may pass by: but that probably there are not.

(2) Two cases not ruled out even given a condition that one believes in the others' existence are the following:

(a) Crusoe, before the advent of Man Friday, puts up a big fire expecting that a ship will pass, having seen ships passing many times before. He intends that the people on the next ship see the blaze and investigate its source. He repeats this procedure often.

Crusoe has, of course, been brought up in a human society and not always been solitary. Nevertheless he is solitary now and can, it seems, perform Weberian social actions in that state.

(b) Further, it appears that, unless (as Winch suggests) it is logically impossible for someone to act meaningfully if he never had any contact with a society, someone could in principle perform social actions according to Weber's definition even if he had never been in any contact of any kind with other real people (or other animate creatures). As long as he had the relevant concepts, he might act with others in mind in the relevant way - or so it seems. Thus, as far as belief is concerned, he might believe with some reason (perhaps good reason) that there are other creatures like himself on the earth, and he might light a fire on the seashore expecting that a vessel (like those he builds himself) with people on it will one day pass, and they see the blaze and investigate its source.

Since it is simply not clear what restriction if any Weber wishes to impose regarding an agent's beliefs about the relevant others' existence, and regarding the others' actual existence, I shall impose only one restriction of this sort. I assume that the agent must believe that the other in question may well exist.¹ At a later point in the discussion I introduce some further restrictions in order to see whether Weber's concept can do a certain kind of work better for us with such restrictions introduced. In advance of a discussion of the work the concept is to do, we can leave things open.

So much, then, for Weber's rather careful and well-illustrated - though still somewhat problematic - account of what social action is.

Our concern will now be with the role, if any, which this particular notion can play in the elucidation of the nature of the social sciences. We start with its relevance to the analysis of collectivity concepts.

¹ Thus one rules out at least his acting with a known-to-be fictional character in mind, deciding quickly, say, so as not to become like (Shakespeare's) Hamlet.

6. Weberian social action and collectivity concepts.

i. A note on 'collectivity concepts'

In what follows I shall focus on what I shall call the 'everyday' notions of a society and of a social group, that is, my own intuitive notions. The ideas of a society and of a social group appear to be paradigm cases of collectivity concepts. They are, however, notoriously hard to analyse, if not to use.¹

It is perhaps not obvious that if the social scientist develops a set of collectivity concepts these must correspond in any way to any concepts of everyday language or intuition. My aim here, however, concerns notions of the latter kind rather than any technical ones, that have been or might be developed. There is to my knowledge no established generic technical notion of social group or society in the social sciences, and I do not propose to try to tell scientists which notions they should employ.²

I believe, however, that often, when sociologists and others give definitions of social phenomena, they do intend to capture, at least to some extent, intuitive notions which we hold in advance of the construction of social theory. Insofar as a technical notion purports to capture an intuitive notion of society or social group, and indeed insofar as it might be taken to capture one of these notions, my discussion will have a bearing on its claim to do so.

¹ For some attempts on this notion - both interesting, neither quite satisfactory - see D. Emmett, Rules, Roles and Relations; A. Honoré, 'Groups, Laws, and Obedience'. See also George Simmel's 'How is Society Possible?'

² I shall in fact hazard at the end of this chapter that Weber's notion of social action is probably not one which should be employed. Even in this case I shall, however, leave it that the onus of proof is on those who claim this is a useful concept: may they produce an interesting theory which makes use of it!

I hope, then, that what follows will help to indicate how we might judge whether a given technical notion is indeed a collectivity concept: whether it captures the 'social' essentials of non-technical notions of society and social group. I believe that in general it is important to place sociological theory in the context of common notions, to know, in other words, what (of those things we are familiar with) it is about, if any.

Insofar as we would intuitively expect a social science to deal with what we think of as societies and social groups and their properties, it is essential to grapple with the everyday notions of social group and society if we wish to understand what social science is in fact about. That is, even if these notions turn out to have no role in any science or if a so-called 'social science' appears to touch not at all on this subject matter. In this section I skirmish with these recalcitrant everyday notions, with a particular question in mind: can Weber's notion of social action help us to elucidate them?

ii. A solitary social actor

We have seen that a solitary Crusoe-like character can perform social actions. Assume that a certain solitary islander - call him Eosurc - once lived in a community and knew the ways of the world; he is now alone on an island and sees no one, but he knows where and what the island is, and he knows ships pass fairly close by three or four times a week. So he lights a huge fire and keeps it burning constantly, hoping that those unknown others who, he knows, will pass by (and who do indeed pass by) will see the fire and come and find him there. If a 'social actor' is one who frequently performs Weberian social actions (and let us define 'social actor' thus for the present), then Eosurc is a social actor, but for all the social actions going on on his island he does not thereby form a society or so it seems.

Perhaps a one-man society is possible. Even if so, it seems clear that 'There is a society' or 'There is a social group' is not best analysed as 'Numerous social actions are performed in a certain area over a period of time'. For Eosurc's case as so far described would fit this, and he, with his single-minded firelighting and otherwise unidentified pursuits, cannot without further ado be nominated a member of a one-member society. Is it possible, nonetheless, somehow to analyse 'there is a society'¹ or 'a society exists' in terms of social actions?. Let us return to Weber's sketch of an analysis.

iii The Weberian social action thesis.

Weber's sketch of an analysis-or sketch of the form of an analysis-went thus:

'when reference is made in a sociological context to a 'state', a 'nation, or to similar collectivities, what is meant is only a certain kind or sequence of actual or possible social actions of individual persons'.

This statement is, clearly, hugely vague. It does not obviously rule out the Eosurc case just considered. It is evident that some weight is intended to lie with the indication that 'a certain kind of sequence of actual or possible' social actions is what is important. Yet Weber obviously does think that something is added by mentioning social action here. I hope to make it clear how minimal a contribution the notion of Weberian social action can make to the analysis of collectivity concepts, even when we try to use it within a certain fairly promising, restricted, many-person framework.

¹ In the following sections I shall tend to use 'society' rather than 'social group', but the points I shall be making apply equally, I believe, to both concepts. 'Social group' is the generic, and more colourless term, applying both to societies and to less all-embracing groups. I use 'society' here for vividness.

I take it that, in considering the analysis of the concept of a society, or of a social group, the underlying question is 'what makes a society out of a plurality of agents?' Clearly the notion of a society applies in its least problematic applications to a plurality of beings, the members, so-called, of the society. Moreover, probably the clearest cases of societies will be societies of agents of a Weberian kind, that is, those whose waking life is in part a series of 'meaningful' actions, whose behaviour will be explained partly at least in terms of their motives, reasons for acting, et hoc genus omne. Given this question, the answer we shall consider is, crudely and briefly, 'Weberian social actions'.

This answer (which we may call the 'Weberian social action theses' or 'the WSAT' for short) will be initially construed as referring us simply to social actions as defined by Weber, not to some species of his genus 'social action'. At a certain point, however, we shall make aspects of the notion more determinate, when this seems a useful line of approach. I should recall that Weber's own very brief sketch of an analysis of collectivity concepts is more complex than the WSAT as characterised above; as my discussion proceeds I shall consider more complex versions of the WSAT, but I shall not be concerned to move in a Weberian direction.¹

¹ While in his brief sketch Weber does give the impression that the notion of social action makes an important contribution to the analysis of collectivity concepts, he in fact builds up more formal accounts of such concepts (like that of a 'corporate group' on p.145) on the basis rather of the somewhat more complex, though not wholly precise, notion of a social relationship (p.118). As it stands, however, this notion can be criticised in roughly the same way as I criticise Weber's notion of social action, insofar as it purports to be a crucial element in the analysis of intuitive collectivity concepts. Our discussion of the formulation of the WSAT in terms of 'demonstrative social actions' will be most relevant here.

iv. A formulation rejected

Preliminary worries with the WSAT abound. First, we need some precision. Is the idea that there is a population of agents, and at least one performs some social action, or what? Further, is the claim about necessary conditions or sufficient conditions, or both? We shall first consider the claim as a claim about sufficient conditions; and as a claim about a type of action most members of the putative society must perform. As a first formulation of the WSAT, then, let us consider the following; Given a population P of (more than one) agents, P is a society if most members of P much of their waking lives perform social actions.

Fairly clearly, this will not do. Clearly, a plurality of persons could perform social actions all the time and there be nonetheless no society of which they are all members. Thus consider Fred in galaxy G, Joe in galaxy G', Xenon in galaxy G'', and Algernon in G'''. These people are in no perceptual or physical contact with each other. They have, indeed, no conception of the fact that there are other galaxies. They each interact with a number of other people in their own galaxy, and often act with these people in mind when they act. They never act, however, with one another in mind. This seems to be a clear case of Fred, Joe, Xenon and Algernon not being members of one society, and certainly not forming a society. Yet, equally clearly, Fred, Joe, Xenon and Algernon would count as a population which was a society on our first formulation of the WSAT. For given the population of these agents, most members of the population much of their waking lives perform social actions.

Our intergalactic 'population' is clearly not a society, and in our first formulation the WSAT clearly does not succeed in giving a sufficient condition for the existence of a society in terms of social actions.

Perhaps there can be intergalactic societies, perhaps Fred and Co. could in fact form a society somehow without ever meeting. Rather than speculate further about this case, let us consider its purpose as served: it has shown that we must reject the first formulation of the WSAT.

v. Another formulation rejected

In what follows we shall limit ourselves to cases in which our putative social group is a plurality of agents all of whom live within walking distance of the others. We shall introduce a group of people whose situation passes through various stages. As first they will perform no social actions, then social action will be introduced. We shall consider whether at any stage they have a clear case of a society. In the case of their social action, unlike that of the intergalactic population, many will involve reference to others with whom they are in direct physical or perceptual contact. In particular, members of the population will often refer to other members present, qua persons present, in doing what they do. We might, in order to have a label for them, call these kinds of social actions, demonstrative social actions. In relation to this stage of the group's life we shall try to test the following formulation of the WSAT: Given a population P of (more than one) agents who live within walking distance of each other, and who much of their time perform demonstrative social actions with respect to one another, P is a society.

This formulation of the WSAT, as the former, is couched in terms of sufficient conditions. We shall deal with the question of necessary conditions for societal existence separately.

Stage 1 Our population consists of a plurality of people living within walking distance of each other. We shall not ask how each comes to be where he is; we shall assume each has a language, not inquiring how. If you like we can imagine each arrived at his own place from a society in which he grew up, but he grew restless and came to settle here.

The emotions and attitudes of these people then, may have to remain somewhat obscure to us, but we can assume they are neither savage nor needy as Hobbesian men. Each grows food, and sufficient food for his needs, off the plot of land in which he has pitched his dwelling. We can picture each person living alone on the perimeter of a forest, out of sight of the others, but in walking distance of them. As the first stage, let us take it that these people, though physically close, never come in perceptual contact; each assumes he is many miles from another person. Here there are no demonstrative social actions performed, and our population does not form a society.

Stage 2. We can let our people come into physical and perceptual contact without introducing social actions. Here is one possibility: every now and then members of our population go on solitary mushroom picking expeditions in 'their' wood. When one is picking mushrooms with his head bent down, he sometimes inadvertently collides with another, who is doing the same. We might suppose that each is stunned for a while by the head-on collision, and when each awakes his first thought is not for the one he collided with, but to get back home and lie down. So they ignore each other; each struggles to his feet when he recovers, and staggers painfully home. As yet, though in a very physical sense these people have come in contact (interacted?) they have performed no Weberian social actions; and at this point it would surely seem that they have no society either. Compare in connection with this situation, Weber's example of two cyclists:

'A mere collision of two cyclists may be compared to a natural event. On the other hand their attempt to avoid hitting each other, or whatever insults, blows, or friendly discussion might follow the collision, would constitute social action' (p.113).

Stage 3. Now let us introduce Weberian social action. After their unpleasant experiences of bumping into things in the woods, the members of our population might begin to tread warily, looking up every now and then to see what was in their path. So now every so often one looks up and sees another coming toward him picking mushrooms, and moves away in order not to collide with this mushroom picker. It appears that we have a set of social actions here now, in particular what we have called demonstrative social actions. If so, we have social actions, but no society.

It could be objected that the avoiders here are not yet social actors, since for all that they do and believe the others they have in mind might be inanimate objects, like trees (which they will also avoid) or (a more appropriate example) like mushroom-picking machines. There seems to be some point to this objection. Perhaps it is sufficient rebuttal to say, by fiat, that our agents move away in order not to bump into that human mushroom-picker, seen precisely as someone with intentions, feelings, beliefs and so on. That is, in this case, though not in the case where they avoid the trees, the subjective meaning of their action involves reference to another human being as such. We might add, further, that in the human case, but not in the case of the trees, each avoider has two reasons for his avoiding action; one is to avoid pain himself, the other is to avoid pain for the other, whom he assumes will be hurt if they collide. If we add in this complex motive for avoidance, the actions involved appear to be truly Weberian social. However, if this is all that went on, I do not think that our population would thereby become a society. The second formulation of the WSAT appears to have been refuted.

Perhaps it will be objected that at the end of Stage 3, though we do not have a society, we do not have something which is strikingly not a case of a society.

Perhaps it will be claimed that this is just the kind of situation in which rational agents, anyway, will become much more of a social group. They will, surely, get together around this point; perhaps they will divide up the labour of mushroom picking, or apportion bits of territory to each individual and agree each only to work his own territory.

In answer to such an objection, I would claim that even if it is true that out of this situation rational agents will make a society, it is certainly not yet a society. This case appear to refute the second formulation of the WSAF insofar as this purports to give a sufficient condition for the existence of a society in terms of the more proliferation of demonstrative social actions. We certainly need to know more than that such actions are often performed in a population, before we decide that it is a society. It could be after all that were the mushroom-pickers ever to actually make eye contact, were they to subject themselves to one another's look, they would panic and rush away from each other. Perhaps each fled from his respective society before because he had eye contact phobia. Perhaps now he will refuse to leave his own home in case he risks encountering the looks of others. I spell out this fantasy in order to show that even if one feels strongly that the situation described contains the germ of a society, the germ needs a special atmosphere in which to flourish; and actual performers of demonstrative social actions need not as a matter of fact be able to provide the right atmosphere.

A more vivid kind of example would be that of some kind of a 'war of all against all'. Not a war formally declared, of course, and something more primitive that was ever envisaged by Hobbes in his talk of such a war. Thus let us imagine ^a chaos of agents who are now massed together in a large forest. They spend their time in the attempt to kill each other, or at least to inflict pain. Thus they bite, scratch, rape and wound those immediately close to them.

They only pause to slip away to tend their own wounds and feed upon berries and sleep. Their need for food and sleep satisfied, they are once again inflamed with anger and misanthropy at the thought of the others, and return to the fray.

Such people - or creatures if you prefer - perform a variety of Weberian social actions over a period of time, but here we seem to have described a paradigm state of nature, a paradigm non-societal state. Once again, we can propose that such a situation will not be stable, that eventually people will start caring for one another or banding together against other bands and so on. But, as the case has so far been described, nothing of this sort has happened. There is, however, a proliferation of Weberian social action; once again the second formulation of the WSAT has been presented with a counterexample.¹

1

Clearly, additional stipulations about the case could alter the picture. Suppose we were to stipulate that these people were originally at least members of a close-knit society, but hostility grew among them so that they unanimously agreed to fight it out, person to person, to the death. Against such a background we might judge (perhaps given further assumptions) that people formed some sort of social unit still. Evidently a given society or social group will often accomodate quite high degrees of hostility and conflict among its members.

vi . Necessary conditions and states of nature.

We might at this point drop the idea of providing sufficient conditions for societal existence in terms of a proliferation of Weberian social actions, and turn rather to the idea that proliferation of social actions provides us with a necessary condition for a society, or at least for a society of agents.¹ This thesis is probably more plausible. Let us now construe the WSAT as follows, then. It is a necessary condition of the existence of a society of agents that there is a population P of more than one agent in which social actions (Weberian social actions) are regularly performed.

Given such a formulation, the following problem arises. Assume that the WSAT is true, that social actions must be performed in a society of agents. Unfortunately it is also a necessary condition for the existence of certain paradigm cases of nonsocial states that there is a proliferation of social actions. Certainly, there are states of nature - or nonsocial states - in which no social actions are performed. Such was our Stage 1 above, but similarly there are states of this kind which involve much social action. In particular, the war of all against all, which seemed a very striking example of a nonsociety, has as its necessary condition a proliferation of Weberian social actions. The problem, then, can be summed up like this.

¹ Cp. Weber: 'A 'state' for example, ceases to exist in a sociologically relevant sense whenever there is no longer a probability that certain kinds of meaningfully oriented social action will take place' (p.118).

It seems that an illuminating statement of necessary conditions for a society or a social group will not also be a statement of necessary conditions for the war of all against all.¹ Since such a statement is not given by the formulation of the WSAT under consideration, it must be admitted not to give an illuminating account of necessary conditions for societies or social groups of agents, if indeed it gives such an account at all.

In general, suppose we have a notion of a kind of action such that a proliferation of actions of that kind is a necessary condition for some paradigm cases of nonsocieties; we cannot, surely, by stating that it is a necessary condition for societal existence that there be a proliferation of actions of that kind, be saying anything very helpful or illuminating about the concept of a society. We may of course be held to be saying something, since clearly some possible cases (like Stage 1 of our imaginary population) may be ruled out, but we leave in cases which are still so far from being societies it is clear very little is being said. The same goes, mutatis mutandis for the more generic notion of a social group.

¹ I do not want to rule out the possibility of a necessary condition which is illuminating though still not a sufficient condition. This is why I do not here simply say that a condition which is necessary but not also sufficient is not illuminating. Perhaps some necessary conditions might be enough to rule out clearly paradigm non-societal states. Possibly some set-ups in which animals co-exist conform to such necessary conditions. The stated conditions should be distinctively necessary for social groups as opposed to paradigm contrast cases.

viii. Conclusions.

In conclusion, if we are looking for an illuminating answer to the question 'What makes a society of a plurality of agents?' the Weberian social action thesis, in the formulations considered at least, does not provide one.

Certain desiderata for a social action concept that can provide us with a real insight into the nature of collectivities may now be stated. That is, the following would be real achievements in our understanding of collectivity concepts.

A. The provision of a sufficient condition for the existence of a social group of the form 'a proliferation of actions of kind K performed by all or most members of a population P makes P a social group or society'.

B. The provision of an interesting necessary condition for the existence of a social group of the form 'a proliferation of actions of kind K performed by all or most members of a population P is necessary to make P a social group'. Views of what are interesting necessary conditions may vary, but

I take it that the type of action concerned should not be necessarily performed in paradigm cases of nonsocial states, like the totally unrestrained war of all against all that we have described above.

In relation to B, it must of course be noted that we have here not solved the issue of whether or not social actions are indeed necessarily performed in any case of a social group made up of agents, and in a sense, because the issue cannot be solved quite readily, whether or not social actions are necessary is itself an interesting question.

I do not propose to follow up this question here.¹ My point has been this. Assume we can somehow, perhaps laboriously, show that Weberian social action is indeed a necessary condition for the existence of any society.² Then, although our result will indeed have a certain interest, its interest will be limited and will not give us any insights into what collectivity really is, and one reasons for that is that it does not even meet the condition on 'interestingness' that I have brought into play here.

It is not at all obvious that we can find a social action concept which fulfils either of the requirements mentioned, let alone both of them. My main point in this section is, however, already made. That is that Weber's notion of social action as he presents it in Economy and Society, does not appear to fulfil the requirements, and at least to this extent can make little contribution to the analysis of collectivity concepts.

Regarding the question of what social sciences are about, we have one conclusion forthwith. If it is a plausible piecemeal claim that they are about social groups and their properties, the global claim that they are about social actions in Weber's sense is highly misleading at best. On the face of it, it is at least as plausible to claim that the study of social groups and their properties would be a social study, as to claim that the study of social actions in Weber's sense would be.

1

I return to the analysis of collectivity concepts later, see pp. 167-167, and pp 298-302.

2

Or perhaps a necessary part of a sufficient condition, or an 'inus' condition, as John Mackie puts it (an insufficient but necessary part of an unnecessary though sufficient condition).

7. Social action and the subject matter of sociology.

i. Introduction

Let us now turn away from the problem of analysing collectivity concepts in an illuminating way, and indeed stop fixing our attention on these concepts altogether. Let us consider whether reference to Weberian social actions can play some other useful role in the description of the extant subject matter of sociology or (what is our real concern) in the description of natural expectations about what a social science would deal with.

I propose to discuss this question first by considering actual and possible studies of suicide.

Our discussion of suicide is meant as a counterexample to the claim, on behalf of Weber's concept, that, insofar as social science might be concerned with types of human action, then their range of interest will be exclusively the range of social actions in Weber's sense.

As far as the extant subject matter of sociology goes:

'The theoretical treatment of suicide is one of the few classical subjects in sociology... it was of great importance in the establishment of sociology as an independent discipline... as a result of Durkheim's work on suicide'. 1

Durkheim's work on suicide was particularly striking, as he himself pointed out, in that it attempted to show, using statistical studies of suicide rates in various countries, that suicide, apparently an individual and 'private' act par excellence¹, was actually closely related to quite specific larger scale features of the surrounding society; for instance, Durkheim claimed to have shown a close connection between relatively high suicide rates in a society and 'anomie' - the lack of well-defined norms in that society.

Though either the truth or the precision of Durkheim's mode of explanation may be disputed, as far as I know it is generally not disputed that Durkheim's explanations, if they are clearly established, would show that the phenomenon of suicide is a proper subject for sociology or, more generally, for a social science.

While the existence of important sociological 'suicide studies' may lend plausibility to the claim that sociology is - at least to some extent - centrally concerned with the study of human action, or types of human action, it at the same time appears to disprove the claim that it is social action which is the concern of social scientist. It is worth discussing the nature of suicide and of social scientific interest in suicide in more detail to show what is true and what is false in both of these impressions.

¹ Cf. Suicide, p.46 (English)

(ii) The definition of 'suicide'

How precisely to define 'suicide' has been an important issue for the sociologist, and definitions have differed.¹ Durkheim's definition of suicide is probably somewhat at odds with intuitive 'pretheoretical' conceptions, a fact of which he is aware and which he makes a point of saying does not worry him. Thus he insists that suicide should not definitionally involve one's aiming at his own death. We should count as a suicide, for instance, one who does not act expressly in order to bring about his own death, but for instance sees his death as an evil which will inevitably come about as he pursues some other aim, such as maintaining the honour of his battalion by continuing to fight in adverse circumstances. Thus Durkheim urges that we should not conceive the suicide as one who 'deliberately kills himself' but rather as one who 'does X' knowing that by doing X he exposes himself to imminent death.

Whether or not this a viable concept, this notion of suicide does involve the agent in doing something intentionally, and with a certain outcome of his action in mind, though he is not necessarily aiming for this outcome. We might well note here that no particular meaningful content in Weber's sense seems to be ascribed to suicidal acts in general by Durkheim. The only condition for an act's being an act of suicide is the agent's knowledge that, in X-ing, for reason R, he will bring about his own death. Whatever the merits and demerits of this account from a sociological or an intuitive point of view, it seems to square with intuitive notions at least in this - 'committing' suicide involves doing something intentionally, where the meaning of one's action is such that it is not necessarily Weberian social action.²

¹ See J. Douglas, op.cit.

² Durkheim's definition of suicide has been criticised by Peter Winch, The Idea of a Social Science (p.110) and Alasdair MacIntyre The Idea of a Social Science, (p.107) among others. For a defence of Durkheim see C.A. Bryant, 'In defence of Sociology; a reply to Some Contemporary Philosophical Criticisms' (p.100)

Now, however the definitional issue is finally decided, it would seem that any natural definition of 'suicide' will not have it being Weberian social action as a matter of logical necessity. Not unless the others one has in mind when one performs such an action may include, as sole member, oneself. It is fairly clear that Weber at any rate did not intend such a construal.¹

It seems then that insofar as suicide is a phenomenon with which sociologists have been closely concerned, and is indeed a phenomenon involving a kind of intentional action, it is nonetheless not a phenomenon which by its nature involves orientation to others. Thus, the existence of (so-called) sociological studies of suicide seems to provide a counter-example to the thesis that insofar as social sciences are concerned with human action, they are concerned with social action. Let us, however, look into this issue further.

iii. 'De Facto' social action

It may be objected that I am wrong to suppose that Weber or our hypothetical describer of social science is saying or would say that the act-types with which social sciences deal are 'analytically' social in Weber's sense, that is, that they are social by definition. They could be saying, rather, that the particular, dated actions with which the sociologist deals are in fact social. That is, they are indeed often tokens of act-types such that one may be an act of that type without being oriented to others, but they are, nonetheless, oriented to others, de facto as it were, rather than de iure.

1

Suppose, as Winch alleges, that at some deep level, all meaningful behaviour is social in some sense, say actually in Weber's sense. Then our discussion here could be taken to address the question: 'do social scientists limit their attention to actions which are social in Weber's sense "above" this deep level'.

Thus a case of suicide, for instance, may be oriented to others in various ways, even though it is true that qua suicide it does not have to be so oriented. One may kill oneself out of revenge, out of spite, out of despair that one will never be loved by a certain person, and so on.

What then of the claim that social scientists are concerned with particular acts which are in fact social in Weber's sense. Do sociological studies provide a counterexample to this claim?.

I take it that the important questions to ask are really these: do sociologists study actions which are in fact social in the relevant sense, and are they interested in such actions qua social actions?.

Prima facie, it is by no means obvious that all actual cases of suicide are in fact social actions in Weber's sense. Of course this is ultimately an empirical matter. However, one could cite the possibility of someone's killing himself in a fit of 'objectless depression'. Moreover there are many cases where the depressed state, though having a clear conscious object, does not have an object that would make an action involving it qualify it as social in Weber's sense. For example a scientist might be severely depressed over his inability to explain some phenomenon; someone might fall into despair on hearing he has contracted a fatal disease. No other person is the object of the person's depression in any of these cases.

We recall that for Weber it is not enough that in such a case the agent is influenced by the people who surround him, unless this leads to the 'subjective meaning' of the act's referring to them. Nor is it enough for thoughts about others to be running through the suicide's head - unless he decides to act in order to spite them say, or because they have spurned him. Our empirical knowledge surely does not make it clear, moreover, that there is always at least an unconscious orientation to others.

Weber might accept such an orientation as entailing that an action were social, but, assuming no problems in the concept of unconscious motivations of some kind, we surely do not know that they occur in every case.

The aetiology of the various forms of depression is a highly complex matter about which there is surely no such knowledge.

Suppose we were to agree arguendo that all suicide cases were in fact cases of social action (obviously very many are, at least). This would not show that sociologists studying these cases were concerned with suicide qua social action. This feature of all cases might never be mentioned by the sociologist, either as part of his characterisation of his initial concern, or as part of an explanatory theory, but in trying to characterise the subject matter of social science we are presumably interested in saying what the precise concern of the social scientist is.

iv. The Durkheimian approach to suicide

How then is it with sociology as we find the subject? Are those who study suicide interested in it qua oriented to others or not? If we look at the classic work on the subject, Durkheim's Suicide, and at a 'Durkheimian' approach in general, it appears that the initial interest is not in individual cases of suicide as such, but rather in differences in suicide rates, in particular in the difference in suicide rates between different social groups. As we have already indicated Durkheim thought that he could explain differential rates of suicide in different societies in terms of 'social' facts - 'social structure', currents of opinion', and the like. Given a certain type of society, as opposed to some other type, and ceteris paribus, there is more (or less) suicide per head of population.

According to this style of sociology, sociologists might be said to be interested in suicide or any other phenomenon of human action insofar as they can explain its incidence in a social group or society by reference to phenomena such as social structure and social 'currents'.

Thus these sociologists will be interested in suicide, say, qua action of a type whose incidence is somehow dependent on or influenced by certain particular kinds of collectivity states.

It is evident that for sociologists of this type, a phenomenon of human action cannot in general be known a priori to be of sociological interest. Thus it might have been that the incidence of suicide in a group was a function not of its social structure, but of the weather in the region in which it occurred.¹ Perhaps it is in fact so dependent, and sociologists who have argued on behalf of 'social causes' are mistaken. If so, then suicide studies would only be 'sociological' by mistake, as it were; they would play no part (as far as explanatory theories for suicide were concerned) in a perfected sociological theory.

This approach gives us another possible definition for the phrase 'social action'. A social action would be on this definition an action of a kind whose incidence in a society is influenced by particular states of collectivities. This seems unlikely to be extensionally equivalent to Weber's notion and it clearly gives a different meaning to the phrase.² As we have seen, Weber does not mean by 'social action' action influenced by certain states of other people, let alone action whose incidence is caused by states of the relevant collectivity.

Durkheim's approach to suicide, then, did not focus on suicide qua Weberian social action. Yet I assume that as this approach has been roughly characterised here, it would intuitively be thought of as a sociological or social scientific theory.

¹ In Suicide, Durkheim discusses and rejects this possibility, which had been mooted by some criminologists, before adducing his own explanations.

² It obviously does not fit Weber's 'individualist' orientations.

The explananda of the theory are the different suicide rates in different societies, and the general form of the theory is that these rates are a function of the kind of society, roughly, in which they occur. Nonetheless the theory is naturally thought of as a study of a type of human action.

Though theories of suicide of a more 'understanding' nature may be found,¹ the very possibility of a study of this general form makes it clear that to say that a social science will deal with types of action only if they are social in Weber's sense, and indeed only qua social in Weber's sense, is at best highly misleading, at worst simply false.

v. Approaches to suicide

If it is objected that it is unfair to use the Durkeimian approach as a counterexample to Weber's theses about the centrality of social actions, because Durkheim's approach is precisely antithetical to Weber's, our reply is that we are here considering what types of theory about human actions - by whomsoever - would intuitively count as social scientific theories, and Durkheim's surely counts as such a theory.

We now briefly consider what a more 'verstehende' theory of suicide might look like. Were one simply to catalogue or categorise motivations agents typically have for suicide, this would not so far seem to give us a social study or theory, but now suppose that according to one's results the motives for suicide turned out to divide basically into, say, two: revenge and calling for attention, one might then claim to have shown that suicides are generally social actions in Weber's sense. (This latter might after all, be true).

¹ See our next section.

² See Lukes, Emile Durkheim, Ch.9.

A book expounding these results could be entitled 'Suicide as Social Action' without clear violence to our concept of the 'social' and might reasonably be thought of as a social scientific work.

What of suicide qua Weberian social action as an explanandum? Might someone not set out specifically to discover what are the ways (if any) in which suicides take others into account? Such an inquiry might of course be undertaken. Indeed if someone took the view that he was engaged in a discipline whose central subject matter was Weberian social actions he probably would limit his enquiries into suicide, if he made any, in some such manner, but any such limit appears quite arbitrary. If one's interest is in human suicide, its modes and reasons, there seems little point in dividing the subject up in advance into the bit which considers orientations to others and its kinds and the bit which considers other orientations.

I suggest that it is plausible to suppose that where the scientist's interest in explaining or categorising suicide is restricted to suicides of a certain type, the type will tend to be more specific than action social in Weber's sense.

Consider, for instance, the following possibility. Someone impressed by famous political self-immolations by Buddhist monks and others could start an inquiry into the conditions under which a person might seriously consider and actually undertake such suicide. Mindful of a popular use of the phrase 'social action' as 'action' which aims to improve one's society, or perhaps 'political action', he could reasonably entitle his study 'Suicide as Social Action' and it would reasonably be thought of as a kind of social scientific study, but the intended topic would not be suicide qua social action in Weber's sense, in spite of the fact that all politico-social action is at the same time social in this sense. The focus would, rather, be action qua action aiming to improve one's society, or qua political action.

Before concluding, I should say that I suspect that many who find the Durkheimian approach to suicide repellent in fact hunger for a kind of account which is hardly that of a 'social science' at all. Such an account might perhaps be provided by a subject we could call 'phenomenological psychology' or it may have to be left to novelists and poets. Thus Sylvia Plath's poem 'Daddy' may seem to give one much more insight into the mind of someone with an urge to suicide than anything Durkheim ever wrote (of course Durkheim did not set out to provide any such insight). The idea of investigating 'typical motivations' and the like may give a better approximation to this sort of insight than Durkheim gives us, but studies like Durkheim's insofar as they give correct results, can surely aid the 'understanding' of suicide which one may crave. For example, if it is found that there is a highly negative correlation of suicide rates with the existence of clear-cut well-known well-established norms in a society, this in itself tells us nothing about individual motivation, but it can point to or suggest hypotheses about individual motivation which can be investigated at the individual level.

Once we get to this level, it is not immediately clear that we are, intuitively speaking, doing social science. We have used clearly social scientific results to arrive at hypotheses about individuals - only this is clear so far.

As the foregoing discussions have shown, results or investigations at the level of individual motivation may single out or focus on actions which are social in some sense; social in Weber's sense perhaps, though their Weberian socialness is not always going to be what is stressed.

The relationship between our investigations at the individual level and at the social group level can interact in more than one way. Not only can the social studies point to explanations at the individual level. Once we have the data about individuals' motivations this will serve to explain, perhaps, how the correlations between large scale variables come about.

Insofar as this would have the psychological data playing an important role in explaining the facts about groups, the psychological data could be considered 'part' of social science.

As we have seen, then, there is the Durkeimian approach to suicide, which may be implemented by psychological hypothesis, 'social' or otherwise, and there are also a variety of possible conclusions or projects on the nature of suicide which could be expressed in terms of suicide qua social phenomenon qua 'social action' in some sense. But other senses than Weber's will be relevant here. This brief sketch of ways one might study the phenomenon of suicide clearly shows that it is not the case that insofar as social studies might be concerned with types of human action in some way, they will always be concerned with suicide qua social action in Weber's sense. This is not even true of theories of a more 'verstehende' type. So a piecemeal theory of social science that says social scientists are concerned with actions only insofar as they are Weberian social actions is surely, once again, at best misleading, at worst, false.

8. Weber's 'social action' as a scientific concept.

I now want to consider the possibilities of Weber's concept in the light of general desiderata for scientific concepts such as those suggested by C.G.Hempel.¹

Two main desiderata for scientific concepts have been stressed in the literature; such concepts for one should have clear criteria of application; and they should also have a role to play in the statement of scientific laws or more modestly of scientific theories.

¹ See, e.g. his 'Fundamentals of Taxonomy'. sec.5

It might seem that we cannot possibly say a priori- in advance of the development of successful theory-that any given notion will or will not prove to be scientifically fruitful, but given a certain background of knowledge of the world one can, surely make plausible judgements along these lines, and I want to suggest Weber's concept of social action is by no means a promising candidate for a fruitful scientific concept.

Considering the range of the performer of meaningful actions, we might draw up two columns, one for nonsocial action and the other for social action (in Weber's senses).

Under 'nonsocial action' we would find: private meditation for one's own salvation; kicking a stone in anger; committing suicide in a fit of 'objectless' depression. Under 'social action' we would find: praying in private for a friend's salvation; asking a priest how he thinks you may be saved; committing suicide in order to earn someone's pity; kicking a man in anger in order to hurt him and thus get rid of one's own feeling of misery.

Thus kicking a man in anger in order to hurt him is social, kicking a stone in anger is not. Kicking a man in order to hurt him is a social action, and so is asking a priest about one's salvation, but private meditation on one's own salvation is not social action.

This is the way Weber's social action category would have us divide things up, but intuitively one might suppose that praying or meditating on behalf of anyone, oneself or another, fell into one realm of behaviour, that of religious acts, and that such acts whether social or not are likely to be understandable in similar terms. Again, the study of aggressive angry behaviour might be expected not to distinguish between stone-kicking and man-kicking insofar as its purview went. In short it is not clear that any theory is likely to need to make the broad social-nonsocial distinction in demarcating its subject matter. The onus is on the Weberian to show that

his concept divides things up into natural kinds, the Weberian social and the Weberian nonsocial.

Thus on the topic of Weber's concept of social action, well-known and often quoted (though also misquoted), I would agree with Rolf Dahrendorf when he says:

'There floats over Weber's "Basic Sociological Terms" a slight air of irrelevance: one may use these concepts if one wishes to, but one is not forced to, and it is difficult to demonstrate that one should'.¹

Except that I would say something stronger: one should not wish to use, or use, this concept unless someone has demonstrated that one should. Dahrendorf is too complaisant: sociologists and philosophers of social science should not complicate their thought and ours with a concept whose introduction has no obvious justification, and of which one is on the present showing tempted to say that it obviously has no justification.

¹ Die
R.Dahrendorf: Angewandte Aufklärung p.159.
I am grateful[^] to F.Stewart for the reference and translation.

9. Conclusions

i. Weber's notion of social action

We have seen that there are types of action other than 'social Action' in Weber's sense which appear to be worthy of the name.

I have mentioned three good reasons one might have for introducing one particular concept of social action. Two correspond to Weber's own justifications of his concept. Negative conclusions on Weber's concept emerge on all counts.

First, a certain concept of action might play an illuminating role in the analysis of collectivity concepts. These concepts are, Weber contends, inevitably going to figure in sociological thinking, and I see no reason to dispute this claim. In any event, if it is true, then the student of sociology or the philosophy of sociology may well want or need some form of analysis of these concepts in order to understand fully what he is talking about. If he needs this, then, I have argued, Weber's concept of social action is not what he needs.

Second, if we could illuminatingly mark off all those types of actions which might, intuitively, provide a focus for a social science, by reference to some precisely specified 'social' type of action, one special concept of social action might justly reign supreme in the social sciences and the philosophy of social science. We could then justify a casual use of the phrase 'social action' in these disciplines, but if there is such a kind of action, Weberian social action appears not to be it. Moreover it appears that the assumption that there is a 'social' kind of action of the required type is itself false.

Lastly, a concept of a certain type of 'social' action might pick out some obviously important 'naturally homogeneous' class of events which were intuitively 'social' and for which one might hope some science could provide explanations and/or understanding. It does not appear likely that Weber's notion of social action picks out a class of events of such a kind.

Weber's notion of 'social action' has, then failed on a number of counts. We may now note two more positive results.

Weber's notion is fairly clearly a notion of an intuitively social phenomenon, and is, moreover, something which is sufficiently strongly 'social' to make it the case that, were there a science or study which focussed entirely and deliberately on such actions, it would aptly be called a social study by reason of this focus. Since it is not part of my aim to legislate about what scientists could in fact achieve, I would prefer to admit that scientific studies aptly named studies of Weberian social action may occur. Nonetheless, as far as our results go, no impressively wide or especially illuminating piecemeal theory of social science has yet emerged.

It seems clear now that Weber's notion of 'social action' cannot provide a satisfactory global theory of social scientific subject matters. Let us summarise our argument:

Given that there are social groups or collectivities: societies, business associations, clubs, and so on, these are clearly enough social things or phenomena of some kind or kinds. Weber himself, we recall, says that sociologists cannot 'afford to ignore' collectivity concepts. Now if these concepts do apply to things in the world, what they apply to will, presumably be subjects or potential subjects for social sciences. The social sciences are quite commonly characterised, indeed, as science 'of society'.

It seems then that a global social action theory could only work if it could somehow be argued that statements about social groups, said nothing over and above certain appropriate statements about the occurrence of social actions, and the only clear way of showing this would be to show that the concept of a social group was analysable rather simply in terms of one's type of social action, in a way which did not presuppose the existence of a social group in the first place.

Now, I have demonstrated that as far as Weber's social action notion goes, social groups cannot usefully be argued to be 'nothing but sequences of social actions'. I have not shown here that no concept of social action can do the trick, but the point we have got to here is quite important for the question of the very possibility of an illuminating global theory of social scientific subject matters, to which we turn here.

ii. A result concerning global theories.

It is an important point that Weberian social action is at least a conceivable focus for scientific work. If there were a theory couched in terms of Weberian social action, it would be aptly called a social science (albeit perhaps a part of 'social' psychology). Thus any global theory must take this possibility into account. Even if this possibility is a weak one, it still poses a problem for a unitary global theory, that given collectivities are conceivable ^{foci} for science, and a science of collectivities would intuitively be a social science,

Now we have seen in the case of Weber's notion that if a global theory must embrace social groups, a theory that social sciences are about Weberian social actions cannot count as a definitive and complete global theory. However, there is the reverse problem also. Collectivities cannot be defined in terms of social actions in Weber's sense, and it is absurd to suggest that Weberian social actions can be shown to be constituted by collectivities. Nor does it seem plausible to suppose that whatever the basic elements of collectivities actually are, Weberian social actions are constituted by them. It is in short hard to see how there can be a non-trivial global theory, in terms of one precisely defined kind of thing, that embraces these two radically distinct phenomena, collectivities on the one hand, and Weberian social actions on the other.

In summary, then, if Weber's 'social action' does not provide a global theory of social science, it also makes it look extremely doubtful that any such theory can be given a priori and definitively in terms of one precisely defined kind of phenomenon, unless that phenomenon is indeed the phenomenon of 'socialness' itself. Discussion of this will be left to the final chapter.

What I have in effect just been arguing is that 'social' things are things of many different (other) kinds. This could also be supported by considering that there are evidently types of action other than Weberian 'social action' which might aptly be called 'social action' and actions of these types do not have to be at the same time Weberian social actions.¹

iii. An outstanding query.

One way in which Weber's distinction between social and nonsocial action might be criticised has not yet been dealt with. Weber presumes that not all actions have to be social. We have noted examples of actions which so far as their given description goes are not apparently Weberian social actions. Thus committing suicide when in a mood of 'objectless' depression; and kicking a stone in anger, and making a hat of a certain kind because one liked a similar one yesterday all seem to be examples of nonsocial actions. Once social action is seen to be a special type of action, the scene is set for the apparently informative claims that sociology is concerned with those actions which are social, and so on.

1 Recall the 'type' of action someone with a Durkheimian sociological orientation might focus on, noted earlier: 'action of a kind whose incidence in a society is influenced by particular collectivity states'.

Yet it has been claimed that it is in fact impossible to draw a distinction between social and nonsocial action. All meaningful action the claim runs, is social and indeed must be social in order to be meaningful. Given that even eating a banana to stave off hunger, for instance, is a social action, the prospects of giving an illuminating account of collectivities in terms of social action appear increasingly slight. Moreover the socialness of sociology is only shown at the expense of admitting the socialness, that is the necessary socialness, of suicide studies.

I wish to defend the distinction between social and nonsocial¹ action made by Weber at least to this extent - it is not the case that all meaningful action is necessarily social^{in Weber's sense}. My defence will be part of an investigation of a more general issue: is there a sense of 'social action' Weber's or another, in which all meaningful action is necessarily social action? This investigation will be the subject of the next chapter.

CHAPTER THREE

ACTION AND THE SOCIAL

1. Introduction: two claims about action.

i. Winch contra Weber.

In his book, The Idea of A Social Science, Peter Winch claims that considerations derived from Wittgenstein's account of 'rule-following' nullify Weber's proposed distinction between actions which are social and those which are not. Writing of Weber's distinction between 'behaviour which is merely meaningful and that which is both meaningful and social'¹ Winch says:

'It is evident that any such distinction is incompatible with the argument of Chapter II of this book; all meaningful behaviour must be social, since it can be meaningful only if governed by rules, and rules presuppose a social setting.'²

This chapter examines Winch's argument.

ii. Two claims about action.

Winch claims that all 'meaningful behaviour' is social. We recall Weber's definition of 'action': this was equivalent to behaviour with a 'subjective meaning'.³ Now, I shall argue shortly⁴ that Winch's 'meaningful behaviour' may be

¹ The Idea of a Social Science, p.116. Unless otherwise stated, quotations from Winch in this chapter, and page references, are to this book.

² *ibid.*

³ See our Chapter Two, Section 3.

⁴ Section 2, (iv).

equated with Weber's 'action'. That is, anything Winch would call 'meaningful behaviour', Weber would call 'action', and vice versa. Therefore we may take Winch's claims about 'meaningful behaviour' as claims about 'action' in Weber's sense. The question remains: what is the relation between 'action' in Weber's sense, or 'meaningful behaviour' in Winch's sense, and 'action' as we ordinarily understand this term?

Now, the notion of 'action' is notoriously hard to pin down. Let us narrow our quarry to the somewhat more technical 'intentional action'.¹ It seems to me that the following claims are at least roughly right (and rough rightness is all I aim for, or need, here). If someone performs an (intentional) action, then this does involve something like the agent's meaning to be doing something, or, in other words, his intending to do something. However, his meaning to be doing so-and-so, while behaving in a certain way, does not guarantee that he does so-and-so. Thus he may be hallucinating altogether, and say, think he is raising his arm - perhaps feel as if he is - and mean to be doing so, when in fact he has his arms clasped rigidly round his body. Behaviour plus 'subjective meaning' does not therefore appear to correspond to action or, more precisely, 'doing A (intentionally)' is not equivalent to 'meaning to be doing A and (physically) behaving in some way or other'. Nor, probably, is it even equivalent to 'meaning to be doing A and behaving in some way appropriate to doing A'.²

There is, then, no clear equivalence between 'meaningful behaviour' and 'intentional action' in the ordinary sense. However, there do seem to be real and close connections, and it is surely the existence of these which led to the

¹ One nodal question here is Wittgenstein's: 'What is left over if I subtract the fact that my arm goes up from the fact that I raise my arm?'. Philosophical Investigations, p.161e, Part I, section 621.

² Some connection (causal?) between the former and the latter may be necessary. For instance, suppose I make an effort to move my hand, meaning to move it, but am paralysed so no such effort of mine can move my arm. However at the same time, coincidentally, my doctor presses a button in a hand-moving machine attached to my shoulder, and my hand moves. I did not 'move my hand'.

formation of the concept of 'behaviour with a subjective meaning' by Weber, and Winch's similar notion of meaningful behaviour. For our purposes here we may simply, and not unreasonably, assume that all intentional action is behaviour with a '(subjective) meaning' in Weber's and Winch's senses, that meaningful behaviour is a necessary condition for action or rather intentional action.

In sum, we may think of the quoted claims by Winch concerning 'meaningful behaviour' as claims about (intentional) action in general. They are certainly about action insofar as it involves meaning to be doing something.

So much for action. Now, once again, let us turn to 'social action'. We are familiar with Weber's definition of this phrase. Now, is Winch arguing - in the passage we quoted at the outset of this chapter - that all action is and must be social action in Weber's sense, and that Weber's distinction between social and nonsocial action therefore marks no real distinction among phenomena? This is one possibility. But I suspect that, rather, Winch is claiming something like this: given the meaning of the qualifier 'social', all actions are and must be social in some way (even if they are not social actions in Weber's sense), and, therefore, no absolute distinction between actions which are social and actions which are not social can be made. This claim could then generate the following criticism of Weber: his stipulative definition of the phrase 'social action' introduces a distinction (between action which is 'social' and that which is not) which obscures the important fact that all action is social.

Our previous discussion has already shown that Weber's definition of a 'social action' is a stipulative one; moreover a number of alternative definitions, all having some intuitive plausibility, have been noted.¹ There remains, then, a striking general thesis which merits our attention: the thesis that all actions are (and must be) 'social' in some way or other. A related

¹ Not all of these can be called 'stipulative', since there appears to be at least one 'everyday' sense of 'social action'. Though different, the different senses for 'social action' considered all seemed relatively apt definitions for that particular phrase, if one had to be found.

thesis has also been noted: all actions are and must be Weberian-social. In view of our previous examination of Weber's notion of 'social action' it will be appropriate, here, to consider both theses together.

The truth of either claim would surely be surprising. With respect to the first, we recall the apparent 'problem of disappearing socialness' which, we conjectured, may lead individualist social theorists to employ technical notions of 'social action' in the first place.¹ It seems that there is nothing obviously social in the nature of actions as such. Of course, particular actions might be social in various ways. For instance, they might be social in Weber's sense. And so, with respect to the second claim, it would seem that, while some actions may be social in Weber's sense, not all actions need be social in that sense. However, it still could be that all action can be shown, by a priori argument, to be social - in Weber's or some other sense - at some deep level.

iii. Some initial defences of Weber.

The idea that Weber's distinction between action which is social and that which is not would be 'bogus', if all action were social in Weber's sense, attracts some quick defences of Weber, which I believe stand.

First, someone may object that it must be possible to make the distinction Weber wants to make between, say, going for a walk because one wants to see the primroses, and going out in order to visit one's friend Joe; between greeting a friend and looking at a peculiarly shaped tree, and so on. Weber's distinction certainly looked as if it could be used, even if its scientific importance seemed doubtful. And, the argument might run, even if all action were somehow shown to be social in Weber's sense, at some deep level, this would not show that Weber's distinction between social and nonsocial actions could not be drawn in slightly amended form 'above' this level.

Second, should Weber's distinction be 'bogus' in the way suggested, the unamended notions may not be without their uses. They could, for instance, be

¹ See Chapter 2, section 1(i).

used to state interesting truths such as, indeed, 'All actions are, necessarily, social actions.'

Moreover, we may want to distinguish between actions qua social actions (in Weber's sense) and actions not qua Weberian social actions. Here the Weberian notion of social action, at least, would be useful. We have in fact already found ourselves wanting to say that sociologists were not always interested in action qua Weberian social action, whether or not as a matter of fact all action is social in this sense.¹ Even if an a priori argument were to show that all actions are Weberian social actions as a matter of necessity, we would clearly still want to say that sociologists were not always interested in actions qua social actions in Weber's sense. Even if all action is social in Weber's sense, then Weber's notion in amended or unamended form might still have some use. However, useful or not, the question of its current range of application remains.

In this chapter, then, we investigate whether Winch's or related a priori arguments show that all actions are, necessarily, social - in Weber's sense or in some other way. Though Winch's book as a whole has been much discussed,² his claim that all action is social in some sense has not, to my knowledge been subjected to thorough treatment either by Winch himself (who writes briefly and obscurely here) or by his critics and commentators.

iv. The argument of this chapter.

Before getting down to details let me block out the main lines of argument in this chapter. First (in Section 2), I argue that Winch's claim that all 'meaningful behaviour' involves rules is best understood in terms of rules-of-language or concepts-as-rules. Under this interpretation the claim has some prima facie plausibility, and, moreover, gives Winch's argument a 'deep' character which

¹ See Chapter 2, section 7.

² See, for instance, the articles and books by D.Bell, A.MacIntyre, E.Gellner, A.R.Louch, and A.Ryan, listed in the Bibliography. Related papers by Winch himself include 'Social Science' and 'Understanding a Primitive Society'. His papers (1964b) and (1970) are replies to critics.

is usually ignored or at best given short shrift by critics (and is glossed over somewhat by Winch himself, as I shall show).¹

In the following sections I examine Winch's Wittgensteinian argument that rule-following involves social settings.² This argument proceeds by two stages, negative and positive. First, it makes the whole idea of rule-following problematic, in particular it becomes hard to see what it could be to follow a rule of language. Second, a positive account of rule-following is suggested, which brings in social settings. I argue that Winch is more successful in the negative part of his argument. I distinguish two positive claims about the 'social' nature of rules, one of which is less bold and, I argue, more plausible than the other. I consider the precise bearing of each claim, if true, on the social nature of meaningful action. In particular, in section 5, I note a relatively plausible form of argument which, if it worked, would give a kind of 'socialness' to all meaningful actions given either of Winch's two claims about the social nature of rules. So, using this argument, and assuming the less bold claim, we reach a relatively well-supported claim about the 'social' nature of all action.

I argue, however, that the necessity of every action's being 'social' in the sense relevant to this argument, would not support the thesis that 'studies of action are always social studies'.

2. Elements in the argument from 'meaningful behaviour' to social settings.

i. 'Meaningful behaviour' and meaningful words.

Winch argues that all meaningful behaviour is 'social' because, to be meaningful, behaviour must be rule-governed, and rules presuppose a social setting. Now, what Winch wants us to understand by 'meaningful' is not altogether

¹ See section 2, (iv); and section 3, (iv).

² The main relevant section in Winch (Chapter 1, section 8) is entitled 'Rules+Wittgenstein's Analysis', and what I call 'Winch's argument', is, evidently, Winch's interpretation of the remarks on rules in Wittgenstein's Philosophical Investigations. I discuss Winch rather than Wittgenstein in what follows because it is Winch who makes the explicit claim that all 'meaningful' behaviour is social.

clear. However, as I shall show shortly¹, one plausible characterisation of his notion of meaningful behaviour brings it close to Weber, and is roughly the following: A's behaviour is meaningful only if something of the form 'A sees himself as doing so-and-so' is true of A. Or, to hark back to our characterisation of Weber, only if something of the form 'A means to be doing so-and-so' is true of A.² Since part of our present interest is, in any case, the bearing of Winch's arguments on Weber's notion of social action, it will be well to consider what we can conclude about 'meaningful behaviour' so understood, on the basis of Winch's argumentation. Now of course what is entailed by 'A thinks he is doing X' of 'A means to be doing X' is not crystal clear, but there are a number of ways in which language and, hence meaningful words, might be held to be involved.

The strongest 'language claim' - and one which may well be espoused by Winch - goes something like this:

(1) A means to be doing X at t,

entails

(2) A has a language in which there is a word or phrase for, or means of describing, 'doing X', whose meaning A knows,

and

(3) A is somehow actually using his language at t, to say he is doing X, he is somehow 'saying to himself' that he is doing X.

We might call this the 'strong language claim' about the nature of 'meaningful behaviour' or intentional action.

There are, we may note, weaker claims that still bring in language. The minimal such claim (which we call the 'weak' language claim) stops short at (1) and (2) above - but thereby leaves the connection between language and meaning to be doing X obscure. An 'intermediate' claim holds (1) and (2) above, but

¹ Section iv, below.

² For present purposes this (necessary) condition is the crucial one.

instead of (3) has:

(3') A says, or would say if asked, that he is doing X.

Clearly, if we accepted any of these accounts as a partial account of meaningful behaviour, a complete examination of the conditions for such behaviour would have to involve an investigation of what it is to know the or a meaning of a word and, correlatively, of what it is for a word to have a meaning. For one could not act meaningfully without (at least) having the relevant items in one's vocabulary.

ii. Meaningful words and rules: a crucial assumption.

Winch does indeed initiate a discussion of what it is for a word to have meaning, and related topics (p.25). Through his consideration of this question he arrives at the more general question 'what is it to follow a rule?' (p.28).

The crucial assumption linking grasp of a word's meaning to following a rule is this: if a word has a sense, there is a rule governing its use.¹ Similarly, to grasp the meaning of a word is (at least) to grasp the rule for its use. To use a word with its meaning, is (at least) to follow the rule for its use in one's use of it. (If a word has more than one sense, there is more than one rule governing its use, and grasping one meaning is grasping one rule... and so on). I shall not discuss these claims further at this point. I now note a variety of lines of argument linking meaningful behaviour and social settings which they make available, given that a connection between rules and rule-following, and social settings, can be made out.

¹ Cf. Wittgenstein, Philosophical Investigations, p.147e, note (b):
'There cannot be a question whether these or other rules are the correct ones for the use of "not". (I mean, whether they accord with its meaning). For without these rules, the word has as yet no meaning; and if we change the rules, it now has another meaning (or none)...'

(1) Assume the stronger language theory of meaningful behaviour. Here meaning to be doing X involves using a word for 'doing X' with its meaning, i.e. (given the assumption noted in the last section) following the rule for its use. But one cannot follow a rule if there is no rule to follow. So, if rules presuppose a social setting, and meaningful behaviour involves following a rule, meaningful behaviour will presuppose a social setting.

(2) However, even the weak language theory viz. that A's meaning to be doing X involves his having a language in which he can say that he's doing X, can lead by a similar route to a claim about social settings. Knowing the meaning of a word (or phrase) for 'doing X' involves grasping the rule for the use of this word or phrase. So if rules presuppose a social setting, so does a meaningful behaviour on the weak theory.

(3) A similar argument applies given the intermediate theory. Evidently we shall have to consider how rules are argued to presuppose a social setting, if we are to be convinced of the conclusion, given any of the assumptions.

iii . Concepts and rules.

Suppose one objects to any 'language thesis' of meaningful behaviour. A tight link between meaningful behaviour and language may well not have to be made in order that 'meaningfulness' be connected with rule-grasping and/or application. Winch himself sometimes seems to want meaningful behaviour to be 'specifically human'¹ so he may want (or need) language to come into its definition for this reason. Clearly, it is a consequence of the 'language thesis' just noted that no creatures 'without language' could exhibit meaningful behaviour (though this, in the days of 'talking' chimpanzees, like the well-known Washoe, may not rule out non-humans). Winch in any case seems to accept, with others, an indissoluble link between having the idea of X and having a language in which one can express the idea of X.²

¹ see p.52, op.cit.

² Cp. Rush Rhees, 'Can there be a private language?' (1954) in Pitcher, ed, Wittgenstein. p.170.

Others, however, would deny this.¹ If one does deny this, or simply finds it dubious, one need not cease to find plausibility in Winch's line of argument.

Suppose we assume that to have the concept of X-ing one must grasp a rule (say for 'co-classifying' things).² Then we could agree that (1) if meaningful behaviour involves meaning to do X, and meaning to do X involves (at least) having the concept of X-ing, then meaningful behaviour involves rule-grasping. Moreover (2) if meaning to do X involved actually applying or using the concept of X-ing, meaningful behaviour would thereby involve rule-application. That meaning to do X involves actually using the concept of X-ing or at least having that concept is surely a plausible claim. If we accept it, together with the idea that concepts are rules, and we are shown that rules in some sense 'presuppose' a social setting, one supposes that we would have to agree that meaningful actions 'presupposed' a social setting also. What this did for the subjective content of action would remain to be examined, as would the precise nature of the 'presupposition' in each case.

iv. The definition of 'meaningful behaviour'.

I wish to stick to, and by, the account of 'meaningful behaviour' in terms of 'meaning to be doing' for the purposes of this discussion. I said earlier that I believe this represents one line of thought in Winch, and I shall now try to justify that claim.

On page 53 of The Idea of a Social Science Winch gives as an example of non-rule-governed behaviour 'the pointless behaviour of a beserk lunatic'³. Since for Winch all meaningful action is rule-governed, one assumes that the 'beserk lunatic's' behaviour will not in Winch's eyes be meaningful. Unfortunately it is not clear what or how much Winch wants us to ascribe to his beserk lunatic. Does this person mean to do things but for no good or particular reason? Or is he just 'flailing about' without any awareness of what he is doing?

¹ e.g. See Gilbert Harman, in Thought, p.84

² Cf. Kant's Critique of Pure Reason, especially the 'Schematism' chapter.

³ As MacIntyre notes, it appears to be his only example of non-rule-governed behaviour. MacIntyre (1967) p.102.

In the discussion surrounding this case it looks as if 'acting in a way involving the participant in the observance of rules' (p.52) is equated with 'having reasons for acting as he does' (p.53). If this were so, then perhaps acting meaningfully would have to be, for Winch, doing (meaning to do) something for a particular reason. Yet in earlier passages it seems clear that what is essential to an action's meaning, for Winch, is a certain bit of behaviour's itself being seen as a certain action by the agent whether or not this behaviour is engaged in for a reason.

Thus Winch remarks:

'The category of meaningful behaviour extends also to actions for which the agent has no "reasons" or "motive" at all in any of the senses so far discussed... Suppose N votes Labour without deliberating and without subsequently being able to offer any reasons, however hard he is pressed...although N does not act here for any reason (my underlining) his act still has a definite sense. What he does is not simply to make a mark on a piece of paper; he is casting a vote' (48-49).

In this case of 'N', Winch says that one requirement on N's casting a vote is that he 'must be aware of the symbolic relation between what he is doing now and the government which comes into power after the election' (p.51). Winch wishes this requirement to be seen as a special case of a requirement for any action with a sense that it 'is symbolic', it 'goes together with certain other actions in the sense that it commits the agent to behaving in one way rather than another in the future' (52).

Now, although Winch himself does not bring this out, we can probably take any intentional action insofar as it involves classification or descriptions in the relevant way as 'symbolic' in Winch's sense. If I see this as voting or describe this as voting, then I am, surely, 'committed' in a sense to calling anything exactly similar in substance and circumstance voting also. More generally, I am committed to calling anything relevantly similar voting. It is important to see that the same points would go for, say, my seeing this as eating a fig. Thus meaning (hereby) to eat a fig just like meaning (hereby) to

vote would have a 'symbolic character'. In short 'meaning hereby to ϕ ' can naturally be construed as having a 'symbolic character' whether or not, as in the case of voting, 'X ϕ 's' entails something about X's intentions and beliefs about the bearing of his action on some expected actual future happening.

This account of Winch's gist may recall remarks of Wittgenstein, and Winch may have been familiar with these remarks:

'...if a word which I use is to have meaning I must "commit" myself by its use. If I commit myself that means that if I use e.g. "green" in this case, I have to use it in others.. If you commit yourself there are consequences'.¹

Wittgenstein may not have meant to put forward here the account mooted in my preceding paragraph. Clearly, however, he meant that a certain 'commitment' is involved in the use of any general term which has meaning and was not restricted to special terms like voting (for the next Government).

It is my gloss on Winch, then, that for him cases in which an agent means (in behaving in a certain way) to be doing something are always cases of 'meaningful behaviour': in particular they are so regardless of their having or not having a particular motive for (meaning to be) doing whatever it is. We need not inquire whether, as with Weber, one's reason for (meaning to be) doing so and so is going to be 'part of the meaning' of one's behaviour, though I suspect that, as in Weber's case, it will be so regarded; but nevertheless here, as in Weber's case, behaviour will be meaningful if, and only if, the agent at least means to be doing something (in behaving in a certain way).

One reason for making this gloss is that to my mind it best fits Winch's own remarks on the matter taken as a whole.² Another - closely related - reason is that Winch purports to be giving us Weber's notion of meaningful behaviour,

¹ In the Moore Lectures, I, Mind 1954; quoted by Konrad Specht (1967), p. 166.

² As well as the passages already cited, see, on p. 117, Winch's reference - in discussion of Weber - to the 'subjectively intended' sense of some behaviour. What one means to be doing, rather than anything to do with motives, is what appears to be at issue in this passage.

and I take it that this gloss on Winch makes his account square well with Weber's notion in Economy and Society.¹ A third reason for construing Winch in this way is that the broad notion of meaningful behaviour thus brought forward appears to allow us to conclude of the widest possible range of actions that they presuppose a social setting, that is if Winch's argument about rules goes through. Moreover, the arguments we shall discuss do seem to bear as much on an action done for no particular reason as ^{on} that clearly done for a reason, just as long as in acting the agent is assumed to have a certain thought, or a certain capacity for thought or language. Lastly, this construal of 'meaningful behaviour' appears perfectly apt intuitively. It gives us, similarly, an intuitively plausible notion of meaningful action, though if by 'action' we understand 'intentional action' this phrase will in fact - as we have argued above - be pleonastic. With regard to this last point, however, the idea that an action done for a particular reason or with a particular point is thereby 'meaningful' is intuitively apt also. Evidently we might usefully distinguish, as Weber does, between different 'Levels' of meaning of behaviour. However, in what follows, we are concerned with what is true of behaviour's 'meaning' at the deepest level we have noted.²

¹ See Winch, op.cit, p.45

² Talk of an action's 'meaning' - like talk of 'social action' has some popularity currently in philosophy of social science proper (doubtless Weber is a major influence in both cases). I need not claim that the account I adhere to in this chapter is in any way privileged except with respect to our discussion here. Evidently quite a number of construals may be quite apt intuitively. For something rather different from those considered here see, for instance, Quentin Skinner's "Social meaning" and the 'Explanation of Social Action', which attempts to apply J.L.Austin's notion of an 'illocutionary act' (a kind of 'speech act') to actions in general.

3. Rule - following as problematic

i. Introduction.

Winch seeks an analysis of the concept of following a rule as a result of pursuing the question: 'what is it for a word to have meaning?' (28). He argues, (negatively) against various proposed analyses of rule-following and (positively) for one which includes the claim that rules 'presuppose a social setting'.

A look at his-and some closely related-negative arguments brings out the special relevance of what Winch has to say to language, and the 'deep' character of his (Wittgensteinian) approach to 'rules' and 'rule-following'.

ii. 'Going on the same'.

What is it to follow a rule? Winch suggests, as 'the obvious answer', the claim that to follow a rule is always to act 'in the same way on the same kind of occasion' (28). This answer, Winch states, is 'correct' but does not advance matters' (28). The reason he give is that 'it is only in terms of a given rule that the word "same" acquires a definite sense' (ibid).

I find Winch's remarks on 'sameness' particularly obscure. Evidently he has a problem of circularity in mind. Perhaps what he has in mind as correct but unhelpful is the following account of 'rule-following': to follow and conform to a rule R, is to act in the same way on the same kind of occasion, viz, in conformity to R. Clearly the reference to 'sameness' here is idle, and moreover no logical relation between following rule R and conforming to R has in fact been made out. The whole 'account' is not only irrelevant but also exceedingly trivial and circular, amounting, in effect, just to this: to conform to a rule R is to conform to a rule R.

I take it that Winch does not here intend to endorse this claim: to follow a rule R is always to act in the same way on the same kind of occasion, viz, in conformity to R. Such a claim would still be circular, as is the claim I suggested Winch may be referring to, though it would, as the other does not, claim

a logical relation between following a rule R and conforming to R, and would hence be more to the point. As far as I can tell, Winch does not want to say that if X follows a rule he necessarily acts in conformity to the rule he is following.¹ Thus he would not, I imagine, consider this latter claim to be correct. I shall return to the issue of the relation between following a rule and conforming to a rule later in this section.²

iii. Stipulating one's future path.

At this point Winch turns to a suggestion about what it is to follow a rule for the use of a word. This concerns one's own stipulative definition. It runs: one follows a rule for the use of a word, if one makes a 'conscious^S decision' about one's future use of a word, and then acts in accordance with one's decision. Winch's example of such a decision is: 'I will use this word ("Everest") only to refer to this mountain' (28). He objects that this 'presupposes the institution of the language we all speak and understand' (28). What this sort of example actually seems to presuppose, essentially, is not so much a shared language, or, of course, the English language, but rather a language which is already understood by the speaker (or thinker). That is, a stipulative definition like that in the example is made in an antecedently understood language- in terms, here, of the words 'this', 'mountain' and so on - for which the rules for the use of the words are already grasped.

We can, I think spell out the gist of Winch's objection here in terms of circularity, once again. We are asking how one may follow the rule for the use of a word. We are told merely: you can do this by (a) following the rules for the use of certain words or a certain sentence (that is, in understanding the terms of a decision) ;(b) certain other conditions. Whatever comes under (b) the

¹ Cf.op.cit., p.32, where he stresses the relations between the concepts of following a rule and of making a mistake.

² In subsections (v) and (vi).

'explanation' itself involves the terms to be explained.¹

Winch also makes what can be taken as a rather different point in connection with the suggestion that one act in accordance with a decision:

'My decision still has to be applied in the future, and it is precisely what is involved in such an application that is here in question' (29).

As I see it, this point applies best to a case where no words except the definiendum occur in the stipulative definition - where it is both stipulative and ostensive. The point then is not that my 'decision' presupposes my grasp of a language, but that 'following a stipulative definition' is following the rule the definition 'laid down' or encapsulated. Once again a circularity is uncovered in a putative solution, or partial solution, to the problem.

iv. Rules and rule-formulations.

At the end of his discussion of following a stipulative definition Winch writes:

'no formula will help to solve this problem; we must always come to a point at which we have to give an account of the application of the formula' (p.29).

It is appropriate here to note some issues which tend to arise in discussions of rules and rule-following² and which Winch could have had in mind in making these remarks.

¹ Winch's point here about rules of language is paralleled by a point about rules in general which I take up in the next section.

² The literature on these notions, in addition to the writings of Wittgenstein and commentators on him, is rather vast. (Some of this interest is connected with the use of the term 'rule' in transformational-generative grammar, particularly in relation, therefore, to the specification of the syntax of natural languages in terms of what have come to be known as 'rewrite rules'. There is also an interest in rules in jurisprudential writings, in connection with the nature of laws. See, for example, the following articles and books (details in the Bibliography):

M. Black, (1958)

N. Chomsky (see, for instance, Aspects of the Theory of Syntax)

D. Emmet (1966)

J. S. Ganez (1971)

N. Garver (1967)

R. D. Gumb (1972)

H. L. A. Hart (1961)

R. Harré (1974).

Consider first the following claim: The clear cases of rule-following are cases where someone is aware of a verbal rule-formulation, and due to his understanding of it, and his desire to do what the rule in question says he is to do (perhaps because he knows disobedience will be followed by sanctions) he does or tries to do what the rule says he is to do. Cases unlike this paradigm in not involving explicit rules are not, strictly speaking, cases of rule-following at all, or at best they are far from central cases. Let us call this claim the view of rules as formulae.¹

Now consider a case where there is no explicit rule but which might seem sufficiently similar to one where there is such a rule, for this model to cope with it. Say I see you writing down a sequence of numbers thus: 2,4,6,8... I may then, without obviously 'saying anything to myself', in particular, without formulating any rule in words, take your pencil and continue the sequence thus: 10,12;14. Now for many people this case would have the following feature. They could easily have formulated the rule they considered relevant before continuing the sequence. Thus I might have said - though I did not - 'Why, you are adding two to the number at the end of the sequence'.

In such cases, even on the view of rules-as-formulae, it may seem reasonable to talk of 'rule-following' at least by extension. The basic reason for this is that the putative rule-follower could easily, if asked, have cited the rule he considered relevant before going on in accordance with it. From here assimilation to the view might proceed as follows: we can reasonably think of a person in this situation as subconsciously formulating the rule to himself, before proceeding. And perhaps similar moves could be made with respect to rules which, once noted

¹ A view somewhat like this has been expressed by Max Black in his 'Notes on the meaning of "rule" '. However, I do not mean to focus on Black's claims here. I am using this bald statement of a view in order to make some contrasts sharp.

and stated, we seem to follow 'habitually' or 'without thinking', as a chess-player follows the rules of chess. It might be argued - reasonably or unreasonably - that the formulated rules are always, somehow, 'at the back of one's mind' ready to be brought out into the light at will, and (therefore) somehow at every stage underlie one's actions in subconscious 'formulations'.

Now on the view of rules-as-formulae, talk of 'using a word with its meaning' as following a rule for the word may seem highly problematic, in spite of the possibility of assimilating to it at least some cases which prima facie seem not to fit. For what would the rule for the use of the word be? How would it be formulated? And could one explain the grasp of a word's meaning and consequent use of the word in terms of a rule-formulation?

We might note the uselessness of appeal to the idea of a word's definition in other terms of a language, as in: ' "vixen" may be defined as (means the same as) "female fox" '. For one thing we may point - as Wittgenstein did - to cases of 'everyday' words that no one can explicitly define this way. Examples may be 'red'¹, 'game'², 'gold'³. And also proper names, like 'Everest'.⁴

Second, it is clear that there must be some point somewhere, for any language, at which the meaning of a term of the language is graspable by some means other than such definition. Otherwise the language learner would be caught either in a vast 'definitional circle' or in an infinite regress.

Assume, then, that every word in a natural language has its meaning and that

¹ Cf. Wittgenstein op.cit. Part I, sections 72 and 73.

² Cf. ibid, sections 66-68, 75.

³ Cf. Saul Kripke 'Naming and Necessity'.

⁴ Cf. Saul Kripke, op.cit., and also 'Identity and Necessity' passim.

one who knows the language knows the meanings of its words and sentences.¹ It appears to be simply false, of many terms in English, that English speakers could easily cite definitions of them in other terms of English; moreover, it seems clear that there has to be some mode of grasping meanings which does not involve such definitions, in order that the definitional circle be avoided. And if one term in a language can be primitive in the relevant sense, then presumably many can.²

In sum, it is clear why definitions of the kind in question cannot be appealed to in answering the doubts of the believer in rules-as-formulae. None of us 'can, if asked, easily cite' definitions of all the terms of our language.

Evidently, however, one may break out of the definitional circle with so-called ostensive 'definitions'. And as we have seen, there are linguistic substitutes for 'pure' pointing. Thus 'what we all did in the garden yesterday, that was a game' or 'That apple is red'. However, the rules-as-formulae view may well not be willing to buy these as the rules that are grasped when language is learnt. For one thing, these 'rules' may seem too particular. Even if that apple was red, it is not the only red thing. So what are the rules for 'red', 'game' and so on?

So we may move on, in desperation, to something like this: 'Games are called "games" '. Or 'Use the word "red" of red things'. I say in desperation, but these seem to be the best we can do - what else is there? Yet it seems absurd to say

¹ In the case of the language of a group, like English, any given person may, of course, lack a fair amount of vocabulary at a given time. One assumes a person at least knows the meaning of all the words in his 'idiolect'.

² I have no need to claim that words in natural languages can never be defined or analysed in terms of other words of the same language. Such definition surely can happen and does, particularly with technical terminology, e.g. the term 'sibling'. The denial of this possibility is surely a bothersome consequence of a denial of an 'analytic-synthetic' distinction such as Quine and his followers espouse. See W.V.O. Quine, 'Two dogmas of Empiricism', in From a Logical Point of View.

the language learner 'could easily cite' these rules as the truth about usage dawns. Perhaps in a sense he grasps these rules, but he surely need not be able to formulate any rule at the time, say, he learns what 'red' or 'toe' or 'dog' or 'Mama' means. In particular, only once he already knows what 'red' means can he make use of formulae for himself or understand the sentence 'Use the word "red" of red things'. And he surely may know what 'red' means in advance of being able to formulate that sentence (for which he needs the word 'word', for instance).

In sum, it appears that grasping the point of an ostensive definition, and hence the 'rule for the use of a word', is not to be equated with being able easily to give a verbal formulation of a rule. More generally, the existence of a tight connection between grasping a rule for the use of a word and being able to cite a rule-formulation looks increasingly improbable.

We have already noted Winch's contention that we cannot give a noncircular general account of following a rule for the use of a word in terms of understanding a formulated rule (or decision) since the latter already involves the former.¹ We now see that the model of understanding a formulated rule and doing what it says one is to do is simply the wrong model for 'following the rule for the use of a word'.

So we now find ourselves with this 'dilemma': either accept the rules-as-formulae view as a correct account of rule-following in general, and reject Winch's idea that the phenomenon involved in every use of a word with its meaning- and in grasping the meaning of a word - is a case of following or grasping a rule; Or, reject the rules-as-formulae view.

Now, even if we took option one, we could still argue for the importance of Winch's topic to that of 'rule-following' proper, and also-of course-to the philosophy of language and theories of action and meaning. For the most fundamental phenomenon here seems to be whatever, rule-following or not, is

¹ Subsection (iii) above.

involved in understanding a word, using a word with its meaning and so on.

Having got this far, however, we might as well stick to Winch's way of putting things. For one, it is not obvious that the rules-as-formulae view is right. Again, there may well be a spectrum of cases in which we can talk of rules, and Winch's way may not be untoward. Indeed I do not find it initially odd to say that in using the word 'red', say, with its meaning in English, I follow a rule for the use of the word. Finally, I see nothing that will obviously be lost in speaking of rules in what follows, provided we bear in mind the real issues and distinctions we have noted here.

Let us, then, try to summarize the situation while retaining Winch's use of 'rule'. Let us continue with the idea that there are - or may be - at least two kinds of rule-following which can be investigated, where one is more primitive than, and indeed appears to be a condition of the possibility of, the other. The first kind is, roughly, 'following a formulated rule'; the other kind is, a more primitive kind distinct from that. It is the more primitive and fundamental kind of rule-following which is Winch's ultimate topic, and it is this, in particular, which is argued to presuppose a social setting.

Before continuing I should stress that this is my gloss on the topic - or ultimate topic - of Winch's argument. Winch himself is not always very clear on the distinction I am making. Though he does at one point make it clear that not all rules have to be explicit. He writes:

'The anarchist...eschews explicit norms as far as possible...
but this does not mean that we can eliminate the idea of a
rule from the description of his behaviour.'¹

However, he goes on, in indicating how varied are the types of rules that can be seen to be involved in our behaviour, to obscure the main, deep issue. He remarks, for instance, 'The anarchist...makes a point of not being governed by explicit, rigid norms.'² But if 'explicit norms' are not at issue, centrally, nor are such

¹ p.52

² p.53

things as the general principles on which one acts, like altruism or anarchy. So this particular remark is at least to the side of the main issue, which may thus get left behind. I think this is a pity, since Winch's - and Wittgenstein's - topic is both difficult and fundamental, and Winch's claim about meaningful behaviour involving rules (using a link through concepts or language) cannot be easily or glibly refuted.

Consider in this connection a comment by Myles Brand on the claim that 'an action...is performed only if a rule is followed'.¹ Brand writes '...what of the case in which a man deliberately acts so that he obeys no rule?' (Brand's stress). This is presumably intended to invoke a case similar to that of Winch's anarchist, though Winch carefully talks of his anarchist as eschewing explicit norms. For on Winch's view of the range and extent of rules, to act with the intention to (meaning to) obey no rule could be considered paradoxical; intention itself will be instinct with rules - and so, in particular, will the intention to obey no rule. This intention will involve a use of a word synonymous with 'rule' or an application of the concept of a rule itself! Or so I think Winch would say, as he could have said in his anarchist case too.

In fact we know how Brand's description would most naturally be taken in real life, and how the following - to make much sense - would normally be construed: 'Persons...have the power to act deliberately contrary to established practices, traditions and rules'.² Nothing in Winch's argument - in its deepest strand - purports to deny that there can be 'social reformers, inventors, and

¹ in Brand, ed. The Nature of Human Action p.17. Brand's comment occurs in a discussion not of Winch but of A.I. Melden's views (particularly as expressed in his paper 'Action'), which have some affinities with Winch's and also acknowledge a debt to Wittgenstein. Brand's comment also exemplifies clearly a possible reaction to Winch.

² Brand, loc.cit.

innovators in general' (ibid). Winch himself certainly does not ignore inventions and revolutions in theory and practice, indeed he uses their occurrence as the key fact that supposedly shows the impossibility of a predictive social science.¹ What he can be seen to argue is that, at a very deep level, every innovation insofar as it involves meaningful behaviour has as its basis something established - the concepts and/or language of the agent.

A final point which may be made here, is this. We normally distinguish between those sequences of actions which are, and those which are not, cases of following one and the same rule. Winch presumably wants this distinction, and this desire is clearly not inconsistent with the desire to see every meaningful action as involving the application of some rule. Having put Winch's concerns in focus, we continue to look at the question: What is it to follow a rule (in general) and a rule for the use of a word, in particular.

v. Meaning and doing.

In this section I introduce a special problem for the idea of following the rule for the use of a word, something not discussed explicitly by Winch. Once again, this time from a different angle, the rules-as-formulae view of rules will be shown to be inapposite for language. And a success-grammatical sense of 'following a rule' will be suggested as most apposite to the kind of 'rule-following' involved in using a word with its meaning. The nature of following the rule for the use of a word, if there is such a phenomenon, will remain, however, extremely obscure, as I shall show when I pick up Winch's thread again in the next section.

Consider the view that using a word with its meaning is not doing something. More specifically 'to use a word with its meaning' or 'to apply a concept' are not pieces of 'meaningful' behaviour in Winch's sense, even though our language

¹ See Winch pp.93-4; for a critical discussion of this see: Margaret Gilbert and Fred Berger (1975).

in talking of these 'episodes' may sometimes make it look as if they are. This implies that we can posit one 'conceptual episode'¹ or application of a concept in thinking or perceiving, without at the same time having to posit some prior, or concomitant, or supervisory, further such episode. It implies, also, that we can posit the 'using of a word with a certain meaning' without at the same time positing some prior or concomitant or supervisory further use of a word, or some additional conceptual episode. According to this view, then, the model of a regularity in meaningful behaviour, that is, behaviour which is accompanied by some thought or intention, which is one's meaning to do so-and-so - is not the model for the repetition of concept application or word use.

This view seems to me to be quite plausible. Using a word with its meaning can surely be involved in - and in part constitute - one's having a certain thought (even if not all thoughts are 'in' language).² But can 'having a certain thought' be a meaningful action? In particular, must it be? It seems clear to me that it need not. Surely in order that I have a certain thought it is not necessary that I first, or at the same time, mean to be doing anything. In particular

¹ I take the term 'conceptual episode' from Wilfred Sellars, who I heard using it in a lecture given at Oxford some years ago. I do not recall the context, but this phrase appears very apt here. 'Episode' has much less of the flavour of meaningful behaviour than does 'application'.

² Cf. Wittgenstein, op.cit. 'One wants to say a significant sentence is one which one can not merely say, but also think', I, section 511, p.140e.

I do not have to mean to have that thought.¹

The phenomenon I believe is referred to by the locutions 'S used word "W" with its meaning' (or '...in the sense of "----" ')² should not be confused with that of 'speaker's reference' as discussed by Saul Kripke³ or with that of a speaker's 'meaning that p' in producing or doing X on a particular occasion, as discussed by Paul Grice.⁴ These latter phenomena both essentially involve 'meaningful behaviour' in a way our quarry does not.

The view that using a word with its meaning is not necessarily a case of meaningful behaviour has one clear advantage in the context of this discussion - it avoids an obvious infinite regress for an account of meaningful behaviour in which this involves 'occurrent' thoughts or 'speech'.⁵ Since this kind of account of meaningful behaviour is what is needed to get the strongest possible link between meaningful behaviour and social settings - as will become clear in the final section - it is worth pursuing the view that to use words with their meaning is not itself to perform a meaningful action in connection with Winch's claims. As I have said, this view about using a word with a certain meaning appears to be independently plausible.

¹ I think it was William James whose brief description of the way our thoughts emerge was: 'the thought comes'. It is surely not our having thoughts-in general- which shows us to be 'free agents'.

² Cf. Wittgenstein, op.cit. section 508, pp. 139-140e.

³ See S.A. Kripke, 'Naming and Necessity' footnote 3, p. 343, and 'Speaker's reference and semantic reference' passim.

⁴ See H.P. Grice, 'Meaning'; 'Utterer's Meaning and Intention'. See also S. Schiffer, Meaning.

⁵ Cf. E. Anscombe, Intention, p. 47, section 27, on the regress in supposing that what action is being performed by an agent is determined by 'an interior act of intention' (my emphasis).

Let us now assume that the model of a 'meaningful action', as this is defined here, is not the model for concept-application or - in particular - for the use of a word with its meaning. There are some important consequences for the nature of rule-following. If the claim is right, then there must be at least two radically different types of rule-following, assuming, with Winch, that using a word with its meaning is following a rule. These two kinds of rule-following are - roughly - those involving meaningful behaviour (the 'rule-applications' are (also) meaningful actions) and those not. It is an interesting question how these two kinds of rule-following - if such there are - relate to the distinction we made earlier in response to the 'rules-as-formulae' view. That is, the distinction between (a) following a rule by doing (or trying to do) what a certain verbally formulable rule says you are to do, because this is what the rule says you are to do, where you formulate the rule explicitly and consciously in advance of doing what the rule says (or at least you could have formulated the rule easily), and (b) following a rule where you could not have produced a statement of the rule in advance of already following the rule.

Now it is hard to see how I can be said to act (or behave) as if I were following an explicit rule, if I am not (intentionally) doing something. For I do not think one can mean to be following an explicit rule in ϕ -ing, if ϕ -ing is not a case of one's own action. Thus it is most natural to write: I mean to be following a rule in doing X, where 'in doing X' is understood as a case of intentional action. Similarly, one surely cannot ϕ for such-and-such a reason, if ϕ -ing is not an intentional 'meaningful' action one performs.¹ It appears, then, that mere 'episodic' rule following, if there is such, must fall into a category which cannot easily-perhaps not even intelligibly - be assimilated to that of meaning (by ϕ -ing) to do what an explicit rule says one is to do.

¹ Cf. Anscombe, Intention

Turning to the case of language in particular ; in discussing indefinable terms, we noted that the category of rule-following as going by understood formulae seemed not to be exhaustive; if there was indeed a rule for the use of these terms, known by those knowing the language, following it could not involve reference to explicit formulae, conscious or unconscious. The two theses being considered converge. These are (a) A person's use of a word with its meaning cannot in general be seen as a case of meaningful behaviour, (b) A person's use of a word with its meaning, if a case of rule-following, cannot in general be seen as or as very like a case of deliberate or attempted conformity to an explicit rule. In fact (a) entails (b), or so we have argued.

We cannot say- perhaps - that all these considerations clearly refute the claim that to use a word with its meaning is to follow the rule for its use. By now, however, the nature of this rule-following, must at least seem highly problematic. Even if Winch would not agree with the suggestions made in this section, he does want to reach this result.

I believe that the depth of the problem may be exposed further by first adopting a stance regarding rules-for-the-use-of-a-word which has some prima facie plausibility, and then showing how this approach, which at first sight seem to have the shadow of substance, in fact hardly has that. The idea is that the relevant notion of 'following a rule for the use of a word' has success-grammar.¹ I should first say something about the possibility of a success-grammatical sense of 'following a rule'. Clearly 'X is following rule R' is at least not always success-grammatical. Consider an example. Jules intends to doff his hat when and only when he meets a lady. For a while he succeeds in carrying out this policy. Then he mistakes John Doe for a lady and doffs his hat at him; he does not realise Teresa Roe is a lady so fails to doff his hat at her. In this case,

¹ By 'following a rule has success-grammar' I mean that 'X follows a rule R' entails 'X conforms to R' - The phrase comes from Gilbert Ryle's Concept of Mind: I will not discuss the relation of my own use to Ryle's here.

we would still say, I think, that Jules was following the rules: doff your hat whenever and only when you meet a lady. It is true that Jules conforms to a certain rule in all this, that is, the rule: always doff your hat when you think you are meeting a lady, not otherwise. But though this rule may be a good 'rule-of-thumb' to enable him to conform to that rule he intends to conform to, it is not the rule he intends to conform to. In the usage envisaged here, then, 'X is following rule R' does not appear to have success-grammer. It might be argued, however, plausibly I think, that it is by no means clear that 'X is following rule R' never is used in a success-grammatical sense. Might one not naturally say, for instance, 'I was trying to follow the rule "never eat between meals" but I couldn't stop eating chocolate bars during the morning'? And Jules in the previous example might complain 'I am trying to follow the rule "doff your hat when and only when you meet a lady" but there are so many men who look like women around that I don't think I always succeed'. In the chocolate-eating case failure to follow the rule is the result of a failure to carry out one's resolve; in the hat-doffing case the failure stems from a mistake in judgement.

It seems possible, at least, that there is an ambiguity in the use of 'following a rule' and related terms in everyday speech, such that sometimes they are used success-grammatically and sometimes not. At any rate, there seem to be two concepts available here - both of which may at times play a role in our thought. With the first 'X follows a rule R' entails 'X makes no mistakes in conforming to R'. With the second, there is no such entailment.

At one point Winch could be taken to be claiming a non-success-grammatical sense for 'following a rule':

¹ 'If it is possible to say of someone that he is following a rule that means that one can ask whether he is doing what he does correctly or not. Otherwise there is no foothold in his behaviour in which the notion of a rule can take a grip; there is then no sense in describing his behaviour in that way, since everything he does is as good as everything else he might do, whereas the point of the concept of a rule is that it should enable us to evaluate what is being done' (p.32).

Now, it is true that if and when 'X is following a rule R' entails 'X is conforming to R' one can not properly ask, say, 'so X is following a rule: is

¹ Cf. Section (ii) above.

he getting things right?' For if he is following a rule (s.g. (= success-grammar) sense) he must be getting things right. But this does not mean that there is 'no foothold in his behaviour in which the notion of a rule can take a grip' (my emphasis). Winch says: 'the point of the concept of a rule is that it should enable us to evaluate what is being done'. But the concepts of rule and rule-following are in any case distinct. Of course it is true and important that the concept of a rule is such that it allows for a concept of 'following a rule', or more generally relating to a rule, which lacks success-grammar. A rule is - intuitively - such that one can diverge from it in one's behaviour. And if one somehow 'aims at' convergence, then one may fail. However, as far as the s.g. concept of rule-following goes, this fact about rules is not somehow denied by it. It is the concept of someone who does conform to a certain rule, who, if he aims at conformity, is successful. The s.g. sense of 'X follows a rule(R)' encapsulates an evaluation of X's behaviour. It evaluates it precisely as conforming to a certain rule, (R)

So much for the existence of a (possible) success-grammatical sense for 'following a rule'. We might give a partial account of it thus:

X follows a rule R on a particular occasion if and only if

(1) one conforms to R in proceeding as he did on that occasion, and (2) some further conditions hold.

and

X follows a rule R on a number of occasions, $o_1 \dots o_n$, if

and only if (1) one conforms to R in proceeding as X did on those occasions and (2) some further conditions hold.

Now, it seems plausible to suppose that if any kind of rule-following is involved in the phenomenon of using a word with a certain sense, it must be rule-following in the success-grammatical sense. The following considerations indicate

this. It seems that using a word with a certain sense (e.g. using 'chat' in the sense of 'cat') is something one either does, or does not do. Thus what is opposed to using a word in a certain sense is using a word in no sense at all. The ways of 'making a mistake' in language are various, but in the present context the place where, say, help and correction from others is necessary is where one as yet attaches no sense to a phrase at all, one just mouths it, say, or says 'I still don't know what it means'. So someone might ask 'is a daffodil a kind of animal?' in an attempt to get clear what 'daffodil' means - in our mouths. We may then, 'correct' his 'mistake'. But now suppose we find someone who evidently uses the word 'daffodil' for some reason as a word for 'cat'. We may, of course, say that he is using the word wrong, by this meaning that 'in English' 'daffodil' does not mean 'cat'. But in certain circumstances - those currently under investigation - we would at least say of him 'he means 'cat' by 'daffodil'.' And in his uses of the word, with that sense, he would not be making a mistake. A final consideration here is simply this. If using a word with a meaning is not a meaningful action, but just a kind of 'episode' it is hard to see how the idea that a mistake may be made could get a grip - except in cases like that where my use of a certain term is simply being said to be different from some other person or persons' usage. But this is surely not what is at issue.

It seems then that the most promising line of approach to rules-for-the-use-of-a-word is the success-grammatical one. So we might consider the proposals that:

- (a) one follows a rule R for the use of a word 'W', in general, only if one uses 'W' in conformity to R.
- (b) one follows rule R for the use of a word 'W' on a specific occasion only if one uses it in the relevant way, as part of a practice of using it in that way generally.

Unfortunately, any 'conformity-plus'¹ account of following a rule meets with certain 'paradoxes' concerning conformity. These bring the problem of saying what rules for the use of words are into the sharpest focus.

vi. Paradoxes of conformity.

We now, in effect, return to the 'success-grammatical' account of rule-following set aside in the section entitled 'going on the same'.² For the sake of continuity in the discussion it will be useful here to restate the success-grammatical account just given in terms of 'sameness' also. Thus (for case (a)) we have:

- (a') one follows a rule R for the use of a word 'W' in general,
only if one uses 'W' in the same way on the same kind of
occasion, viz. in conformity to R (plus some further condition).
(b) can be restated similarly.

In what follows, then, it is possible that the argument will in effect be a clarification of Winch's own worries about the vacuity of accounts in terms of 'going on the same'. In any case the argument is important and perplexing. It encourages consideration of almost any attempt, however initially implausible, to salvage the idea of following a rule for the use of a word. Let us, then, consider the problems involved in (a') above.

1. An apparent epistemological problem. There might at first sight seem to be a considerable epistemological problem in assessing whether the 'sameness' condition is satisfied. This problem can be posed by considering, first, cases where it does not appear to arise. It seems at first sight that it will often be obvious that someone is saying the same thing (uttering the same word) on the same kind of occasion. Thus suppose someone says 'chien' only when (and where)

¹ It appears that some further conditions will be necessary. Intuitively following a rule R and 'merely' conforming to R are distinct phenomena.

² Subsection (ii) of this section, above.

there is a dog passing by. Then he is - it would seem - clearly doing the same thing on the same kind of occasion.¹ Or, in other words, this use of the word conforms to an easily stateable rule, that we would express as something like: 'chien' is only said of dogs. Here, we have been able to re-identify situations in terms of our concept or rule for 'dog'. The claim we are interested in, however, is that as long as behaviour accords with some rule it is a candidate for rule-following behaviour. And certainly we want to allow that men can innovate. Say someone in a group starts using the term 'MCP'. Might he not use this term in a rule-following way, or at least use it in the same way on the same kind of occasion? But perhaps we have ~~no~~ way of specifying what way by saying something of the form: 'MCP' is only said of ϕ s, where ' ϕ ' is a term other than 'MCP'.

Surely we don't want to say a term 'T' is used in a rule-following way only if it is used in the same way as a pre-existing term in some language. We want to allow for innovation; and in any case this smacks of a vicious circularity. It surely pre-supposes that the pre-existing term is used in the same way on the same kind of occasion. Ultimately, then, the epistemological question will be raised of the pre-existing term: how do we know it is used in the same way on the same kind of occasion? It can't be by correlating it with the use of a pre-existing term. At some point at any rate, we must run out of such terms. This epistemological problem can be solved, but at the risk of presenting a more intractable problem. Unfortunately this latter problem is with us, since the solution to the former problem consists only in pointing out certain quite indisputable facts.

2. A paradox. As Winch puts it, 'any series of actions which a man may perform can be brought within the scope of some formula or other if we are prepared to make it sufficiently complicated' (p.29. My emphasis). Can it be

¹ In Chapter 2 of Word and Object, W.V.O. Quine argues, in effect, that there are a number of descriptions under which this person can be said to be doing the same thing on the same (kind of) occasion. This is consonant with the point we are about to make (see 2 below).

proved that any finite sequence of actions can be covered by a formula? I suggest one foolproof way to provide a 'formula' for a sequence of actions: take the given sequence, and stipulate that to begin with one performs the actions of the sequence, in order, and then say that one next does...(what you will, maybe something quite different in kind from the actions so far 'covered' by your rule). We may note that as far as someone's writing out a sequence of numbers is concerned we need not resort to this trick. It can be proved that a standard kind of function can be found which will generate any finite sequence of numbers (however 'randomly placed' the numbers look).¹

I am not sure how to back up the more general claim (that any series of actions is subsumable under some formula or other) except by reference to the trick referred to. In any case it does seem quite safe to conclude that any series of actions can be covered by a formula. The same surely goes for a series of 'episodes' that are not necessarily actions, as long as each episode can be identified under some description - not necessarily the same description. The 'formula' first posits the episodes in the sequence in which they occurred, and then posits any other sequence of episodes whatsoever. If this is so, then the notion of conformity to rule would appear to be quite empty.²

In answer to the question: under what conditions is one conforming to a rule in some series of actions? We now have: under every set of conditions. Someone might point out that a 'conformity' conception of rule-following still has some bite. For not any conception of what rule a person was conforming to would be correct. Clearly, if I say 'ouch', stepping on a pin, and then go to do some work, I am not thereby following (conforming) to the rule. 'Never say "ouch" when you step on a pin'. Thus even this conception does not leave things

¹ More precisely, for any finite sequence of numbers $a_1, \dots, a_n$, there is a polynomial function $f(X)$ such that $f(1) = a_1, f(2) = a_2, \dots, f(n) = a_n$. I thank Saul Kripke for pointing this out to me and showing me the proof.

² This is reminiscent of a point Leibniz makes in the Discourse on Metaphysics, the topic being the uniformity of nature. See also Wittgenstein, Notebooks 1914-16 tr. Anscombe p.83e.

completely open.¹

3. Another paradox. Another closely related problem for the view that takes rule-following as conformity to rule, a problem that Winch does not explicitly state, is that any finite series of actions can be brought under an infinite number of rules. (The same will go, once again, for any finite series of episodes.)

One can show this as follows. Accept that every finite series of actions can be subsumed under some rule or other. Now suppose I have done a, and b. Then I do something else, call it 'x'. Whatever x is, a rule will cover my doing a, b, and x. For every finite series (a,b,c; a,b,d; a,b,e, and so on) is covered by some rule. Each possible continuation produces a new series, and each series is covered by a rule. But then the original series was covered by as many rules as there are possible one-act continuations. But more than that, since once we've done a, b, and say x, we again have a choice of what to do. And whatever we do will be covered by a rule. So the original series was covered by as many rules as there are possible continuations at any stage. Assuming that there are an infinite number of possibilities even at the third stage, we can say the series a,b, is covered by an infinite number of rules.

In conclusion: were we to say that a man follows a rule if and only if he does the same thing on the same kind of occasion according to some rule (not necessarily a rule he states or could state) then we will find no sequence of actions or episodes, indeed no given action, or episode, that is not an instance of rule-following. Moreover, this conception gives us little line on the idea that in performing a series of actions someone might be following a particular rule. For since any finite series of actions is covered by so many rules, we have this choice. Either agree that if X's actions are subsumed by R,

¹

This, however, will be seen to be cold comfort shortly.

X follows R, and accept the consequence that in each series of actions he performs X follows innumerable rules at the same time in performing those actions. Or deny, the consequence, and hence be forced to deny the original assertion. Suppose we find the consequence unpalatable, This whole account certainly seems not to square with any intuitive notion of rule-following, as opposed to merely conforming with a given rule. As Winch says: 'That a man's actions might be interpreted as an application of a given formula is in itself no guarantee that he is in fact applying the formula.'¹ Where, then, do we go from here?

The assumption that there are some conditions in addition to rule-conformity might be assumed to remove the sting of this objection. The fact that any series can be seen as rule-covered is, however, very worrying. Are the extra features going to be such that they do not allow rule-following to be a merely subjective phenomenon? That is one problem we are left with. Since the conformity condition has absolutely no bite when taken alone, it seems that all weight must be left with whatever supporting conditions we can find, if we are to give a satisfactory account of the s.g. sense of 'following a rule'.

Thus our 'solution' to the epistemological problem - how are we to tell when someone's behaviour does conform to a rule? - leads us to see the crucial importance of 'further' conditions on following a rule. There is no need to find a way of telling when someone conforms to a rule - all ways of behaving, all episodes, conform to an infinite number of rules. What then can it be to follow a rule (success-grammar sense), apart from conforming to a rule? This question now assumes the centre of the stage, and the notion of conformity to some rule is relegated to a minor role, even in an account of the success-grammatical sense of following a rule. We now turn to Winch's positive account.

4. Winch's positive account of rule-following.

i. Introduction: Winch's examples.

Winch's positive remarks are even briefer than his negative ones, and, I find, less suggestive and satisfactory. However, in this section I do what I can

¹ Op.cit.p.29

with them, formulating a set of clear conclusions (which may not exactly be Winch's) and relating these as I can to the nature of meaningful behaviour.

Winch centres his positive remarks around an example, which I shall quote in full.

'Imagine a man - let us call him A - writing down the following figures on a blackboard: 1 3 5 7. A now asks his friend, B, how the series is to be continued. Almost everybody in this situation, short of having special reasons to be suspicious, would answer: 9 11 13 15. Let us suppose that A refuses to accept this as a continuation of his series, saying it runs as follows: 1 3 5 7 1 3 5 7 9 11 13 15 9 11 13 15. He then asks B to continue from there. At this point B has a variety of alternatives to choose from. Let us suppose that he makes a choice and that A again refuses to accept it, but substitutes another continuation of his own. And let us suppose that this continues for some time. There would undoubtedly come a point at which B, with perfect justification, would say that A was not really following a mathematical rule at all even though all the continuations he had made to date could be brought within the scope of some formula. Certainly A was following a rule; but his rule was: Always to substitute a continuation different from the one suggested by B at every stage. And though this is a perfectly good rule of its kind, it does not belong to arithmetic.

Now B's eventual reaction, and the fact that it would be quite justified, particularly if several other individuals were brought into the game and if A always refused to allow their suggested continuations as correct - all this suggests a very important feature of the concept of following a rule'. (pp.29-30)

I shall first make some somewhat critical comments on Winch's example, then consider the use he wishes to make of it.

Much of Winch's example is couched in terms of justification. We are in effect asked to consider when B would be justified in saying that A was not really following a mathematical rule. Winch may seem to move to a non-epistemic conclusion: 'Certainly A was following a rule; but...' Yet we could construe this as having tacitly prefixed 'He would, with perfect justification say that..'

Looking at the details of the example it seems possible, to fill in details in such a way that A could still be following a mathematical rule, in spite of what goes on. The given set-up is that A is asking B to continue a certain series of numbers. We are told nothing about his motives. Suppose that his motive is to fox B, who is always boasting about his mathematical quickness. He thinks up a formula such that presented with the beginning of a number-series

which is in accordance with the formula, B will naturally go in a certain way, but will be wrong; such that when presented with a certain larger stretch B may well be wrong again, and so on, for a long series of numbers. If B knows that A is a good mathematician, that his energy and cunning know few bounds, he may suspect him of being up to some trick but it would seem that he might well feel uneasy at any given point about saying that A cannot be following a mathematical rule. For suppose A has worked out some nefarious formula of the type described...

Note that the way I set up my story of what could be going on in the A and B case, A was guided in his devising of his formula for a series by a rule something like: 'for a long while produce, at intervals, a continuation different from the most natural one'. Even if this were true of A, the formula he developed could surely still be considered a mathematical one, albeit a very complex one. It would make no reference to 'natural continuations', 'B', and suchlike.

As far as justification goes, it seems that once B begins to suspect A he might hit on testing that the rule A was following was not as Winch says by getting A either to write out a formula for the series in some hidden place, or to write out a **long series** of numbers that he holds to be in accordance with it, again in some secret place, where B and the others cannot see it, but where it can be checked on later. Unless B and his friends try getting A to write out the next numbers before they do, or to write out the formula, it seems that they have not tried hard enough to check that A's rule is or is not 'do something different from that suggested by B'. They have not tried hard enough, that is, to feel confident that this is his rule. Again, I do not see that it is immediately clear that Winch is right in claiming that B's reaction ('it isn't a mathematical rule') is specially justified if 'several other individuals were brought into the game and if A always refused to take their suggested continuations as correct'. What if these others are far less capable mathematically than A? Why should their failure here be relevant?

ii 'In principle discoverable.'

Of his example, Winch says that it suggests 'a very important feature' of the concept of following a rule. He then says two things, one of which, he says, is more specific than the other. In fact these seem to be two rather different claims. They are:

- (1) 'one has to take account not only of the actions of the person whose behaviour is in question as a candidate for the category of rule-following, but also the reactions of other people to what he does.'
- (2) 'More specifically, it is only in a situation in which it makes sense to suppose that somebody else could in principle discover the rule which I am following that I can intelligibly be said to follow a rule at all.'

The first claim is indeed rather general. In what way are we to take into account others' reactions in deciding whether or not someone is indeed following a rule? Is it argued - as might appear to be the case - that others' reactions are 'merely' crucial for the discoverability by someone else that a subject is following a rule? Or is it argued, rather, that others' reactions are somehow actually constitutive of someone's following a rule? These would seem to be rather different claims, and the bearing of the former on the latter is not obvious.

Looking to Winch's claim (2) for an answer, we are told that 'more specifically' it only makes sense to say that X follows a rule if it makes sense to suppose that someone else could in principle discover the rule. Thus it appears that it is something about what constitutes following a rule that Winch wishes to argue, not merely something about conditions for the discoverability by someone else that someone is following a rule. The precise claim is that it must make sense to suppose that someone else could in principle discover the rule.

What is meant here by 'could in principle discover'? In his explanation of this Winch draws on the idea of 'going on as a matter of course'. In his

example, Winch notes A's friend B continued the series with 9,11... as a matter of course. He appears to assume that one who is following a rule goes on as he does —

'as a matter of course'. And he claims: 'going on in one way rather than another as a matter of course must not be just a peculiarity of the person whose behaviour claims to be a case of rule-following. This behaviour belongs to that category only if it is possible for someone else to grasp what he is doing, by being brought to the pitch of himself going on in that way as a matter of course'. (p.31)

of
^

This suggests a construal of Winch's example A and B and A's not following a mathematical rule somewhat different from that we have given. It could be that Winch wants us to take it that A's claim is this: given 1 3 5 7 as an initial, given series, '1 3 5 7' is the way he goes on as a matter of course. Then the issue becomes whether or not he is going on this way as a matter of course, and not whether or not from the outset he had in mind, and was trying to act in accordance with, a particular mathematical formula.

Recall my comments on Winch's example. I 'saved' A's mathematical rule-following by having him go on in accordance with an understood formula which he had developed in order to fox B, intending that the formula generated first the numbers 1 3 5 7 but then a series of numbers which would not naturally occur to B as the continuation. Thus A as I conceived him in my 'filling out' of Winch's case was not 'going on as a matter of course' on the basis of the nature of the of the first few numbers. He is not seeing '1 3 5 7' as the natural continuation after an initial '1 3 5 7', but rather as the natural (right) way of developing the series on the basis of the special formula that he had invented to fox B.

As we have argued above, there is good reason for not allowing into a general discussion of following a rule, the assumption that we know from the start what it is to understand a formula or rule-formulation, and then to do what the rule in question says one is to do, or to act in accordance with the formula. Understanding formulae, etc, is the most fundamental phenomenon at issue.

Let us take it, then, in Winch's example, that A is not (and is known not to be) following an explicit rule or formula from the beginning, but purports to be continuing the series as he does as a matter of course on the basis of the numbers given. Winch's claim was, it seems, that given the facts about B and friends (that they do not go on as a matter of course the way A says he does), one would be justified in saying that A was not following a mathematical rule (in the relevant sense). A different claim then took over, no longer just a claim about justification: unless someone else could in principle discover the rule A is following A is not following a rule. This condition can now be restated as follows: unless someone else could be brought to go on as a matter of course just as the putative rule-follower, A, goes on (or says he goes on), then A is not following a rule. What, then, is the precise nature of the condition that someone else could in principle latch on to what the putative rule-follower is doing? Is the condition supposed to be proven unfulfilled in the A and B (and friends) case? Is it then something like this: if a number of people (say with similar background to A, e.g. trained mathematicians) try hard to get the hang of what A is doing, when he is supposedly following a rule in a series of actions, and they fail, then A is not to be following a rule? An immediate question arises: what if B 'latched on' to A's principle, but C, D, and everybody else did not? Would B be enough? Or must everyone (with the relevant background) latch on? Or everyone intelligent enough? Winch is not clear on these points. He states, of the use of 'words like "Everest" and "mountain" ', that given a certain sort of training, everybody does, as a matter of course, continue to use these words in the same way as would everybody else' (31) (my stress). But he states as a necessary condition '...it must be possible for someone else...' (31).

There appears to be one idea of 'in principle discoverability' by someone else, in which whether you say 'someone' or 'everyone' makes no difference: 'someone else' or indeed 'everyone else' are here a sort of rhetorical fiction.

The idea is roughly this: if X is following a rule in what he does then it is possible for a person Y to learn to go on as X does as a matter of course, provided at least that Y has the relevant abilities (perceptiveness, mathematical knowledge, and so on) and is placed in the relevant circumstances. Here 'in principle discoverability by 'another'' may be a trivial consequence of the claim that X is following a rule; it appears that it will be so if and only if 'X is following a rule in what he does' entails 'X is going on the way he does as a matter of course'. On this account someone might claim that a person, Y, could in principle learn to go on as X does even in the case of terms X uses for sensations of his which have no behavioural correlate,¹ and even given the stipulation that among the 'relevant circumstances' here must be counted 'having X's sensations'. Suppose that only X can have X's sensations. Even then if X is going on as a matter of course in what he does, it is (therefore) possible for someone Y (who has X's sensations) to go on as X does as a matter of course. It just so happens that in this case the only person who can in fact do so is X himself. To point this out is, in effect, to point out the extreme thinness of the interpretation of the 'in principle discoverability' condition that we are considering. However it may be with this account, it appears that Winch himself wants to claim that 'in principle discoverability' in some stronger sense is necessary to rule-following. What might this stronger sense be? Here I suggest two possibilities. First, Winch might claim that it is in principle possible for someone else, Y, to learn what X is doing if and only if Y, given his status as

¹ These kinds of terms have generally been assumed to be central to Wittgenstein's 'private language' argument. See for instance, Norman Malcolm's review of the Philosophical Investigations. For an exposition of Wittgenstein's private language argument that regards sensation-language as a mere special case, see Saul Kripke, 'Wittgenstein on Rules and Private Language: an Exposition' (forthcoming).

another human being, could, with the relevant training and background and physical makeup, get the hang of what X is doing, though he may in fact never actually succeed however long he goes on. This would probably rule out a 'pure sensation' language, but if such 'in principle discoverability' of X's rule were construed as an entailment of 'X follows a rule' it would in itself not allow anyone ever to say 'Well, we have tried long and hard...therefore it follows logically that he is not following a rule'.

A stronger sense than the last would be this: it is in principle possible for someone else to latch on to what X is doing if and only if, should actual other people try to get the hang of what the putative rule-follower is doing, they would, by and large, sooner rather than later, given any relevant background training, actually succeed. (Given this condition, again, a language of 'pure sensation' names may well not be possible). From what Winch says, it seems that he might well opt, in fact, for the strongest condition noted.

iii. Social settings and rule-following.

Rather than pursuing this issue here (we shall return to it later),¹ we must look at how Winch brings in a social setting for rule-following. For so far we have only got to the claim that if others-in-general try to get the hang of what X is doing, they must (sooner rather than later) actually succeed. Both of these claims could be true, it seems, without X being a member of any human group, or living in any 'social setting'.

Winch argues from the notion of 'making a mistake'. He writes that if it is possible to say of someone that he is following a rule that means that 'one can ask whether he is doing what he does correctly or not' (p32).²

He then claims that:

'a mistake is a contravention of what is established as correct; as such, it must be recognisable as such a contravention'.(p.32)

Leaving aside the notion of 'establishment' (i.e. being established) for

¹ Section 5.

² Cf. Winch, op.cit., p.58, 'the test of whether a man's actions are the application of a rule is...whether it makes sense to distinguish between a right and a wrong way of doing things in connection with what he does'.

the moment, we have here a claim about 'the recognisability as such' of mistakes. Now it is not clear that a lone 'rule-follower' could not recognise mistakes as such, if he made mistakes. Suppose he has 'discovered' the natural numbers and is playing around with various mathematical series. He is going up by 5s-- 5, 10, 15... at a certain point he goes: 55...65. Then he realises he has skipped a number and starts again at 50: 50, 55, 60, 65. It isn't clear that he couldn't recognise his mistake. Moreover, insofar as mistakes must be recognisable 'in principle' by others, this, as we argued of 'in principle discoverability' - even taken in a strong sense - does not demand an actual social setting for the rule-follower.

What then of the claim that 'a mistake is a contravention of what is established as correct'? Winch does not make entirely clear what he means by 'what is established as correct' or what the argument is for this claim about mistakes. What he says is:

'If I make a mistake in, say, my use of a word, other people must be able to point it out to me. If this is not so then I can do what I like and there is no external check on what I do; that is, nothing is established. Establishing a standard is not an activity which it makes sense to ascribe to an individual in complete isolation from other individuals. For it is contact with other individuals which alone makes possible the external check on one's actions which is inseparable from an established standard'. (p.32)

Now it seems that if I am 'following a rule', and therefore someone else could get the hang of what I'm doing (and hence learn this is how to go on...), then, if I were to make a mistake (to go on wrong), someone else must in principle (having got the hang of what I'm doing) be able to point it out to me. (Similarly if I know how to go on, I should be able to point it out to myself.) But must someone else already know what I am doing (know the rule?) in order for me to be following a rule at all? If not, how, if someone else sees how to go on, can he be establishing anything that was not there before? In cases where the other learns the rule from me, he can only endorse, not establish, what

is right. If he is able to grasp what I am doing must there not already be something to be grasped, be something I am doing?

Winch, after the quoted passage, then qualifies his statement:

'It is of course possible, within a human society as we know it, with its established language and institutions, for an individual to adhere to a private rule of conduct'.(pp.32-33)

So it is not a necessary condition of X's following a rule, according to Winch, that someone else does indeed 'grasp' what is going on in every case of rule-following. But it seems to be implied that in general others must already know the rules we are following.

Winch then sums up what he takes to be Wittgenstein's view and what is also, evidently, his own view) with two points. (1) 'It must be in principle possible for other people to grasp (a person's private)...rule and judge when it is being correctly followed'. (2) 'It makes no sense to suppose anyone capable of establishing a purely personal standard of behaviour if he had never had any experience of human society and its socially established rules'.(p.33)

It should be clear that Winch has really not presented an argument for the crucial (2). Winch, then, sets up an intriguing problem about language and rules based on Wittgenstein's work, but gives an obscure two pronged 'solution'. He claims first, that others must be able 'in principle' to latch on to what a rule-follower is doing; and second, that a rule-follower must have had experience of a human society with its socially established rules. The second claim alone involves an actual requirement of a social setting. As we recall, meaningful actions were said to presuppose social settings because rules did. But so far one can hardly count it proven that rules, and rule-grasping and following, do presuppose social settings. In order to examine the bearing of Winch's strongest claim, if true, on the nature of meaningful actions, I now try to give some more specific content to that claim. In view of the brevity of Winch's own remarks, what follows will be quite speculative with regard to his intentions.

iv. Established rules and rule-following.

1. A simple account

What might Winch have in mind when he writes of 'established rules' and 'establishing a standard'? Let us first consider one simple and minimal notion which uses concepts Winch deploys. This account runs thus:

A rule is established in a society (or social group) S if and only if

- (a) most members of S generally agree in their behaviour in a certain respect;
- (b) most members of S go on as a matter of course in this behaviour. (An example: most members of S use the word 'red' (as a matter of course) to refer to the colour red).

Now, this is a rather thin notion. For one thing, though this account mentions 'society S' what the society's members have to do for there to be a rule 'established' in their society might be done by a set of totally unconnected, socially unrelated people (say the intergalactic 'population' considered in the last chapter). More generally, calling this an account of a socially established rule might seem rather inappropriate.

Probably Winch has a richer notion in view when he speaks of socially established rules. Perhaps, for instance, he thinks of the rule-conformity in question as resulting from interaction between people. Let us consider how this might come about, to develop a richer account of '(socially) established rules' that is relevant to Winch's concerns.

2. A 'rational reconstruction' of rule-establishing.

Remember that we are ultimately interested in putative rules of language. Perhaps we can 'reconstruct' the 'social establishment' of such rules somewhat as follows. (This will be very sketchy).

One person in a small gathering, call him A, points to a couple of flowers in turn, saying, 'snowdrop!' each time he points. Some others, B, and C, feel they see how the word in question is being used. They catch A's attention and point to certain flowers saying 'snowdrop!' as they do so. A nods agreement.

Then perhaps someone, D, points to what we would call a 'crocus', and says 'snowdrop'. A shakes his head in disagreement, but he looks to B and C for confirmation; they also shake their heads confirming his judgement (see below on 'pointing', 'indicating disagreement', and so on). D may look puzzled. He points to another object and queries 'snowdrop?'. This time the others nod agreement, and there are smiles all round. Now someone may go off on an orgy of picking out 'snowdrops'. The others, perhaps, are more sedate: they use the word when they need to.

If anything like this ever happened, we might naturally say that a rule for the use of the word 'snowdrop' had been established among the people concerned. Note that for this account to make sense, whatever indicates 'this is wrong', say a certain gesture with the head, must itself be 'taken' in the right way by the parties concerned. Thus this account of 'establishing a certain rule' appears to rest on the assumption that certain 'rules' or modes of interpretation are already in some sense 'established', are 'pre-established' if you like.¹ In order for the kind of 'teaching' process referred to to get under way, there must, it seems, be some 'primitive' understandings which underlie its development. In order to teach a new way of doing something we must already share some ways of seeing/doing things with our pupils. Perhaps sometimes some crucial understandings have to be 'negotiated' but there can surely be no negotiation without some agreement somewhere. This account also relies on a gesture of pointing as (here) primitive, and also a gesture or utterance of questioning.²

¹ They would in fact probably count as established on the simple account first considered.

² Cp. Philosophical Investigations, §185 where Wittgenstein describes someone who 'naturally reacted to the gesture of pointing with the hand by looking in the direction of the line from finger-tip to wrist, not from wrist to finger-tip'. Any account of a rule's becoming established in the way envisaged appears to presuppose a homogeneity in natural reactions of some kind - or some homogeneity of 'pre-established rules'.

A further point to stress here is this. Winch argues against A.J.Ayer who claimed that 'some human being must have been the first to use a symbol';¹ Winch takes Ayer to be wishing to imply that socially established rules cannot have been presupposed by this use. Winch objects that: 'from the fact that there must have been a transition from a state of affairs where there was no language to a state of affairs in which there was language, it by no means follows that there must have been some individual who was the first to use language'. (pp.36-7)

My 'rational reconstruction' above was not intended to imply or presuppose that there must have been some individual who was the first to use language, or to follow a rule of language. I do not have to claim that the first user of 'snowdrop' in our story was clearly following a rule or had the grasp of a rule for 'snowdrop' or was already using 'snowdrop' with its meaning, before he had the grasp of a socially established rule for the use of 'snowdrop', that is, before he had reached an agreement with the others in their use of 'snowdrop'.

In our story he himself was unsure about one case, initially, and turned to others for confirmation; thus perhaps he too needed some confirming interactions with others before going on as a matter of course in his use of the term 'snowdrop'.

We can now perceive a kind of sense for 'established rule' where the qualifier 'socially' established has some force. Let us produce a rough partial account so as to have something to work with in this chapter:

A rule is socially established in a society S only if:

- (a) most members of S generally agree on their behaviour in a certain respect;
- (b) most members of S go on as a matter of course in this way,
- (c) (a) and (b) are the result of interactions between members in which

¹ See Winch, p.36; The quotation from Ayer is from 'Can there be a private language?'.
.

a common understanding¹ about what should be done, or may be done, has been arrived at.

3. Socially established rules and rule-following.

One can, Winch claims, adhere to a 'private rule' of conduct only on condition that one has had experience of human society and its socially established rules. For the sake of argument here, let us say that a human being adheres to a private or unshared rule only if:

- (1) There is a certain way he has of proceeding; such that
- (2) He goes on as a matter of course in so proceeding, and
- (3) at least given some further experience, and training, and the same degree of intelligence, other human beings could (sooner rather than later) learn what he is doing so as to get to proceed as he does, as a matter of course, but
- (4) no other human has yet learnt from him what he is doing, by interacting with him.

To give a concrete example: One uses 'snowdrop' to refer to (what we also call) snowdrops. This is a private name one has given them.

We have elucidated Winch's strongest claim about rules more precisely than before, if risking manifold errors of interpretation: one can only adhere to an unshared rule (as so far defined) if one has experience of a situation of socially established rules (as so far defined). We note that Winch does not say that any particular rules or types of rules must have been established, only that some must have, and one must have experienced this. Having got at least this far - but not

¹ On how the nature of such 'common understandings' may be construed, see the next Chapter.

seen the dependency of unshared rules on socially established rules made out, let us consider what its consequences would be for the content of any meaningful action. In particular: how, if at all, will meaningful actions, as such, be social?

v. The content of meaningful behaviour.

At this point we can recapitulate the arguments from meaningful actions to 'social settings' sketched at the outset of this discussion (in Section 2 (ii.) of this chapter).

The core of one was this: acting meaningfully involves one in applying a rule (via what might be called a linguistic and/or a conceptual 'episode'); rule-application 'presupposed' a 'social setting'. So acting meaningfully must, to the extent that it involves rule-following, presuppose a social setting also.

The core of two other (very similar) arguments was that acting meaningfully involves having a grasp of, as opposed to actually applying, a rule

We have now established a somewhat precise account of what is meant by 'rules presuppose a social setting'. That is: you cannot (as a matter of logic) follow or even grasp, an 'unshared rule' without experience of 'socially established' rules; We have given interpretations of the terms in (scare) quotes. Now let us consider what bearing the truth of 'rules presuppose a social setting' would have on the 'socialness' of actions.

One can, apparently, adhere to an unshared rule, and since no specifications about which rules are (or must be) socially established have been given it seems that we may assume that one who acts meaningfully could have involved in any given action of his only unshared rules. But perhaps this is not so. Perhaps at any rate an argument can be given to refute the assumption. So I shall consider two kinds of case, one where someone's meaningful action at least prima facie involves only the application of rules which weren't established in his present or past community, and one where established rules are involved.

1) Eosurc divides the exotic plants on his island into two kinds: the Yfaels (as he calls them) and the non-yfaels. Other people could learn this distinction easily enough, but no one else in fact has a grasp of it yet. He has discovered two special ways of eating plants, which he calls 'chomping' and 'champing'. At this moment he is going through the motions of chomping ~~an~~ yfael, and means to be so doing. Let us take the strong assumption that he is therefore applying the concepts of yfael and chomping at this moment: he is seeing his behaviour in this light. Where, if anywhere, lies the 'social' character of his action?

According to the Winchian line, Eosurc must have had experience of a society with its socially established rules in order to adhere to any rule at all. So by this argument a necessary condition of Eosurc's applying the concepts of yfael and chomping is his having had such experience. To have had experience of ('a society with its) socially established rules' may be held to involve at least (a) having the concept of a socially established rule and (b) having once realised that one was in the presence of (for example, when being taught) such a rule. Let us take it that these are the requirements on Eosurc, and that a 'socially established rule' is as we said above.

Must Eosurc, in intentionally chomping ~~an~~ yfael, have socially established rules in mind in any way? Or can his action be seen as social for some reason other than what he has in mind? A positive answer to the first question here hardly stands out from what we have so far found in Winch. Clearly Eosurc may have socially established rules in mind here: he might be chomping ~~an~~ yfael in order to prove to himself that he could do something that none of the socially established rules he has had experience of ever envisaged. Then the notion of a 'socially established rule' itself would be essential to characterise the meaning of his action. His action would then, indeed, be social in Weber's sense. However, this possibility proves nothing about the general case. In the next section, we

shall consider a fairly subtle argument in relation to Winch's weaker claim about the 'in principle discoverability' of rules. If the general form of this argument is indeed correct, and if Winch's stronger claim about the link between unshared rules and socially established rules could be supported somehow, then - as will be clear - this argument would in fact also lead to the conclusion that, in chopping on yfael, Eosurc did indeed have others in mind, in a sense, in acting thus. We may here point out a rather simple way in which Eosurc's action could be argued to be social, in a sense.

Suppose we were to call all, and only, people who have (or have had) experience of socially established rules 'social beings'. (This is a not inapposite sense for the phrase). Then we might define a 'social action' as the action of a social being. Then so far as Winch's argument goes, if we could arrive at a satisfactory way of showing why one could not follow a private rule if one had no experience of socially established rules, we would have come at least quite close to showing that all meaningful actions whatsoever were social actions in the sense just defined.

2) What of the case of someone applying socially established rules as he acts? Consider Weber's woodcutter, chopping wood in order to get in his day's exercise. Assume he learnt the notion of chopping wood, exercise, and so on, from his parents and his teachers. It appears to be no more necessary for him than for the unshared-rule follower, to have others in mind when he acts. Albeit these rules are established, why need he think of their establishment? Prima facie he just sees what he is doing in a certain light, that is, as 'chopping wood (for exercise)'. Insofar as to follow a rule at all he too must have had experience of society and its established rules, he too must, we may assume, have the notion of an established rule and have applied it on occasion. But for him as for Eosurc the only plain conclusion about an action's 'socialness' from what Winch has said (or we have so far got from him) seems to be that in a certain sense of 'social action' (defined via a special sense of 'social actor') all his actions are social in this sense.

We have now been through Winch's explicit argument. There seems to be a glaring lacuna in it. It is simply not clear why it was claimed that one without experience of socially established rules could not follow a rule at all. And if this is not substantiated, nor are the conclusions about the socialness of actions which seem to be derivable from it. I believe that more can be said on behalf of the socialness of all action that Winch has said, however. In what follows I shall no longer concentrate on the movements of Winch's discussion, but what follows may possibly still represent something that lies behind his explicit remarks.

5. Further considerations on the social nature of action.

i. An argument for the social nature of all action.

Let us hark back, now, to Winch's claim that a necessary condition on rule-following was the 'graspability' in principle of the putative rule follower's way of going on, graspability, that is, by people other than the rule-follower. There is a possible line of argument which using this, the weaker of Winch's positive claims, leads to the claim that all actions are social, given certain not too implausible assumptions. It is worth isolating this line and seeing what it does and does not show about the nature of all meaningful actions. Consider then, the following, as an hypothesized partial analysis of the concept of 'following a rule':¹

A person follows a rule in what he does if he (a) goes on in a way other humans do or would latch on to given a certain amount of time, trouble, (and perhaps training).

Following a rule without mistake might be construed as above plus:

¹ Winch uses on the whole the language of analysis and logical relations between concepts. Wittgenstein's own positive views on rule-following are couched in a different manner. Cf. a passage Winch himself quotes from Wittgenstein: 'Here perhaps one really would say: "...it is not a rule' (Wittgenstein, op.cit. 6.237, p. 87e; Winch, p. 32), (my stress). For an appropriate construal of Wittgenstein, see S. Kripke, op.cit., (forthcoming). The argument I am about to give does not depend on differentiating between these construals.

(b) others who had latched on to what he was doing would agree that each step he took was a correct one.

Note that nothing has been said yet about established rules in the sense discussed above or in any other sense. Nor does the argument I shall give refer to them. We need not move from 'discoverable' rules to find fairly tight argument for the 'socialness' of all meaningful action. The argument briefly and roughly is this: the concept of following a rule is explicated as above; 'applying a rule' is explicable in a similar way. Now assume that in employing any concept, in particular in identifying anything as a ' ψ ' rather than an ' ϕ ', one is applying a rule. When I think 'that's a ϕ ' I may of course not consciously think 'if I am right in so saying, I have grasped a certain rule and am applying it correctly'. But we could say that a belief to the effect that I have grasped a certain rule and am applying it correctly is presupposed by any such thought. (A homely dialogue might illustrate this, Juliet: 'Romeo, you are an M.C.P.' Romeo: 'You're crazy, you have no idea what an M.C.P. is'. Juliet: 'Of course I do, and I'm right in saying you are one. Do you think I would have said you were an M.C.P. if I hadn't known what one was?').

If a belief that I have grasped a certain rule is presupposed by any rule-application (including in particular those involved in intending to ϕ) then, it could be argued, the belief that something like the stated conditions for rule-following obtain is presupposed.¹

It looks, then, as if some sort of socialness in 'every' meaningful action may now be argued for. What do I have, in some sense, in mind when I act? What thoughts, beliefs, intentions, motives, underlie my action? If the belief that, among other things, I am doing something (for instance, and in particular, making a certain identification) that other people would agree was right and correct,

¹ The 'paradox of analysis' could pose problems here, but I shall ignore this, presenting the argument as if (and in the hope that) it does not.

then is not the meaningful content of my act sufficient to make it social?

This line of thought, it seems to me, gets close to a conclusion that all action or all meaningful action is 'social' in a sense, as a result of the thoughts an agent must have in acting. Of course the less actual rule-application we allow in any meaningful act the more force this argumentation loses. But it is not difficult to accept a view of meaningful action whereby some concept at least is always applied when such an action is performed, such that the act's meaning varies with the concept applied.

ii. Are all actions social in Weber's sense?

1. The consequences of the argument in 1.

This argument does not lead to the conclusion that all action is social in Weber's sense. For we find that the belief at issue here, such as it is, does not appear to be a necessary part of the specification of (Weberian) meaning of the action. It lies, if you like, below the level of meaning. For given that we would not have motives or intentions at all did we not grasp and apply rules for the use of words and/or concepts and given that applying such a rule presupposes a belief about others, the belief about others is neither a motive for action, or an intention in acting, or a motive for a motive ... and so on. That is, it is not necessarily a motive. Our imaginary character Eosurc might, certain that he follows a rule, confidently invite Man Thursday to note his referring to a certain plant as an 'yfael'. There, in the teaching situation, the availability of one's practice to others plays a role in one's motivation. But we do not always teach, nor are we always learning our conceptual tools. In other situations, we can afford to be less 'altruistic'; and, surely, often are.

I would conclude, then, that this argument does not show all action to be Weberian social. Though I should be happy to admit, if the argument goes through, a belief in certain claims about others is shown to underlie any meaningful act.

2. A 'Strawsonian' argument from facts about self-reference.

I want briefly here to consider a somewhat different line of argument for the Weberian socialness of all actions. The question at issue here is: am 'I' always referred to, in a description of the meaning of my actions? (e.g. if I intend to climb a tree, to bell the cat, to see Fred, do I ipso facto intend that I see Fred etc? What about intending that the door be opened? Mustn't even that be in effect intending that I bring it about that the door be opened?) It seems not implausible to suppose that this is so. But if it is, then someone might try to construct an argument, à la Strawson, in Individuals, about the 'awareness of others' being somehow a component of every meaningful action. Strawson argued that having the notion of myself as a locus of intentions, sensations and so on involved having the notion of other centres of experience and intentions who are not myself...; the concept of the 'I' presupposes, or carries with it, the concept of the 'not-I'. So, accepting Strawson, and accepting the presence of 'I' in the description of the meaning of actions, it follows that any performance of an action requires that the agent have the notion of some agent or conscious being other than himself. Even if this is so, it fairly clearly does not entail that all actions are Weberian social. The point about this 'Strawsonian' line, is similar to that about the view that to act meaningfully is to follow a rule is to have a belief about other human beings. Even if these things are so, they have no evident entailments for the meaningful content of particular actions of the agents concerned.

In conclusion, I assume that the claim that all meaningful actions are Weberian social, if not definitively quashed (how could it be?) is by no means proven. We have considered at least two lines of argument which though they strain to, fail to reach the desired result.

iii. Conclusions: two further concepts of 'social action'.

In this chapter, two main implicit notions of 'social action' have emerged, and are possible candidates for the title. One was the notion deriving from a definition of a 'social actor' as one with experience of a society with its established rules. All the meaningful actions of such an actor are to be called 'social', on this view. (Insofar as all agents had to be social actors, all actions

then, would have to be social also). The second notion, implicit in the argument just considered, is this: a social action is a meaningful action which presupposes the belief that others (other human beings?) would have certain reactions...(On this account, insofar as the relevant argument goes through, all meaningful action would be 'social'). Neither of these notions is the notion of Weberian-socialness nor do they imply Weberian-socialness.

Looking at these uses of 'social action' for their own sake it seems that neither is a hugely 'counterintuitive' use. Whether we should adopt them depends, evidently, on our purposes. We have as ^{an} outstanding role ~~such~~ a concept might perform, that of contributing usefully to collectivity-concept analysis, ^{by providing illuminating necessary} or sufficient conditions for a collectivity in terms of a type of action. Neither of these two notions seems to do that trick.

Finally: suppose all action is indeed social in one of these senses, does that mean that any study of any kind or form of action, including suicide, say, or grinding one's teeth, is therefore apt to be called a social science? Against this is first, the possibility that the (omnipresent) social aspect of actions may not be the focus of the study. But assume we ignore this reply for a moment. The claim might be: if it is actions that are studied, then since actions are social, something social is studied (or used in explanations). I myself feel inclined to say something like this here: the intrinsic socialness of these phenomena is so weak, given that it is there at all, that it is jarring to call a science a social science by virtue of its dealing with phenomena of this kind. So even if 'action' or 'meaningful behaviour' is a natural kind of thing, (as it surely is), and is social in a sense, a certain strength of socialness is lacking. This is true particularly of the argument which seemed to beg fewest questions, the one considered at the beginning of this section.

As for the argument, alluding to established rules and 'social actors', the socialness that would be involved in all actions, should the argument go through, is, intuitively, of a 'borrowed' kind - borrowed from the socialness of

societies and such phenomena as socially established rules. However, Winch gives no proper account of 'socially established rules' or other features of societies. We next turn to some ideas of Durkheim, who put features of societies clearly at the centre of the social scientific stage.

CHAPTER FOUR

A DURKHEIMIAN CONCEPTION OF A SOCIAL PHENOMENON.

1. Introduction

i Contents of this chapter.

We have examined certain claims about social science's subject matter, and found them wanting. These claims were of the form, roughly: social sciences are sciences of 'social action'. In spite of the inadequacy of these accounts, someone might feel bound to cling to them or to something of the same form. For he may believe that the alternative is to 'reify' things which are not, to be led into dubious statements about 'group minds' and the like, and to say that social sciences are about ^{such} nonexistent^s. Some such criticism has often been levelled at Emile Durkheim for his claim that sociology deals with certain 'sui generis' entities, which he calls 'faits sociaux' or as we shall say 'social facts'.¹

1 Use of the phrase social 'fact' is somewhat jarring and sometimes I shall use as an alternative 'social phenomenon'-Durkheim himself refers to his 'faits sociaux' as 'phenomènes' and surely uses 'faits' and 'phenomènes' as virtual synonyms in this connection. I intend that 'social fact' be so interpreted where it occurs in my discussion. Cf. S. Lukes, Emile Durkheim, p.9.

The work involved in the present chapter was undertaken in the (apparently unusual) belief that a constructive approach to Durkheim's notion of a social fact would pay off, particularly in the context of our overall enquiry concerning the demarcation of the realm of social sciences.¹ It is true that much of Durkheim's own writing on this subject is obscure, and many of his phrases smack of unacceptable reification. My aim in this chapter is to present a notion of a social phenomenon which is recognizably Durkheimian but free of these faults. I find my neo-Durkheimian notion not only unobjectionable but also appealing. More appealing, certainly, than the accounts of the subjects of social science discussed so far.

I shall not claim that my notion is Durkheim's. Rather, starting from certain key remarks in Durkheim's discussion of 'social facts', I shall develop a notion which seems to capture quite closely the spirit of those remarks, and also, in another sense, to capture an important range of phenomena, phenomena very aptly called 'social'.

I shall contend that a science of the phenomena covered by my Durkheimian notion would clearly be a social science, as would a science of some species of them. Indeed these would be social sciences par excellence. Any global theory of social science subject matters, therefore, must somehow admit these phenomena within its purview.

1 Though Durkheim's presentation of his concept has been the subject of a great deal of (often useful) criticism, attempts at any sort of reconstruction of his notion are virtually non-existent. Notable exceptions are Fauconnet and Mauss (1901), and Lacombe (1926). For a clear summary of some main criticisms see Steven Lukes, Emile Durkheim.

The clear socialness of Durkheimian social phenomena will confirm our view that social action notions are going to be in difficulties as far as providing adequate global theories goes.

Presentation of my Durkheimian notion will take some while: it is quite a complex notion; some of its elements are important and interesting in themselves, some require defense as opposed to certain alternatives which someone might propose.

ii Durkheim's project

Durkheim's project in introducing and discussing the notion of a 'social fact' was very relevant to our own. He was trying to demarcate the realm of sociology (which in his view was, in effect, to demarcate the realm of 'social science' as a whole).¹ He wished to do this by reference to one clearly specified kind of phenomenon. Thus Durkheim:

'...we must accurately distinguish social facts and show what it is that gives them their identity, if we are to avoid reducing sociology to nothing but a conventional label applied to an incoherent collection of disparate disciplines.'²

1 Cf. 'The various disciplines dealing separately with different forms of social phenomena did indeed prepare the way for social science... but social science, in the strict sense, came into being only when it was clearly perceived that the branches... (political economics, the study of morality, religion, law)... were...parts of a whole...embracing all social phenomena.' (from Montesquieu et Rousseau, tr.Giddens, in Giddens ed.Emile Durkheim.p.53) and see the next quotation.

2 Giddens tr. from 'La sociologia ed il suo dominio scientifico'.

So far his project is not so obviously distinct from Weber's implicit project in defining 'social action', but, importantly, Durkheim was not working within the constraint Weber imposed on himself, that of giving an account of social science in terms of a kind of action performed by individual persons. Part of Durkheim's motivation, rather, entailed that he would shy away from any such constraint, espousing rather a conflicting one. For he wished to show that 'social' phenomena were distinct from 'psychological' and 'biological' ones, and hence that there was clearly room for a science of the social-sociology distinct from the already existing sciences of psychology and biology.¹

Now, 'psychology' means different things to different people, but evidently on many interpretations, if the subject matter of a science were the actions of individuals, this science would be that of (individual) psychology.²

1 See(1) 'a discipline may be called a 'science' only if it has a definite field to explore...before social science could be deemed to exist, it first of all has to be assigned a definite subject matter' Montesquieu et Rousseau tr.Giddens, in Giddens, ed.p.57.

(2)'...There is in every society a certain group of phenomena which may be differentiated from those studied by the other natural sciences'. The Rules of Sociological Method, p.1 (references to this work are to the English translation by S.A.Solovay and J.H.Mueller, except where otherwise stated). (I have sometimes preferred to use my own translation).

2 Cf. Durkheim 'psychology...is properly the science of the mind of the individual' Rules p.xlix.

Durkheim does not choose a type of human action as the type of social phenomenon. As we shall see, his basic type is rather a kind of 'collective' practice. Among the things such practices involve essentially are a multiplicity of agents.¹ (For this reason we might dub Weber's a 'micro-theory' and Durkheim's a 'macro-theory' of social science's focus for the simple reason that the number of humans involved is necessarily larger in the latter case). Durkheim's type of attempt must, then, have a particular interest for us in the context of our dissatisfaction with certain 'social action' notions.

Quite apart from the main thrust of our enquiry, the construction of a concept of a social phenomenon which is not restricted to actions is of some independent interest. Both in and outside social science related literature, the phrase 'social phenomenon' is quite often used, but rarely defined. In discussions of the philosophy of language, for instance, we come across the claim that language is a social phenomenon. This might itself be a helpful claim if a special notion of 'social phenomenon' were well understood. As I have already indicated, the basic 'overarching' non-technical concept of the 'social' is so wide that it can hardly be used to make a substantive theoretical point here. For this reason more specific notions need to be explored.²

¹ See... 'in order that there may be a social fact, several individuals, at the very least, must have contributed their actions'. Rules, p.lvi

² The linguist F.de Saussure spends some time on an account of language as a social phenomenon, in his Course in General Linguistics. His interesting remarks are very similar to some of Durkheim's, they share some of the obscurity of the latter.

2. Durkheim's basic conception of a social fact

i. The basic conception versus the 'definitions'.

It is common and natural to focus an account of Durkheim's concept of a social fact on his 'definitions' of 'fait social' in the Rules. Nevertheless I shall work here not so much with these as with what I shall call Durkheim's basic conception of a social fact, which is indicated in less central positions in his text. I should first justify my claim that Durkheim has such a conception; this is not part of any standard view of his ideas, and to my knowledge has not been noted by other writers.

At the end of Chapter I of the Rules, the locus classicus for his account of social facts, Durkheim presents the reader with two alternative ways of 'defining' 'social fact'. These run as follows:

'A social fact is every way of acting, fixed or not, capable of exercising on the individual an external constraint; or again, every way of acting which is general throughout a given society, while at the same time existing in its own right independent of its individual manifestations'(p.13).

These definitions have been much criticised in the literature, on grounds of ambiguity, unclarity, even downright unintelligibility.¹ In particular, Durkheim's reference to 'constraint' was evidently intended to cover a multitude of rather distinct phenomena, not all of which characterized all social facts. Indeed, in an apparent attempt to save constraint of some kind as a feature of all social facts (or perhaps in an attempt to show that some one kind of constraint characterized all social facts) Durkheim was reduced at one point to equate the 'constraint' of social facts with their reality: 'they are things which exist in their own right'.²

¹ See especially J.Monnerot, R.Lacombe, Georges Sorel, E.Benoit-Smullyan, and, for a summary of critical parts,-S.Lukes, Emile Durkheim.

² 'These are things which exist in their own right'.Rules, p.xxii of French ed, my translation; see p.lvi, English ed.

In the preface to the second edition of the Rules, in reply to some critics of his 'definitions', Durkheim indicates that these 'definitions' were intended to present some salient features of social facts, features by which a sociologist would be able to pick them out (see pp.xx-xxi,Fr; pp. liii-lv,English). And in the main text of the Rules we find such remarks as: 'A social fact is to be recognised by the power of external coercion which it exercises or is capable of ^{exercising} over individuals' (my stress) (p.10). In 1898- after the first edition of the Rules and before the second-Durkheim wrote 'we wished only to point out a convenient sign by which the sociologist can recognise the facts falling within his field'.¹ Durkheim's 'definitions', then, are not best understood as definitions strictly speaking, that is as explaining or partially explaining the meaning of the 'defined' term. Moreover, Durkheim refers to them, even qua accounts of salient features, as only 'initial' and 'preliminary'.

Summing up the role of his definitions, Durkheim writes:

'It was a matter of restricting the field of research as closely as possible, not of subsuming it under a sort of all-embracing intuitive principle' (p.liii)

It is obviously somewhat dangerous for anyone else to try to delimit the 'intuitive principle' or principles (the French has simply 'intuition' (p.xx)) with which Durkheim himself was operating. However, I believe that one can articulate at least one such principle or intuition, and I shall take this as my own basis of operations. I call this principle Durkheim's 'basic conception of a social fact'. In order to justify the attribution to Durkheim, let me first note some quotations. Here is one from the main text of the Rules:

¹ 'Individual and Collective Representations' p.25; in Sociology and Philosophy, tr.Pocock.

'Here, then, is a category of facts with very distinctive characteristics: they consist of ways of acting, thinking, and feeling... They constitute... a new species, and the characterisation 'social' should be given to them and reserved for them. It is an appropriate characterisation for them; for it is clear, that, not having the individual as their substrate ('substrat') they can have no other substrate but society, whether it be the political society in its totality, or one of the partial groups it encloses, religious orders, schools of politics, literature, professional corporations, etc.' (my translation of p.5, French ed; see p.3 English ed.)

We may note the form of Durkheim's claim here. In the underlined passage he seems to be saying, roughly, that a truly social phenomenon is one whose 'substrate' is either a whole society or a lesser social group within a society. More precisely, he argues thus: (1) ways of acting whose 'substrate' is a social group are aptly called 'social'; (2) they alone are aptly called 'social' since they alone do not already fall into a category other than that of the social. One contrast case Durkheim probably has in mind is that of certain actions by individuals which may seem to have social aspects of various kinds, but nonetheless by virtue of their nature as actions by individuals can be seen to fall into the category of 'psychological phenomena'. Actions which are 'social' in Weber's sense would presumably fall into this class.

Rather than at once querying the precise sense of 'substrate' in the above passage, let us turn here to a later passage (in the second edition preface) which ties 'social phenomena' to social groups in two specific ways.

'If as is agreed, this sui generis synthesis which constitutes every society gives rise to new phenomena, different from those which occur in isolated minds, it must be admitted that these specific facts reside in the very society which produces them, and not in its parts, that is to say, in its members'. ¹

Here, then, Durkheim refers to two relations that phenomena may bear to a society or social group, what we may call the relation of production and the relation of inherence (or 'residence in'). It seems that Durkheim claims for his 'social facts' that they both inhere in and are produced by a society or social group.

Is either of these relations prior in Durkheim's mind, can one single one be argued to be definitive of the social fact for him?. One might, on the basis of the first quotation, feel that it is the relation 'having a society or social group as a substrate' which is definitive of the social fact. What, however, does Durkheim mean by 'substrat' in this passage? Solovay and Mueller, his translators, appear to wish to translate 'substrat' as 'source'. (What Durkheim wrote was: 'il est clair que, n'ayant pas l'individu pour substrat, ils ne peuvent en avoir d'autre que la société .') Solovay and Mueller translate: 'it is clear that since their source is not in the individual, their substratum can be no other than society') If 'substrat' meant 'source' this might indicate that the relation of production was prior for Durkheim. However, my own reading of the three paragraphs surrounding the second quotation I have just given, which use the word 'substrat' several times, would have 'substrat' be the correlative of the relation of inherence, and (in the second quotation) Durkheim appears to want to contrast the notions of 'inherence in' and 'production by'. Such a reading of 'substrat' in the first quotation could make the relation of inherence - whatever it is exactly - seem to have been prior in Durkheim's mind.

Now, I believe that for Durkheim inherence in a society or social group was, intuitively, a necessary condition for a social fact, if not a sufficient one. Production by a society on the contrary will certainly not be a sufficient condition, for a given action of a given individual might count as produced by a society under some conditions, but is not of the right type to be a Durkheimian social fact. However, Durkheim did in fact lay stress on both aspects of his social facts, in discussing their distinctive and specifically social character: he has stressed both their production by, and their 'inherence in', social groups. Moreover, I believe that consideration of a twofold account leads us to an important and intuitively highly social range of phenomena. On the basis of my reading of Durkheim, then, I am going to take the following

as a characterisation of Durkheim's 'intuitive definition' or basic conception of a social fact:

A phenomenon P is a social fact if and only if it inheres
in and is produced by a society or social group.

This is, obviously, a somewhat obscure statement. I shall make an attempt to clarify it further shortly. One point may, however, be stressed at once: the basic conception makes the notion of a society or social group central to the idea of a social fact, thus we might say-though somewhat misleadingly-that for Durkheim a social fact is a kind of 'societal' fact. This is somewhat misleading because a social group need not be a society, and societies as such are not so clearly essential for Durkheim's social facts as are social groups of some kind. In what follows I shall, in fact, write 'social group' as short for 'society or social group', in discussing Durkheim's notion.

There is a clear difference between Durkheim and Weber here - for the concept of a society or social group was in no way involved in Weber's concept of the fundamental social phenomenon - the 'social action' in Weber's sense.¹

ii. Towards an example.

Can a real and important kind of phenomenon which clearly corresponds to Durkheim's basic conception of a social fact be described in precise detail? I believe that it can; the rest of this chapter will be largely taken up with the reasoned construction of such a description. On the basis of this we shall present a relatively precise neo-Durkheimian concept of a social phenomenon.

1. Our problem, then, is to find, and clearly describe, a group of phenomena which 'inhere' or 'reside' in social groups and which can reasonably be said to be produced by such groups. There is one obvious problem in this statement of the problem. What is the relation of inherence, what is it for something to 'inhere' in a social group? We must find an answer to this question, so we can proceed.

¹ I shall later argue, in effect, that there is no need to veer towards Weber and others on 'individualist' grounds; see pp 298-302.

I have already suggested that 'inherence in' and 'substrate' are correlative notions for Durkheim, but what are these ideas? As with other terms Durkheim employs, there may be ambiguities and confusions in his usage, and terms like 'substrate' seem prone to obscure a mix of logical and ontological categories.¹ Nonetheless a fairly simple idea can be extracted from some things Durkheim says, an idea which also seems to have some point.

In an important passage in the second edition preface² Durkheim argues as follows (as I see it):

(1) on occasion certain elements or substances combine and the resulting mixture or combination possesses new properties, i.e. properties previously lacked by the individual uncombined elements. For instance, two flexible and soft metals are mixed and the resulting mixture is hard.³

(2) it is freely acknowledged that the proper conception of such a situation is that the new properties in question 'reside in' and the combination of the elements, and not in the individual elements which are combined.

Now, the phrase 'reside in' here may seem to mean simply 'is a property of'. This is the simplest interpretation of the phrase. Against this interpretation, someone might argue that if it did mean this, Durkheim's 'argument' would be truistic and pointless. It would amount to nothing more than 'the properties of a combination of elements...are properties of the combination of elements'. Against this objection, I think it can be shown

¹ Cf. Jonathan Bennett's criticisms of Locke on 'substance' in his 'Substance, Reality, and Primary Qualities'.

² p.xvi, Fr.ppxlvii-xlvi, Eng.

³ Cf. 'The hardness of bronze is not in the copper, the tin, or the lead, which are its ingredients (qui ont servi a le former) and which are soft and malleable bodies; it is in their mixture (elle est dans leur mélange)' Rules, p.xlviii, English; p.xvi, French.

that Durkheim could have a useful point here. To help see this, let me distinguish between two senses of phrases like 'the combined elements'. Someone saying 'P is a property of the combined elements e and f' could mean one of two distinct things:

(a) 'P is a property of the individual elements e and f (which are in fact combined)'; or (b) 'P is a property of the combination (which is in fact made up of elements e and f)'. What Durkheim is saying in the above argument need not be truistic, then, if 'resides in' means 'is a property of'. He may mean to stress that properties of a combination of elements are - or at least may be - purely properties of the elements in combination or the combined elements (sense(b) above), and not of the combined elements (sense(a) above).

In general Durkheim wants to argue at least that just because a thing A is wholly composed of elements e and f it does not follow that A cannot itself have properties that e and f lack, not that 'A's properties' are somehow really rather properties of its elements (and he is surely right in so arguing). With regard to society, Durkheim wants to argue that, by parity of reasoning, we should be able to distinguish between, on the one hand, properties of the individuals out of which a social group is composed, and, on the other hand, properties of the group itself or 'members of society as they are socially combined'.

These claims may seem obvious, but I believe that the distinction about society that Durkheim is making is in danger of being obscured in a passage in a book on Durkheim himself. I have in mind a passage in Steven Lukes' 'Emile Durkheim' (p.7) where Lukes first uses the phrase 'the substratum of associated individuals' and later continues with 'its constituent elements or 'substratum'. Here 'associated individuals (elements)' is apparently being equated with 'constituent elements', but for Durkheim the social substratum is rigorously conceived of, as Lukes's first phrase indicates, in terms of a set of individuals in combination or taken collectively.

It is perhaps because he saw the possibility of the distinction mentioned that Durkheim felt able not merely to maintain that 'society' could be 'and was) something distinct from any one individual, but also that social phenomena can have as their 'substratum' 'not the sum of...individuals taken one by one, but the collectivity of those individuals'. This latter phrase in quotation marks, is a translation from a criticism of Durkheim by Lucien Herr.¹

Now, is 'inheres in' equivalent to 'is a property of' or not? In other words, are all of a thing's properties inherent? We recall that Durkheim, in the passage discussed, talks of 'new' properties of combinations, properties lacked by the elements. Is an inherent property of a thing perhaps one that its elements lack? My own inclination is to say 'no' to both questions. The conception I shall spell out, below, which accords with these answers, is the most appropriate to Durkheim's view of social facts that I have come up with.

Consider the following case: a number of small black squares are put together to form a large black capital 'H'. Now, assuming that the property of being 'H-shaped' 'resides in' the H itself and not in its parts, the squares, what of the property of blackness?. We have ways of putting things which may lead to a denial that blackness 'inheres in' the H. We might say 'the H's blackness is the blackness of the squares' or 'the H is black just because the individual squares are black'. Perhaps we can expand these comments along these lines: any random assemblage of those parts would be black, just because the parts are black. We cannot, however, say similar things about the H's H-shape. We could say: the H is clearly more than a mere 'sum' of its parts. It might look, from this, that we are distinguishing here between properties a thing shares with its parts and those it does not, but this is not so.

¹ Herr apparently found Durkheim's views (as he saw them) unintelligible. See L.Herr, 1894, Review of the final edition of the Rules, Review Universitaire, 2.1., quoted in S.Lukes, Emile Durkheim, p.315

Consider a large square made up of smaller squares. It would not be appropriate to say 'The big square's squareness is the squareness of the small squares' or 'The big square has the shape it has just because its individual parts are square'. For the shape of the squares in combination depends upon the way they are combined: a random assemblage of small squares need not be square. So the big square, too, is more than a mere sum of its small square parts. A notion of inherence can be culled from this which I think fits well with Durkheim's text and intentions, and makes sense.

I shall pause with this rather informal criterion of inherence. I shall say that a property P of a thing T inheres in T rather than in T's parts or constituents if it is at least not the case that T possesses P just because T's parts possess P. More briefly:

a property P of a thing T inheres in T if it is not the case that T possesses P just because T's constituent elements or parts do.

2. A general strategy. Now we have an account of inherence to work with, let us consider a general strategy. We have already seen something of the difficulty of analysing the notion of a social group (in Chapter 2). Meanwhile our Durkheimian social phenomena must be clearly seen (a) to be inherent properties of social groups and also (b) to be produced by such groups.

Now suppose there were a group of inherent properties of social groups which could intelligibly be said to produce other such properties. It seems to me that these other properties would be examples of things produced by the social group they inhered in. This I shall take as an assumption: If an inherent feature f of a social group S produces a phenomenon ϕ , then this is in effect a case where S produces ϕ . If this does not seem immediately obvious, I hope that the example arrived at by the strategy it suggests will clearly be of the right type.

Our strategy, then, will be as follows. First, find some features of social groups which 'inhere' in those groups. Then show how these features

can give rise to others, which also inhere in social groups, indeed which inhere in the same groups. This will give us a kind of phenomenon which is both produced by and inheres in a social group. Such phenomena will reasonably be labelled 'Durkheimian' social phenomena. Once this task is completed we can consider the apparent claims on social science of these phenomena.

3.An initial requirement. Durkheim does not deviate in the Rules from speaking of social facts as 'ways of thinking, acting or feeling'. I do not think of this reference to 'ways of thinking' and so on as itself part of the basic conception, but rather as the beginning of the (arrested) development of that conception into something more concrete, and more intelligible. It is possible that Durkheim's basic conception itself arose out of his sensing the special nature of certain ways of acting, and his feeling that this special nature was the nature of the social, or the social most properly, or usefully, conceived. Clearly, Durkheim believed that there were ways of thinking and so on which fitted his basic conception as outlined above, and he may have believed that these 'ways' alone did, even could, do so. See his reference to ways of existing (*manières d'être*) - like the distribution of a population over its territory - as 'crystallized "ways of acting" ' (*manières de faire consolidées*)¹

I see 'ways of acting' entering the discussion, then, as Durkheim makes his conception more 'concrete'. The reference to 'ways of acting' and so on seems clearly to be less tentative and clearer than other aspects of Durkheim's attempt to be more concrete, like his ambiguous, much-criticised, references to 'constraint' as a constant feature of social facts.

I shall take it as a requirement on our production of an example of a Durkheimian social fact, then, that this be an example of a way of acting, thinking, or feeling.

¹ Rules, English, p.12; French, p.13. (For the whole discussion the French Edn. pp.12-14 should be consulted. The English edition inexplicably omits a paragraph crucial to the sense of the whole).

3. Collective practices.

i. Introductory: examples.

I shall propose as inherent features of social groups practices of a type I call 'collective'. First, I introduce an 'everyday' notion which seems aptly named that of a 'collective practice'. A working definition of 'collective practice' will then be proposed and justified. I shall argue that collective practices in my sense inhere in the social groups they characterise.

A caveat is in order. Durkheim himself uses phrases like 'collective practice' and sometimes writes as if all 'collective practices' are social facts.¹ This does not, of course, show that he would have considered everything that will be a 'collective practice' according to my working definition to be a social fact. I shall not, myself, try to argue that collective practices, as I define them, are all by their nature 'produced by' social groups. I argue, rather, simply for their inherence in such groups, and according to my own neo-Durkheimian conception of a social fact, collective practices as such will not necessarily be social facts.

First, let us informally consider some examples of what we might ordinarily think of as 'collective' ways of thinking, feeling, or acting. (I shall henceforth use the phrase 'collective practice' as short for 'collective belief, feeling, or practice'.)

We sometimes ascribe ways of acting and so on to whole societies and nations, that is to very complex, highly organised groups. For instance, we ascribe the belief in a Pantheon to the ancient Athenians. We say that they approved of the institution of slavery, and that they spoke Greek. It is natural to label this belief, this attitude, and this practice 'collective'.

¹ See for instance Rules p.7 (English) p.8 (French) where he refers to the beliefs, tendencies and practices of a group taken collectively ('pris collectivement').

We can also ascribe such practices to smaller groups within societies, groups whose members are members of a wider society also, and perhaps of other 'partial' groups as well. These groups may have names, like the Cambridge 'Apostles', or lack them; they may have a code of written rules, or a constitution-like a Trades Union-or they may not. Thus someone might claim (I am not sure if the claim is true) that whereas most Greeks were indifferent or antipathetic towards male homosexuality, a group of Athenian intellectuals of whom Plato was a member strongly approved of it. Someone might say that in the opinion of this group most contemporary females were insufficiently educated to be adequate objects of a deep attachment, and so on.

The smallest possible group, a 'dyad' - two-membered- can have its collective practices: a married couple may divide the housework equally between them, think they are very enlightened, dislike conservative politicians, and so on. Families can have their 'traditions' - say of attending a certain football match each year, or going to a particular University. Traditions seem to be collective practices (think of some connected locutions: 'The Churchills go to Harrow', 'we never miss the Boat Race')¹.

It appears that groups with collective practices can exist unconfined by societal or national boundaries. Consider the 'community' of Scientists. There may be confusions about its collective beliefs, attitudes and practices, and perhaps there really is no such community, but many people are at any rate under the impression that 'scientists' (including scholars) have collective practices. Perhaps there is a collective practice of not accepting a theory without good reasons. Perhaps there is a collective attitude of disapproval towards plagiarism. This would not - unfortunately - mean that no one ever plagiarised another's work. Indeed plagiarism could be rampant in spite of such a collective attitude. Evidence of the existence

¹ For a recent philosophical essay on tradition see David M. Armstrong's 'The Nature of Tradition'.

of such an attitude might be, say, the vigorousness and indignation with which plagiarism was denied by plagiarists wishing to maintain their image as worthy of the status of scientist or scholar.

Another such example might be that of a 'worldwide' religion such as Roman Catholicism. There are Roman Catholics in many countries but we can ascribe to Roman Catholics 'as a whole' many beliefs, practices and attitudes. For instance we might say 'Roman Catholics believe in transubstantiation'. Now, there is something rather special about this kind of example, a rather general possibility is clearly relevant here. There are many millions of Roman Catholics. It seems possible that few of them actually believe in transubstantiation. Now for the sake of argument, assume that this is so. It could still be true, on one natural interpretation of this utterance at least, that 'Roman Catholics believe in transubstantiation'. For as long as the doctrine of transubstantiation is part of the official doctrine of the Roman Catholic Church one can say that Roman Catholics believe in it - in a sense. This sense seems to be something like this: 'it is part of the official doctrine of the Roman Catholic Church, to which Roman Catholics belong, that transubstantiation is possible'. (it is surely true that for at least some points of doctrine one will still count as a Roman Catholic even if one does not believe or accept that point.)

Now, there are plenty of 'collective practice' statements which cannot be given the special kind of interpretation which we are now envisaging. Suppose we say of a certain married couple that they believe all Americans are wealthy. Here I think we mean at least that each one individually thinks all Americans are wealthy. (As I argue later, we mean more than this). It would surely be absurd to claim that such a statement says or may naturally be taken to say that there is some 'dogma of the house' such that whether or not the individuals John and Angela actually think of Americans as wealthy, they are the Joneses and it is a tenet of 'Jonesism' that all Americans are wealthy. So it seems that not all 'collective practice' statements can be

taken in what we might call, for convenience, the 'dogma-related' sense. We may specify this sense, roughly thus: we are to insert 'are supposed to' after 'The so-and-so's' in statements of the form 'The so-and-so's think(or feel, or act) thus'. This would display openly the dogma-related sense of a statement.

I propose to concentrate on those collective practices which are ascribed in non-dogma-related statements. The two 'kinds' of practice uncovered here are surely very closely related, however.

I would hazard that, for one thing, often someone making a dogma-related statement could make his point otherwise, by means of certain 'ordinary' collective practice statements. Let me briefly justify this suggestion with respect to a simple example.

A governess reproves her charge 'Ladies do not giggle'. Whether or not ladies actually giggle may not be the point; a dogma-related interpretation may be given to this statement: ladies are not supposed to giggle, but what does this dogma-related statement, in its turn, come down to? What may be intended, in particular, is something like this: members of a certain wider society disapprove of ladies who giggle, and will be chagrined if they come across one. Or something more 'definitional' might be intended (as John Mackie pointed out to me). Perhaps there is a stress on 'ladies' as opposed to 'giggle'. That is: one will not count, in the wider society, as a lady at all, if one giggles. In either case a fact about an 'ordinary' collective practice is what is basic to the situation. In this case the fundamental collective practice is not, that of 'ladies' themselves, but rather of a wider society in

which some people are deemed to be ladies.¹

To sum up our discussion so far: superficially similar ascriptions of practices to groups of people can be divided into dogma-related and other. Not wanting to cast our net too wide we shall concentrate on the latter, which we in any case suppose to be more basic. The relevant kind of practice can be ascribed to a wide range of entities, from whole societies and world-wide communities, to married couples and intellectual cliques.

Collective practices, and indeed the groups which have them, may be stable or fleeting. The practice may change though the group remains: Plato and his friends might alter their views on homosexuality, but remain a cohesive intellectual clique; the 'American people' became disillusioned with the Presidency. Or the practices (the collective practices) may disappear because the group does. Thus the married couple divorce, and no longer divide the housework between them; they may still each individually dislike conservative politicians, but it is likely that we will no longer ascribe a collective attitude to them. One reason may be that we no longer have a 'group-name' or 'description' for them, which seems to fit. Or, what must be the nub of such a point, they no longer form any sort of 'collectivity' (probably) and it seems that typically it is to social groups, collectivities, or social 'units' of some kind, that collective practices are ascribed.

1 Clearly a somewhat different account will have to be given of cases like that of 'Catholics believe in transubstantiation'. For one thing, it is presumably not a matter of what people in some wider circle than that of Catholics believe or feel about Catholics. I do not have a precise account of this case at present. The dogma-related statement here could (perhaps) come down to a statement about what Catholics themselves believe they ought to believe (and do), namely (say) whatever the Pope says is the case or is to be done. It seems possible that each one could hold this general view about what he ought to believe and do, yet 'lapse' sometimes in particular cases, for various reasons. There is obviously room for more discussion of this kind of case.

We have seen the range and general nature of the entities to which collective practices are ascribed. We might briefly consider their possible contents.

Though, typically, ethnographers may make special mention of relatively idiosyncratic collective practices, beliefs, and so on, there is no obvious reason to suppose that these cannot be quite commonplace, or even universal in human societies. After stating that a certain tribe had some particularly unusual belief, say that fire is a benevolent god, one might after all have occasions to note that 'They are quite aware that fire burns humans', 'They avoid burning buildings where possible' and so on.

Collective practices like the last two are surely ubiquitous, at least common to many a place; however they are surely collective practices.

Some recent sociologists have, in fact, placed enormous stress on the importance of the 'shared presuppositions' of everyday life in whatever group one is studying, including one's own. Many of these 'presuppositions' will surely be commonplace collective beliefs and practices, such as the belief that fire will burn.¹

1 See, for instance, A. Schutz, Collected papers: I. The Problem of Social Reality, pp. 13-14: ' "rational action" on the commonsense level is always action within an unquestioned and undetermined frame of constructs of typicalities of the settings, the motives, the means and ends, the courses of action and personalities involved and taken for granted. They are, however, not merely taken for granted by the actor but also supposed as being taken for granted by the fellow-man'. See also H. Garfinkel; Studies in Ethnomethodology (p. 36) 'the familiar commonsense world of everyday life - In the social sciences, and in sociology particularly, it is a matter of essential preoccupation'.

Collective practices may be relatively simple or complex. Thus Orthodox Jews 'keep kosher'. Keeping kosher is a complex of activities, involving obedience to the laws of 'Kashruth'. These laws dictate, among other things, that one not mix meat and dairy products in any one meal, in particular that one wait for some hours after eating meat before eating dairy products, and so on. Thus in some groups 'keeping kosher' is a collective practice and, say, not eating milk and meat together is, also. It is clear that, in this last case, an understanding of the simple practice will depend upon an understanding of the complex of which it is a part.¹

It is clear from the foregoing that collective practices, beliefs, and attitudes can have almost any content: they could be beliefs which every normal human has, or have a content extremely complex and unique to a particular group. It is also clear that collective practices are, intuitively, important. If we want to know about an 'alien' tribe or nation, we want to know about their collective practices, those they share with us and those more idiosyncratic. 'Culture-shock' is to a large extent shock at finding oneself in groups with collective practices one is not used to. Obviously it will be useful for the social scientist, and the philosopher of social science, to have some fairly precise characterization of what a collective practice is.

To this end, consider some cases where we would not talk of collective practices.

If just a few people in a certain group believe that so-and-so, we shall hardly say that the group believes that so-and-so on this basis, and similarly for practices and attitudes. Unless a majority, perhaps a largish majority of the members, participate, it seems that we can hardly attribute the practice to the members in general.

1 For an anthropological attempt to explain 'the abominations of Leviticus' in general see Mary Douglas's essay of that title.

Is it, then, perhaps a necessary condition on a group's ϕ -ing, or on the members ϕ -ing collectively, that everyone ϕ 's, or ^{that} a clear majority of the members ϕ individually? If in some or all cases it is necessary, is it then sufficient, and so are collective practices merely practices widespread in groups? To this issue we now turn.

ii. On the logic of ascriptions of collective practice.

Consider the following sentences:

- (1) The Zuni believe that the north is the region of force and destruction.¹
- (2) The Joneses do not use contraceptives.
- (3) The Germans are at war with the British.

These sentences differ in a number of ways but are similar in one respect at least. One who hears any of them will probably take it that the tribe, nation, or couple referred to are being referred to as in some sense a unit, or as a number of individuals taken collectively.

Sentence (3) has a unique feature of the three. In order to describe it let us define a 'direct individualistic inference' as (a) an inference to 'all or most members of G ϕ ' from a sentence about a group of the form 'the G ϕ ' or (b) an inference from 'the G do not ϕ ' to 'most members of G do not ϕ ' or 'few if any of the G ϕ '. In other words we infer from the fact that (as we might express it) the group does ^{something, that all or most of the members individually do (do not do)} (does not do) $\wedge$ that very thing. Now it seems that sentence (3) above, unlike (1) and (2), does not admit of direct individualistic inference.

1 See Primitive Classification by E. Durkheim, and M. Mauss, tr. R. Needham, p. 44.

From 'The Germans are at war with the British' it seems that we cannot infer of each (or of almost every German) that he is at war with the British. Perhaps it makes no sense to say of an individual that he is at war. Or perhaps in general it just need not be the case that if a tribe or larger unit is at war with X some or most members of that unit are at war, individually, with X. (Cp. 'The tribe has a large army'. Perhaps each member of the tribe has a large private army. But there need be no private armies at all in order for the whole tribe to have its army). This does not mean, of course, that the Germans' being at war with the British is not just a matter of a complex of facts about individuals. For the claim that some statement does not admit of direct individualistic (logical) inference is simply a claim about the form an individualistic inference from that statement may have (or rather the form such an inference may not have) and not a claim that no kind of individualistic inference or individualistic analysis is possible. In not admitting direct individualistic (logical) inference, for whatever reason, statements like that made in (3) seem at least to that extent to refer to what we might call the activity of a collectivity. However, in developing a notion of collective beliefs and practices I shall not attempt to deal with activities of this type. My reasons for this are similar to those which applied in the case of dogma-related statements. First, the truth of such statements is, it seems to me, almost certainly going to be a function of the truth of the essentially simpler kind of statement I shall deal with. Second, for the sake of considering one phenomenon with at least relative thoroughness, I do not want to cast my net too wide, and there may be quite distinct problems attached to these different kinds of statements. Indeed, collective practices of the kind I try to isolate in this section almost certainly have their own usefully discriminable kinds.

(1) and (2), unlike (3) do seem to admit of direct individualistic (logical) inference. Thus it seems that there is an interpretation of these sentences for which they entail, respectively,

(1') All or at least a large majority of the Zuni believe
that the north is the region of force and destruction,
and

(2') Neither of the Joneses uses a contraceptive.

However, even given that such an inference is warranted, it is clear that there is an interpretation of these sentences (perhaps indeed the most natural interpretation) for which (1) and (2) are certainly not equivalent to (1') and (2') respectively.

The following considerations show that this is so. One who hears (1) will not expect the utterer of it to be thinking or immediately afterwards to utter($\emptyset$):

($\emptyset$) Perhaps each member of the Zuni believes that the north is the region of destruction but at the same time is afraid to tell anyone else that he believes this, since he is afraid that the others will mock him, and that they certainly do not believe it. Perhaps each one keeps his beliefs totally secret and is quite ashamed of it!

We would not expect to hear $\emptyset$ after (1), and if we did we might take ($\emptyset$) as an implicit criticism of the statement that (1), as embodying the thought, perhaps, that the utterer of (1) has got his facts wrong and is saying that the Zuni believe such and such when in fact it is only the case that each individual believes that thing and it is not the case that they believe it collectively. If the utterer of (1) was faced with a response like ($\emptyset$) he would probably take it as a criticism of himself, or perhaps as evidence that he had been misunderstood.

similarly, you would not expect a hearer of (2) to reply 'I know - neither of the Jones's uses a contraceptive, but each thinks that the other does' (ψ).

($\emptyset$) and (ψ) might be argued, perhaps, not to contradict (1) and (2), respectively, outright, but rather to contradict one of two senses which (1) and (2), which are ambiguous, may bear. Alternatively, it might be claimed that no contradiction is ever involved, but a conflict with a certain implication of the

speaker may be occurring when for instance an utterer of (1) is surprised by a response of ($\emptyset$). That is, it might be claimed that it is not a matter of logical implication of the negation of ($\emptyset$) by (1) that is ever at issue. I need not argue with either of these suggestions here, since the main point I want to make is about a kind of thought we might have, and this could be carried by sense or by sense-plus-implication as far as I can see. It is fairly clear that when someone says (1) or (2) he may well mean to convey something with which ($\emptyset$) or (ψ) is in conflict.

It should be clear from the foregoing examples that we have a notion of a collective practice (including collective beliefs and attitudes) which are not just matters of widespread correlative individual beliefs or practices, though they do appear to be at least this. It is of course a further question what distinguishes collective practices from practices merely general throughout a group. On the basis of the foregoing it seems that something like the 'openness' of the beliefs in question is an important aspect of the collective case.

iii. Elements for a definition of 'collective practice'.

1. Group common knowledge.

A precise account of a notion close to one intuitive idea of a collective practice can be derived from the idea that the 'openness' of such practices is important. Take the case of the Zuni's (supposedly) collective belief that the north is the region of destruction (call this the belief that p). We assume that if this collective belief exists then all or at least a large majority of the Zuni believe that p, but it seems necessary also that it is not the case that each individual Zuni thinks (mistakenly) that almost no Zuni believe that p.

Moreover, it seems not nearly enough to say that each Zuni merely lacks the belief that others do not believe that p.¹ It seems closer to the requirements of the case to say that each Zuni must know that almost everyone in the tribe

¹ For more on the weakness of this condition, see Section 4(vi) below.

believes that p. But even more than this degree of knowledge seems to be required. For example, if each knows that the others believe that p, but also thinks that they are under the impression that all apart from themselves lack this belief, this would not seem right either. It seems that members must have more insight than this into their fellows' knowledge of a collective belief.¹

I contend that the way to rule out (false) beliefs about lack of belief in p, or (false) beliefs about lack of belief in beliefs in p, and so on, in this connection, is to give an account of collective practices roughly as follows. First, we define 'group common knowledge in a population' thus:

It is group common knowledge in a population P that..., if
and only if there are two or more members of P and (a) all
members of P know that...; (b) all members of P know that (a)
and so on.²

Some notes on my use of the term 'population' are in order here. (1) Any condition on people, including conditions stipulating particular people as members, defines a population (of people). Thus, 'Tom, Dick and Algernon', 'those who have walked on the moon', 'the Nuer', 'the Russians', and so on, define populations. (2) Two such conditions, C_1 and C_2 may be true of exactly the same people, yet it may be group common knowledge among those satisfying C_1 , but not among those satisfying C_2 , that p. For, in the sense defined above, it may be group common knowledge among, say, the Soviet spies in London, though not among those seated round the dinner table, that the Soviets are planning to tap the phone of a certain Conservative

¹ Cf. David Lewis, Convention, p.59 where he considers a sequence of possible (mistaken) ascriptions of lack of beliefs, lack of belief in belief, and so on. 'The cases become more and more unlikely, but no less deserving of exclusion'.

² Readers of D.K. Lewis's Convention and S. Schiffer's Meaning will be familiar with the sort of notion involved here. Schiffer, p.131, defines what he calls 'mutual knowledge* in a group' similarly. His 'S and A mutually know* that p' is defined and discussed at his pp.31-6. Cp. my section on 'individual common knowledge'. See also Lewis, Convention, pp.52-6 (his definition of 'common knowledge' is on p.56); and 60-68. I am much indebted to these discussions, the first that I knew of on the subject.

politician. For secrecy may be so great that Soviet spies are not even known to one another, and thus the assembled company are not aware that they comprise precisely the Soviet spies. Then, if in the definition of group common knowledge just given we replace 'members of P' by 'the Soviet spies in London' all the conditions will be satisfied. On the other hand, if we replace 'members of P' by 'those seated at the dinner table', condition (a) in the definition will still be satisfied, but none of the conditions from (b) onwards will be. (3) To fix ideas I shall make what seems to me to be the natural stipulation that having or lacking group common knowledge is a property of the population P in question.

(1), (2), and (3), imply that populations are such that two conditions may be true of exactly the same people but determine different populations. A population is not, then, merely the set of its members.¹ Social groups are populations in this sense.

Having established a notion of group common knowledge in a population we next partially define 'collective practice of a social group' thus:

a practice Pr will be a collective practice of a social group G only if it is group common knowledge in G that Pr is a general practice in G.

This partial account of collective practices gives these practices a high degree of 'openness' within the group. Various doubts about my notion of group common knowledge are, however, possible. I shall therefore shortly attempt to

¹ If one wishes to think of a population as simply the set of its members, one must use the term in some manner inconsistent with (1), (2), and (3); for example, someone might deny (3) and say that group common knowledge is a property of a population only relative to some defining condition; in such a terminology, the single population comprised both of the dinner guests and Soviet spies is aware qua being the Soviet spies of the Russian plans, but not qua being the dinner guests. I believe that there is room for more work in this area. Others who have used concepts like my group common knowledge in a population, in particular S. Schiffer and D. K. Lewis, have not discussed the logic of 'population' talk, or have done so only tangentially and unsatisfactorily (cf. D. K. Lewis, Convention, pp. 64-68).

justify the use of this concept further.¹ Part of the justification will involve an argument for the realism of the notion, in the light of certain considerations about knowledge.

Meanwhile one issue concerning our group common knowledge condition may be dealt with at once. Assume that 'normal' adult members of human groups can indeed fulfil the conditions on group common knowledge. Yet, it might be argued, (a) children and certain adults (say, some who are very unintelligent) may be members of social groups yet, (b) they cannot be parties to the knowledge definitive of group common knowledge: they will not, for instance, have the relevant knowledge about most members of their group. Yet we have just required that all members of a group with a collective practice be parties to certain items of group common knowledge.

It could perhaps be argued that children and certain special categories of adults - those in particular we conceive of as incapable of participating in group common knowledge - are not truly members of the social groups to which they are 'attached' by, say, being children of or being cared for by group members. However there is something at least somewhat odd in ordaining by fiat that certain people will not count as, say, members of the Zuni tribe, when those who should know - the (other) Zuni members - count them as such. In any case, were we to allow this move, it still would not be sufficient, for the following reason.

It seems that not only those incapable of being parties to group common knowledge in a social group may fail to be parties to it.² For instance, suppose one Zuni member goes to Oxford and while there ceases to believe that the north is the region of force and destruction; he then returns home to his tribe, and happens to hear few people express this belief. He thus comes (not so unreasonably)

¹ See especially subsection (iv) D, below, and section 4.

² I am indebted to Saul Kripke for the points made in this paragraph and the next.

to believe that the others lack the belief he in fact lost going to Oxford. Must his fall from participation in the relevant group common knowledge entail that the Zuni, as a group, no longer believe that the north is the region of force and destruction? (Assume he too ceases to talk about the matter). Surely not.

Various moves concerning this point are possible. Some, however, will not do. David Lewis, for instance, argues (in a different but related context) that we can ensure that something like our condition of group common knowledge obtains in a population by using the following device: 'anyone who might be called an exception might better be excluded from P'.¹ Thus he uses the condition to specify the population in question.² But it is hard to see how to avoid a circle here. How can the common knowledge condition be defined antecedently to the population? ^{The population is} / here supposed to consist of those people among whom it is common knowledge that ϕ -ing, say, is general in the population...; On this account we cannot fill in the blank here, and specify the population, without circularity.

There is a case then, for amending our partial account of collective practices so that it is not only unnecessary for everyone in G to conform to the practice, but also unnecessary for all to know that most conform, for all to know that all know that most conform, and so on.³ On the other hand one can certainly argue that there are circumstances in which the present strong conditions hold. In particular, one would expect them to hold in small groups made up of intelligent, sane adults, where there is much interaction between

¹ Convention, p.77.

² Cf. Schiffer, p.131. 'I propose to borrow from a suggestion put forward by David K. Lewis, and to treat the mutual knowledge* condition as partially defining membership in group G'.

³ David Lewis appears to adopt such an amendment, in his account of conventions in his article 'Languages and Language' see pp.5-6, though he does not there explicitly acknowledge a change.

members and where evidence is easily available to all, and known to be available to all, that a certain practice is indeed general throughout the group. It may be that larger groups, or groups with members who cannot be reckoned to understand their circumstances, are held to have collective practices insofar as they resemble to a greater rather than lesser extent more homogeneous small groups. I shall continue to couch my discussion in terms of the simpler, bold account just given, for simplicity's sake, having noted the issues.

2. Recognitional capacities versus concepts.

Is it a requirement on collective practices in a social group G/that members of G be able to recognise one another as members of G? That it is a requirement is suggested by some remarks by David Lewis, and by Stephen Schiffer.¹ In this section I argue that it is not. Let us first look at the remarks referred to.

Lewis has just claimed that we must have 'common knowledge' of our conventions. Lewis's 'common knowledge' is defined somewhat as our 'group common knowledge' is.² Now, Lewis claims later (p.64) that the knowledge involved in conventions may be 'confined to particular instances, taken one at a time' and goes on to argue that the agents involved must 'be able to recognise instances of S and derive the proper particular attitudes' (p.67). What goes for 'instances of S' (the situation to which the convention relates) goes also, according to Lewis, for 'members of P' (the population in question). He describes a population which has 'propensities to come up with the right particular beliefs regarding any new case that is presented in sufficient detail' as having the essential component of knowledge of convention (p.67). For Lewis, then, 'ability

¹ D.K.Lewis, Convention, last para, p.66-67, end of last full para; S.Schiffer, Meaning, pp.130-1, esp. Comment 2, p.131.

² In Convention, it is in fact defined in a more complex manner. See Convention, p.58.

to recognise' members of one's group appears as a necessary part of something very like our group common knowledge. I shall argue that in no natural sense of this phrase is it a necessary part of group common knowledge in our sense. I shall further argue, that no 'recognitional capacity' need be posited in addition to the requirement of group common knowledge, in the definition of 'collective practice'. Before explaining my position, let me turn briefly to Schiffer.

I find Schiffer's remarks about recognitional capacity puzzling. He seems either to be saying what Lewis is saying or to be saying something else, which is also false. He has just given an account of, roughly, what it is for a term 'X' to mean 'p' in a group G. (p.118-130). According to Schiffer a necessary and sufficient condition for this is that there be (what I call) group common knowledge in G that two other conditions hold.¹ Schiffer then writes: 'it follows from conditions (1) and (2) that members of G are able to recognise members of G'. To my mind, this follows neither from Schiffer's conditions (1) and (2) nor from them plus the general condition that they be group common knowledge. However, the total phenomenon he is trying to capture seems to be something we would call a 'collective practice'. Perhaps, then, Schiffer has the idea

¹ These conditions are: '(1) if almost any member of G utters something M-intending to produce in some other member of G the activated belief that p, then what he utters might be X. (2) if any member of G utters X M-intending to produce in some other member of G the activated belief that p, he will intend the state of affairs E (which he intends to realise by uttering X) to include the fact that X is such that there is a precedent in G for uttering X and meaning thereby that p (or an agreement (or stipulation) in G that X may be uttered to mean thereby that p)' (p.130). (The precise meaning of the technical terms used here does not affect the points I shall make).

that, whether in addition to or in consequence of group common knowledge, collective practices involve the recognitional capacity in question. In my view, this too would be wrong. Let me justify and clarify my own position.

From the existence of group common knowledge in G, as defined, it follows that each member of G has the concept of a member of G. Does it follow from the fact that there is group common knowledge in G that members of G are able to recognise other members of G? I would say not, but one can be more specific than this. What is recognising? As I understand it, it involves at least this: one recognises a person if one realises who he is. Presumably one recognises something as a thing of a certain kind if one realises that it is of that kind. Typically, if not by definition, one recognises something by the way it appears to the senses, for instance, by the way it looks.

Now, membership in a particular social group need carry with it no 'external marks' on the basis of which members may be recognised. So, if one is thinking of recognition by 'external appearance', those who are parties to group common knowledge in a social group need not be able to recognise each other in this way. Nor must they be able in practice to recognise each other, in general, when they meet, even given ample opportunity for conversation, observation of behaviour, and so on. Imagine cases involving a group of spies. Such a group may, clearly, have items of group common knowledge, and, indeed, it would seem, collective practices: members of the spy division may, say, (collectively) believe that to be detected in the course of duty - to be recognised - is worse than committing murder (in the course of duty). Each one may be so good at his job, he manages to fool even the members of his own division.

Is there any obvious kind of recognitional capacity we have not yet considered?. Perhaps someone will still argue that, given group common knowledge in a social group G, members will be 'able to pick out' other members. I am not sure, however, what meaning can be given to this suggestion. Perhaps all that

is meant is that members of G will have the concept of a member of G.¹ I have already agreed that this is true. But this certainly does not entail ability to tell by external marks, or to be able in practice to recognise who are the members of G around one.

I conclude that in no obvious or natural sense of this phrase is 'ability to recognise one another as members of G' a condition on group common knowledge or on collective practices in G. There is however a condition which might - I think erroneously - be thought to entail such a condition. This is possession of the concept of a member of G. But this is, at any rate, not a genuinely additional requirement on collective practices, given that the group common knowledge condition is fulfilled.

3. Self-ascription of group membership.

I now argue for the inclusion of one further condition on collective practices. This may, however, be otiose. Whether or not it is depends on whether or not members of a social group could lack knowledge of their own membership in the group. Assume that this is so. The question then arises whether it is a condition on collective practices that each member of the social group in question knows he is a member (or more mildly, that the 'normal adult members', most of them, know this). I am not sure that I can 'prove' anything here but I find such a condition appealing, and shall try to show why, briefly.

First, let 'p' be any item of 'common sense' that anyone will rightly expect everyone else to know. Suppose that N knows that p, knows that anyone who is a normal human knows that p, knows of the Y tribe, and hence knows that all normal

¹ Lewis, however, explicitly denies that in order for population P to have a convention, members of P must have the concept of a member of P. [I take issue with Lewis on this in the next chapter, section 4, (ii).]

adult Y members know that p. N could know of all this and still not be a member of the Y tribe. Now suppose (perhaps per impossibile) that N is a member of the Y tribe, but does not know it. He is in exactly the same position with respect to the Y's knowledge that p as any non-member of the tribe who knows what we've said he knows. If we now (again perhaps per impossibile) suppose every Y member is in N's situation, I personally find it hard to ascribe to the Y tribe any collective beliefs at all. Perhaps also it is hard in these circumstances to talk of 'the Y tribe' (as a social group) at all.

If we do make this an explicit condition on the collective practice of a social group, then, it may be redundant; however it is possible that some essential features of social group existence are not in themselves essential to collective practices, so it would be worth making this condition explicit, that is: each member of the social group in question must know that he is a member.

It is worth noting that by this one move - the requirement of knowledge of group membership - we have changed the situation in an interesting and important way. Now that Y members must know they are such, as well as knowing what all Y members know, and so on, they will have sufficient grounds for locutions like 'We Y believe so-and-so'; their knowledge of their collective practice is now expressible this way. It is perhaps because this thought is not rational without sincere self-ascription of membership that we find it odd to refer to people who do not think in this way in 'collective' language at all. If the use of the pronoun 'we' is the clue to our intuitions here then perhaps a further condition should be added: each member of G must know that he is not the only member of G. This seems, independently, reasonable enough.

iv. A working definition of 'collective practice'.

A. The definition.

The minimal account of collective practices that we have now arrived at is as follows:

A practice Pr is a collective practice of a social group G only if:

- (a) G has two or more members
- (b) most members of G conform to Pr
- (c) it is group common knowledge in G that (b)
- (d) each member of G knows that he is a member of G
- (e) each member of G knows that he is not the only member of G.

For our purposes here we shall adopt a working definition of 'collective practice' in terms of the above conditions. It is not essential for our purposes to enquire whether the set of conditions here maintained to be necessary is indeed sufficient. Even the account we have so far is quite rich in its content and implications. We shall shortly spend some while defending the realism of the notion as defined so far. In the rest of this section we discuss three further important issues which can usefully be taken up at once.

B. Inherence.

Recall our account of 'inherence', applied to social groups: a property P of a social group G inheres in G rather than in G's members, if it is not the case that G possesses P just because G's members do.

Now, I take it that collective practices are properties or features of social groups, and in my definition a 'collective practice' was, accordingly, defined as the practice of a social group. In every day parlance, moreover, we refer to collective practices using such phrases as 'That couple believe that p', 'This tribe will not countenance q', and so on. To some extent, indeed, in its expression of thoughts about collective practices, our ^everyday language may have some flavour of 'group mind' talk. However, a people's (collective) belief that p is at least partially a function of all or most individuals concerned believing that p. This might make one think that a group would believe that p

just because its members did. However, as I have argued, a group's belief is not wholly a function of what the individuals in it believe. I have also argued that it is not anything very mysterious that is involved here, in addition. Or so our working definition indicates.

A 'group belief, attitude, or practice' is not wholly a function of the correlative individual beliefs, attitudes, or practices; in addition it depends on a rather complex set of facts about the cognitive states of individuals and the relationship between these states. We can, therefore, conclude that collective practices, on our account of them, do inhere in the social groups they characterize.

C. The question of 'individual common knowledge'.

In this section I argue that group common knowledge is a necessary aspect of collective practices, as opposed to what might be called 'individual ^{common} knowledge', a closely related and important phenomenon.

Let us say that individuals F and H commonly know that p iff:

F knows that p

H knows that p

F knows that H knows that p

H knows that F knows that p

F knows that H knows that F knows that p...and so on, ad inf. Here and in what follows we read 'F knows that H knows...' etc. as 'F knows of H that he (H) knows...' etc. Exactly what it is to know of a particular person that... needs an account; I take this notion as primitive here, however.¹

We can extrapolate from this paradigm to cases involving any number of persons greater than one. Let us use the idea of orders of knowledge. Knowledge of knowledge is second-order knowledge, and so on. Other-related knowledge is knowledge of another person's knowledge. As long as each bit of knowledge in a given chain of nth-order knowledge is other-related (except the first) we shall say we have a bit of nth-order other-related knowledge.

We can now generalise relatively briefly from our two person case. For any population P, the members of P commonly know that...if (1) each member of P knows that..., and (2) every chain of higher-order other-related knowledge of each member's knowledge involving members of P occurs.

¹ On some conditions for such knowledge see David Kaplan, 'Quantifying in'.

Thus, where P comprises individuals 1..5, if members of P commonly know that..., then, say, 1 knows that 2 knows that 3 knows that p; 2 knows that 4 knows that 3 knows that p, and so on. The intuitive idea is that no one's knowledge of p is hidden from anyone else, no one's knowledge of anyone else's knowledge of p is hidden from anyone else, etc.

I shall say that it is individual common knowledge in a population P that... iff the members of P commonly know that...¹

I would argue that the phenomenon of individual common knowledge is an important phenomenon which often occurs, though it is less likely to occur the larger a population grows.² Moreover, it may be argued that individual common knowledge is of more immediate practical importance in human interactions than is group common knowledge, which in itself does not entail that anything is commonly known by any two individuals. I do not, however, make any use of the notion of individual common knowledge in a population in my account of collective practices.

There is a simple reason for not defining a collective practice wholly in terms of individual common knowledge. If it is individual common knowledge in P that..., then, as we have noted, each member of P knows ~~of~~ each other member.

¹ Note that on this definition of individual common knowledge in a population P, every member is involved. Some particular members of P, say F, H, and G, might commonly know that p, without all members of P knowing that p, or knowing of some other member's knowledge. In such a case it would not be individual common knowledge in P that p, on our definition. Let P' be the population consisting of F, H, and G. Then it is individual common knowledge in P' that p.

² See section 4 on the possibility of common knowledge.

But a tribe of, say, a thousand members, will not be such that each member knows of each other member.¹ So since such tribes can, according to our intuitions, have collective practices, our intuitions cannot be captured wholly by reference to individual common knowledge. Now, it is not a requirement of group common knowledge as we define it that each member knows of each other member. Thus as far as this aspect of the matter goes, it is possible that our notion of a collective practice be captured by reference wholly to group, as opposed to individual, common knowledge.

Now, obviously, our account of collective practices will be simplest if we can put it wholly in terms of group common knowledge, and do not have to make it disjunctive. That is, if we do not have to say 'a social group G has a collective practice if either there is group common knowledge or individual common knowledge in G of a certain general practice (and perhaps some other conditions also)'.

I believe that, as far as my intuitions about particular judgements go, they do support a definition wholly in terms of group common knowledge.

I hope I can make this judgement persuasive by appeal to the following case, where there is individual common knowledge of a group's general practice, but no group common knowledge.

Let us assume that the Transylvanian spy division has been severely depleted by the enemy. Three men, A, B, and C, are in fact the sole surviving members of the division. Each knows that he is a member of the division, but, when he meets the others at a cocktail party, does not know that they are. Now suppose that they find themselves (rather awkwardly) speculating on the nature of the Transylvanian spy division among a group of other people. One of the other guests says 'I know for a fact (from a Russian spy I know) that all Transylvanian spies refuse vodka at parties!' It is now (we may assume) individual common knowledge

1

Cf. Schiffer, op. cit., p. 130, comment 2.

in G (= the Transylvanian spy division) that all Transylvanian spies refuse vodka at parties. For A, B, and C, commonly know that Transylvanian spies refuse vodka at parties and they are the only surviving Transylvanian spies. However, it is not, we shall assume, group common knowledge among Transylvanian spies that they refuse vodka at parties. As far as A knows, he alone of his colleagues in the spy division knows that all Transylvanian spies refuse vodka at parties - for he, he thinks, is the only such spy in this conversational group.

Moreover, he knows that this new information has a very unusual source.

He may reasonably believe meanwhile, that he is not the only surviving spy.

Assume that the same goes, mutatis mutandis for B and C.

Now, it seems to me that in the above case we could not say that Transylvanian spies had a collective practice of refusing vodka at parties. Yet it is individual common knowledge among Transylvanian spies that all such spies refuse vodka. This shows, I suggest, that individual common knowledge cannot do duty for group common knowledge in an account of what it is for a social group to have a collective practice.

Group common knowledge is, in any case, of considerable practical importance. Merely given certain facts about recognition, group common knowledge in a population quickly gives rise to individual common knowledge in that population, or in subpopulations thereof. To give an example, X and Y meet in England and exchange a few words in English. In the course of their conversation they discover each of them is in fact French; it is group common knowledge among Frenchmen that they speak a particular language, namely French. Given these facts, X and Y now commonly know that each of them speaks French; it is now, in other words, individual common knowledge in the population consisting of X and Y, that French is spoken in that population.

Some final questions about, in effect, the necessity of the group common knowledge condition, will be dealt with at the end of the next section, which

debates the condition's realism, and compares it with a number of alternative accounts which might be offered.

D. Collective practices and the definition of 'social group'.

Whereas our definition of the collective practice of a social group involved group common knowledge in a social group, group common knowledge in a population is at least definable without positing that the population in question be a social group.

The 'population-analogue' of a social group's collective practice, as presently characterized, would thus be as follows:

For some population P, with 2 or more members:

- (1) it is group common knowledge in P that most members of P regularly ϕ .
- (2) each member of P self-ascribes P-membership.
- (3) each member of P believes he is not the only member of P.

Here it is not stipulated that P be a social group. Let us for now think of this set of conditions as conditions for the collective practice analogue (cpa) of a population - any population whatsoever.

Two interesting questions then arise: (1) can a population P have a cpa, where P is not in fact a social group? (2) if not, we should consider whether a social group might be simply defined in terms of the cpa of a population, thus:

a population P is a social group if (and perhaps only if)

P has a collective practice analogue Pr.

We shall approach (1) through a consideration of (2), starting by noting the appeal of the conjecture it expresses.

The proposed account is attractively simple, and it has one striking intuitive justification. Given the definition of 'collective practice analogue' members of any population P which has a collective practice analogue have grounds for use of the pronoun 'we' when talking of the population as a whole (e.g. 'we sailors') each also has the grounds for attributing many shared items of knowledge to those in the population, including himself, even though the population itself

need not be specified in terms of anything shared - it might be the population consisting of Fred, Joe, Xenon, and Algernon.

There are some possible objections which I think can be rebutted. For instance, the ideas that, on this account, the whole of humanity could - in principle - form a social group, that there can be highly transient - even momentary - social groups, that mere dyads could be social groups, that very hostile people could form social groups. None of these - as I have already indicated - seem to me to be truly powerful objections.¹

However, there are some excessively worrying aspects of this account, which must lead to its rejection. One problem involves commonsense propositions.

Suppose some people possess a feature F (like having brown hair) such that all normal adults will know that they have F, if they have it. Suppose further that F-ness doesn't carry with it any obvious special tendencies, attitudes, and so on. Now, all normal adult humans with F will be parties to general human common knowledge of commonsense propositions, assuming that such exists (if it does not, a similar example can be constructed involving, say, common knowledge among normal adult English people).

Given all this, if p is any commonsense proposition, it will be group common knowledge in the population of those who have F that p. Moreover, the other conditions stated on collective practices will be fulfilled (we assume each of those with F knows that he is not the only one with F). So there will be a collective belief analogue among brown-haired people that, say, a deep cut in the hand is painful. Surely, one might feel, whether or not the population of brown-haired people is a social group, or could be, it will not be for this reason.

¹ Cf., on the last two counts, Georg Simmel: (in) 'small groups, such as the marital couple,... a certain amount of discord... is organically tied up with the very elements which hold the group together' (from 'Conflict and the Web of Group-Affiliations'; repr. in On Individuality and Social Forms, p.74).

Another implausible case of a putative social group is as follows.¹

Take some population, P, say, the population of brownhaired people, and some proposition p which is believed by some members of P. Suppose that it is reasonably believed by each P-member ^{who believes that p} that he is at least not alone ^{among P-members} in believing p (let 'p' = 'gentlemen prefer brunettes'). Now consider the population P', specified thus: P' is the population of those P-members who believe that p. The belief that p will surely be a collective belief analogue in this population. For it is analytic or tautologous that all the P-members who believe that p believe that p! We may surely expect all P-members who believe p to know this, and so on. It does seem quite counterintuitive, however, to call this particular population, P', a social group by reason of this belief. For how could the members' knowledge of analytic or tautologous propositions, albeit ones involving defining properties of the population, make a population a social group?

The foregoing does not prove that a fairly simple account of a 'social group' in terms of something like our collective practice analogues for populations cannot be given.² The specific, quite attractive suggestion mooted in this section must, however, be rejected.

¹ This case derives from a point made by Saul Kripke at a seminar I gave on this topic at a philosophy workshop in Banff, Canada, June, 1977.

² See, for a tentative account of this kind, pp 289 - 302, below.

4. In defence of common knowledge

i. Is 'common knowledge' possible?

It could be objected that neither individual nor group common knowledge, as defined above, are possible.

This might be argued on some such ground as this: no real creature, it might be argued, can have an infinite number of propositions in its head, and common knowledge requires numbers of creatures to do this.

Another objection might run as follows: infinity of propositions apart, surely we do not all have in our heads such interminably long or complex propositions as is implied by our having say a piece of 500th order other-related knowledge. (Similar objections might, evidently, be levelled at the 'double K' principle - that knowing that p (Kp) entails knowing that you know that p (KKp). It might be thought that this leads to the conclusion that nobody ever knows anything, since it gives unrealisable conditions.)¹

It is my view that there is a basically plausible kind of account of at least a kind of knowledge, according to which it is reasonable to suppose that common knowledge occurs. Features of this account which are of importance here are that (a) it does not necessarily involve one's currently entertaining, or indeed one's entertaining at any point in one's life, the proposition that p , in order that one knows that p . (b) Nonetheless, on this account, knowledge does involve belief, of a kind. Moreover, (c) in spite of its not necessarily involving current entertainment of p , this kind of knowledge can be shown to be a powerful possession, a tool against irrationality and temporary confusions.

In what follows, I shall sketch the kind of account I have in mind. The fact that this account makes common knowledge a possible phenomenon seems to me to be a point in its favour; common knowledge as defined does not seem to me

¹ The 'double K' principle, as formulated here, is by no means generally accepted by philosophers, nor does anything I have to say presuppose it. However, my arguments for the possibility of common knowledge would also stand as arguments for the possibility of knowledge, even given the truth of the double K principle.

to be something monstrous in conception, as some might feel.¹ I should note at the outset that the rough account I shall give is vulnerable to the famous Gettier counterexamples. Edmond Gettier has given counterexamples to the claim that knowledge is, roughly, justified true belief.² However, as we gather from Kuhn and others, it is not contrary to scientific practice to help oneself to a falsified theory if one does not have a better one, and if the 'bad' theory seems satisfactory in many ways. It does seem to me likely, moreover, that there are modifications to be made in an account of the kind I have in mind which will satisfactorily take care of the counterexamples without radically altering the account, and which will allow it still to support common knowledge as it does in the admittedly rough and partial sketch which follows.

ii. Knowledge.

We shall consider then as a partial analysis of 'X knows that p' the following: (1) p; (2) X has good evidence that p; (I shall say more about what (2) amounts to in a moment.)

On this spare account individual common knowledge that p between two persons F and H involves the following:

(1) $K_H P$, $K_F P$, that is:

(i) p

(ii) F and H both have good evidence that p.

1 See the discussion in the next section of some who evidently do feel this.

2 See Edmond Gettier(1963) Is Justified True Belief Knowledge? This brief paper has provided a focus for many recent discussions of knowledge and, in particular, attempts to give an analysis of 'X knows that p'. For such a discussion and some further references on this subject see, for instance, Gilbert Harman's Thought. (Are Gettier's examples really counterexamples to a simple justified true belief account of knowledge? Sometimes I have doubted their force; certainly I would be inclined to deny the force of some 'Gettier-type' examples Harman claims are examples of 'no knowledge' (in particular see Harman op.cit.pp.142-4))

(2) Every possible chain of higher-order other-related knowledge of F and H's knowledge, involving F & H, occurs, that is:

(i) $K_{H^*}p$, etc. (that is, every chain beginning with some order of H's other-related knowledge, involving H and F).

(ii) F has good evidence that $K_{H^*}p$ etc.
and similarly for H's knowledge of F's knowledge.

What, we may now ask, is 'having good evidence that p'? If it is to be knowledge that we are talking about, this should be interpreted as involving at least the following two components. First, if X is to have good evidence that p, then there must be good evidence available to X that p. 'Available to X' here means something like 'present to X'. Perhaps we should demand this: (a) X must be aware that a certain state of affairs E occurs in circumstances C, where (b) E in C is good evidence that p. I would say that if (b) obtains then, at least, by going on sound principles of reasoning (inductive or deductive) one would conclude that p on evidence of E in C. In other words, E in C strongly probabilifies or proves p according to sound principles of reasoning. Second, this good evidence for p must not just be something X is aware of. It must also be 'available to' him in another way at least. In particular he must have those (correct) standards of reasoning which would enable him to reach the conclusion that p, reasoning from the fact that E occurs in C alone.

Thus, 'having good evidence that p' entails at least (a) being aware of the relevant evidence (E in C) and (b) having (correct) standards of reasoning enabling one to reach the conclusion that p from this evidence.

Turning to aspects of common knowledge, we note that on the partial account of knowledge just set out, if S knows that J knows p, then S has good evidence that (at least):

(a) J is aware of a total situation which is good evidence that p.

(b) J has the (correct) reasoning standards to enable him to reach the conclusion that p.

(c) p.

iii. Knowledge and belief

It might be objected that this partial account of knowledge omits, in particular, any reference to belief that p on the knower's part, but whereas knowing is more than believing, does it not have to involve belief? On the other hand, it might be argued, insofar as knowledge must involve belief, the notion of common knowledge that p must be quite unrealistic. I shall argue that lack of realism here is not a necessary consequence of the assumption that the 'knowledge' in common knowledge involves belief - of a certain kind. I shall argue first, independently of this objection on behalf of belief, in favour of a modification of our account of knowledge so far; I then argue that knowledge now entails a kind of belief, a kind apt for real common knowledge, and for collective practices.

It seems to be a logical consequence of X's knowing that p, as this has been defined so far, that something like the following will hold of X (call it situation S):

(S) X could conclude that p, by applying his current standards of reasoning to the evidence he is presently aware of.

Under what circumstances, though, would X conclude that p? Suppose he has no impediments from lack of memory, time, concentration, perseverance or interest, and is not influenced by any kind of bias or prejudice about whether or not it is the case that p. Let us say that in this case he is reasoning 'effectively'. Now suppose that he is putting his mind to considering whether or not p, and starts from what is in fact good evidence for p. Will he then reach p? Not necessarily, even given fully effective reasoning on his part. The following example surely shows this.¹

Say X is told all the axioms of number theory. He also has the principles of deductive reasoning - these are his reasoning standards. Can he reach all the

¹ I am grateful to S.Kripke for the example.

theorems of number theory? Surely not. It is usually not for lack of memory or time that mathematicians fail to prove theorems they are interested in.¹ Does X know all the theorems of number theory? Plato might say that he does (cf. the discussion of geometry in the Meno). Yet common sense would surely give a different answer. No one yet knows all the theorems of number theory (or no one on earth anyway). What is the solution here - what error has Plato fallen into with respect to commonsense ideas?

Perhaps we should add something like the following condition for knowing - it is a further specification, if you like, of the way in which the evidence qua evidence for p must be available to X: X must know how to draw the conclusion that p - at least in some more substantive way than having the relevant principles of reasoning.

Now there are obviously many cases where in fact the 'proofs' of our conclusions, given our evidence, are 'elementary', where no flash of insight or special creativity is necessary to see how to arrive at them. Holmes's 'elementary my dear Watson' is a clever put down. Where flashes of insight are required, only once one has had them, is the 'deduction' or inference elementary. It is this kind of elementariness of inference that knowledge seems to require.

Let us now, therefore, alter our partial account of knowledge in the following way. We shall say that if X knows that p then:

(a) p;

(b) X is aware of some state of affairs, E in C, which is good evidence for p;

(c) X has the (right) standards of reasoning (to enable one, with sufficient insight, reasoning effectively, to infer p from E);

(d) p is an elementary inference for X, from E in C (i.e. X has the necessary insight to infer p from E).

¹ The recent proof of the '4-colour' problem is highly unusual in its use of humanly impracticable computations, such that it could not have been completed without the aid of a (non-human) computer.

In order to convert this into a set of necessary and sufficient conditions for knowledge, we assume that further clauses or modifications will deal with Gettier examples. Doubtless other clauses and clarifications will be needed too. Among these clauses, however, I would claim that there must not be a condition that X has entertained, or will ever entertain, the proposition that p, or even that he has the capacity or perserverence to do so.

Simple examples, which have nothing to do with common knowledge, seem to show that explicit entertainment of p cannot be a necessary condition. Consider X, quite a normal individual, who knows that his uncle is dead; X is sitting in his study. He has never entertained the proposition, or said to himself 'my uncle will never again enter this study', but surely he knows that this is true. Further, perhaps as a matter of his psychological make-up (say his inability to think explicitly about his uncle any more), X could never bring himself consciously to entertain this proposition. Even in such a case I think we would say X knew that his uncle would never enter his study again. Then there are the cases in which we talk of self-deception. X thinks to himself 'he will come again soon, I know it'. Yet we might still want to say that X all this time knows that his uncle will not come again. Far from p's having to be asserted or entertained, it may be hotly denied, yet still be known: or so at least some of our ways of talking suggest.

An important point about how we should construe 'effective reasoning' above may be usefully made here.

We have just considered an example in which X knows that p though he is actually incapable of entertaining p. In the example the problem envisaged was what we might call an 'emotional block', and our intuitions on the example establish that we allow that one can know a proposition that one cannot entertain for this reason. Apart from emotional blocks, another allowable reason might be the 'length' or complexity of p. (Say the proposition that N is a natural number where N is a 5,000 digit number and is specifiable by no easy algorithm.)

Given that we may be said to know propositions that we cannot entertain, I nonetheless assume that it must be true that were we 'reasoning effectively' in a broad sense we would entertain them - we would 'arrive at' them and 'take them in'. What this comes down to, in other words, is that our evidence and our reasoning standards must be such that only lack of certain specific kinds of things like time, perseverance, emotional calm, etc. prevent us from ever being able to reach the conclusion. To give an example of a lack that counts against knowledge, we might cite lack of concepts (like the concept of a natural number or the concept of knowledge). We have already cited lack of 'elementariness'. Regarding complexity, this view has the consequence that, to put it somewhat paradoxically, the proposition known can only be 'too complex' in a rather simple way. In my view the sums of very 'big' numbers, and items of knowledge of a very high order, are in this class. Almost everyone has the capacities such that the calculations involved are of an elementary kind for them.

It is possible that someone might still dispute that one can know the truth of propositions that he cannot actually entertain. In particular, in the context of discussions of common knowledge, it may be argued that highly complex or 'long' propositions cannot be known. Though my disagreement with this view has just been expressed. I shall consider some alternative accounts of common knowledge that have been proposed on account of it, before concluding my discussion of the reality of common knowledge.

To turn, at last, to the question of belief. It is a consequence of X's knowing that p as defined so far that something like situation S1 is X's situation.

(S1) X would conclude that p if he effectively applied his current standards of reasoning to the evidence he is presently aware of.

Is S1 a situation in which X believes that p? I myself think that it is. (Note If we were to say that in S1 X has a latent belief, I can hardly envisage an objection to this locution, except perhaps precisely that the belief here is not latent but is actual or more than latent.)

The claim that S1 is a situation in which X believes that p is perhaps suggested by the following kind of example: Professor Smith asks Professor Jones: 'do you think that Tweed is an intelligent student?'. Jones has never consciously considered this question, but all of Tweed's essays for him have been excellent. Jones replies quickly 'yes, I think he is very intelligent'. It would surely seem wrong even as a piece of pedantry for Jones to reply: 'I don't think anything about the matter at this moment, but give me a second or two and I'll figure out a view. Hm, yes-I now think, though I didn't before, that Tweed is very intelligent.'

It is, it seems to me, by no means obvious which is in fact to be considered a more important aspect of belief: explicit expression or certain factors, whatever exactly they are, that may underlie and give rise to such expression. The strangeness of the second reply considered above perhaps indicates that explicit expressions are to be considered as mere occasional 'epiphenomena' of beliefs or 'views' proper.

Finally, regarding the claim that knowledge involves belief, it seems to me that either this is false or the belief involved has to be no more explicit than that envisaged here.

I would argue, then, that our account of knowledge so far entails that if X knows that p, he believes that p. That is, X would assert p given (only!) the effective operation of his standards of reasoning, in an elementary way, on the evidence of which he is aware.

iv. An example

Consider how in a two person case the conditions for individual common knowledge might be fulfilled.

T and B Jones have agreed to stop using contraceptives until B gets pregnant. Each has ample evidence of the other's desire for a child, sincerity, and openness. B is aware of the above facts. So if she thought about it would conclude from them (rightly) that (1) neither she nor T is using a contraceptive; (2) that T is aware of facts enabling him to conclude that (1); (3) that T is

aware of facts that would enable him to conclude that (2); and so on: Given effective reasoning the sense could be continued indefinitely. The same applies, *mutatis mutandis*, to T. In other words p (neither B nor T uses a contraceptive); $K_B p$; $K_T p$; $K_B K_T p$; $K_T K_B p$...and so on. The inferences involved for T and B are elementary, given a good memory, time and perseverance. Thus the Joneses have individual common knowledge that p, for at least one substitution instance of p; they also have the corresponding beliefs.

There is also, for similar reasons, group common knowledge in the social groups, or unit, comprising T and B (Mr and Mrs) Jones. (They 'all' probably just refer to this group as 'us' or 'we'). It will be group common knowledge 'in' this unit, the Joneses, that the Joneses are refraining from using contraceptives, since everyone in the unit knows that everyone in the unit knows - and so on.

So much, then, for cases.

v. Objections reconsidered.

Let us now consider the objections to the very possibility of common knowledge noted at the beginning of this section. First there was the objection concerning infinity: no real creature can have an infinite number of propositions 'in its head'. Two variant interpretations of this spring to mind. First, no real creature can ever include in one 'judgemental episode' an infinite number of propositions; an infinite number of propositions can never be surveyed on one occasion by a real creature. Second, we might not bring in the idea of a single 'judgemental episode', but simply put the claim thus: no real creature can ever have entertained an infinite number of propositions. This second claim makes the real point of the first explicit: the real issue is surely the incompleteness of the 'survey' supposed.

Suppose that either variant is true. This will not entail the impossibility of common knowledge. Given our account of knowledge, at any rate, and the definitions of common knowledge, the latter does not require anyone to entertain an infinite number of propositions or to have entertained them. Indeed,

not one single member of the long series of propositions about knowledge which must be known for there to be common knowledge need ever be entertained by parties to common knowledge.

Similarly - to turn to the objection regarding complexity - on our account one is not forced to see parties to common knowledge as ever actually entertaining any extremely 'long' or complex propositions, or even as being capable of doing so, since certain special limitations of time, memory, concentration, and so on, are not held to rule out knowledge.

Finally, perhaps common 'knowledge' as we have explicated it is not really knowledge, and the corresponding 'belief' is not really belief. Nonetheless it is clear that 'common knowledge' as we have explicated it is possible, or, at least, not open to the objection stated, and presumably all of us have it in many connections. Surely also, given the possibility of the phenomenon outlined in this section, it is not necessary to retreat to the idea that 'common knowledge' phenomena must be defined in terms of people's 'not disbelieving' certain propositions, as some have recently supposed. Before commenting further - and I hope helping finally to halt - this recent trend, we can usefully note some further features of common knowledge as I understand it. The kind of knowledge (and belief) I claim is involved in common knowledge phenomena is, in spite of its limitations, a powerful kind of possession. I shall argue briefly here that if one has (this sort of) knowledge that p , then one has a tool against any (irrational) tendencies one may also have to assert that not p .

It has been proposed¹ that if you know that p , then it is not the case that you believe that p is false, or, in standard notation, ' Kp ' is argued to entail ' $\neg(B\neg p)$ '. Let us assume that correct inferential standards will never take one from a total state of affairs E to both p and $\neg p$. Then on our account

¹ See e.g. O.R.Jones (1975)

of knowledge (and associated belief), if Kp , then Bp , and not $B\neg p$ (for the relevant kind of belief). However, not all aspects or kinds of belief are necessarily involved in knowledge, and it seems that Kp does not rule out certain aspects or types of belief, that not- p , for instance, present assertion of not- p , plus a 'feeling' that one means what one says, that is, that one is sincere. Knowledge and its correlative kind of (rational) belief is not that powerful. On the other hand it is hard to see how other aspects or kinds of belief can destroy this one (given that the relevant faculties of reasoning remain). Thus T.Jones in our previous example might on some occasion suspect B is deceiving him and is secretly using a contraceptive, but if he still really knows that she isn't doing this (as he still might, if this suspicion isn't based on weighty new evidence) then he will be so much the less able to keep up with these suspicions for long, particularly if challenged and/or forced to think things through calmly.

vi. On some alternative accounts of 'common knowledge' phenomena.

In his paper, 'Languages and Language' David Lewis presents a common knowledge condition pretty much equivalent to ours in terms of group common knowledge, but he goes on to remark:

'perhaps a negative version...would do the job: no one disbelieves that (1) - (5) hold, no one believes that the others disbelieve this, and so on'.(p.6)

It seems clear, however, that this 'version' will do very little. Though some phenomena are indeed, rightly, ruled out as common knowledge situations, far too much is allowed in at least for our purposes here. Briefly, this condition can be fulfilled, with respect to a proposition p , by people who have none of the concepts involved in p , and in fact every population will have 'common knowledge' of an infinity of propositions they have no notion of. (What they have might better be called 'common ignorance' than 'common knowledge'.) A remote tribe in the Amazon, which has had no contact with industrial civilization, would, on this view of the matter, have group 'common knowledge' of the existence of televisions, typewriters, washing machines. Yet what they in fact have is no

beliefs about televisions (and so on) at all, since they lack the concept of a television. The phenomenon in question here, whatever one calls it, does not seem to be a very important one, nor to be nearly positive enough for our own needs in introducing an account of 'common knowledge'.

An alternative to an account like ours, which is not as weak as Lewis's, has been suggested by Christopher Peacocke (in 'Truth Definitions and Actual Languages')¹. He suggests that (for his purposes) one use a 'notion of common knowledge that finitely generates this paradigm:

it is common knowledge between x and y that p iff
all these conditions hold:

- (a) x knows that p
- (b) y knows that x knows that p
- (c) x does not disbelieve that (b)
- (d) y does not disbelieve that (c)
- (e) x does not disbelieve that (d)...

- (a') y knows that p
- (b') x knows that y knows that p
- (c') y does not disbelieve that (b')...' (p.174)

Here again is a use of the idea of 'not disbelieving that'.

In favour of his account Peacocke says:

'This definition does not imply that if x and y commonly know that p, then they are capable of highly complex propositional attitudes' (ibid).

Now, we have argued that our own account does not imply that parties to common knowledge are capable of expressing highly complex propositional attitudes or of in any clear way 'consciously representing to themselves' such attitudes. We have argued that propositions which never could be entertained by X can, nonetheless, be known (and believed) by X, and see no reason for saying that the propositions to be known in common knowledge, as defined by us, cannot be known and believed. Insofar as this is so, we need not feel moved to accept either Lewis's or Peacocke's account in preference to our own.

One unattractive aspect of Peacocke's account in general is that the cutoff point for requirements of a piece of knowledge that is given is surely

¹ In Evans and McDowell, eds. Truth and Meaning, see especially P.174

wrong. The account might appeal to someone who thinks our own account in terms of knowledge of knowledge ad inf., imposes too strong a condition and is unrealistic because the 'cutoff point' comes too late (i.e. never) and thus we require beliefs no person can express or consider. However, if the cutoff point comes too early, then some cases may be allowed in as common knowledge when, intuitively speaking, they should be ruled out. The stated account seems to have this fault. Let 'p' stand for 'x has stolen a book from Blackwell's'; we can imagine Peacocke's conditions being fulfilled for some x and y, regarding p. We can also, at the same time, imagine x wondering whether to confront y with his knowledge that p. I am imagining that x is not sure ^{whether} y knows that he (x) knows of y's theft or not. Peacocke's conditions are satisfied in this case, but this does not seem to be any kind of paradigm case of 'common knowledge', or in particular of the phenomenon we are trying to capture.¹ However, any cutoff point requiring more knowledge, and to that extent less counterintuitive, will at least, surely be arbitrary.

Since I do not feel the need to move to references to lack of disbelief in the definition of common knowledge, I shall not investigate the feasibility of doing this ^{in a way} that avoids both relative vacuity and the use of an arbitrary, or otherwise unacceptable, cut off point.

Before concluding I should note a rather different approach to the definition of 'common knowledge' phenomena which I have not yet mentioned. David Lewis takes the approach in Convention (though not in the later 'Languages and Language') and an attempt to improve on Lewis's Convention account has just recently appeared.²

¹ It may serve Peacocke's own purpose adequately; It is not necessary to investigate that question here.

² See Jane Heal 'Common Knowledge' Philosophical Quarterly, April, 1978. See also Schiffer, op.cit. P. 134. (Schiffer does not try to define common knowledge phenomena in terms of . finite conditions, but does give a finite set of conditions which he argues guarantees his 'mutual knowledge'.)

The general idea of these definitions is to define common knowledge in terms of a finite body of information which would enable those who have it to construct as large a section of the regress in our type of definition as is practically important for them to use.

I see no reason to adopt such a definition here.¹ An account of a finite set of conditions of the type mentioned could, however, obviously illuminate any account of common knowledge of the more abstract (and, I would suggest, more fundamental) type I have given. (Steven Schiffer evidently sees his own work in this area in such a light). I shall not pursue this aspect of common knowledge further here.

Finally, I should like to say something about what one could call 'common belief'.

One could define phenomena of 'common belief' analogous to those of common knowledge as defined, and certainly if there is a 'common belief' that a practice is general (though it is not) this may be of importance. However I think that the concept we have defined can take care of what is likely to be most important about 'common beliefs'. An example may serve to illustrate what I have in mind.

Suppose it is 'commonly believed' that almost everyone in a society S refrains from adultery (practice Pr) when in fact a large majority do commit adultery (-Pr). -Pr is an actual general practice, Pr is not. There will nonetheless most likely be a further, important 'collective practice' here, in our sense - there will be a collective belief that most people in S refrain from adultery; that is, everyone will know that there is this general belief

¹ Heal explicitly worries about some things I do not find worrying, like knowing ^lvery 'long' or complex propositions, and also about how common knowledge can figure in reasons for acting. On the latter topic see section 5 of this chapter, subsection (vii).

in the society, everyone will know that everyone knows, and so on, and this of course may have real effects: it does not in general (ex hypothesi) prevent adultery; but it may facilitate it (interestingly enough), since everyone can rely on the fact that they are unlikely to be suspected of adulterous aims or actions. It may, also, prevent adultery in particular cases: that of mistaken conformists, for instance. It may be a source of comfort to those who prefer monogamy but ceteris paribus prefer to conform. It may cause pain to adulterers who feel alone in their activities.

There may still be a case for a definition of a kind of 'collective practice' in terms of 'common belief' which is not also common knowledge. However, it seems that the conditions for our present definition are often fulfilled and, moreover, the cases involving common belief without also the related collective belief are going to be rather special and nonparadigmatic cases and should probably be distinguished from the kind our definition picks out. For one, the beliefs involved in common knowledge as opposed to 'common belief' may be thought of as having a basic kind of stability: there is at least no mistake to be uncovered and overturn the whole edifice. I shall not therefore pursue this idea further here.

vii. Conclusion: on common knowledge and collective practices

Our minimal conditions on collective practices already capture an important and widespread real phenomenon, whether or not some variant may be either closer to our common concepts or better adapted to certain scientific needs. The importance of the collective practices of social groups to the life of members of those groups was obvious from our initial informal characterisations of them; a host of actions of individuals ... and interactions between persons will be built on them, for instance. I believe that our working definition of the 'collective practice of a group', in terms of common knowledge, already indicates how collective practices informally conceived of have that importance: they obviously help to form a 'solid' shared 'social world' that individual

^{agents}ts can take for granted - as taken for granted¹ - in dealing with each other. These then, are the phenomena I take to 'inhere' in social groups, when they characterise them.

5. A social group produces social phenomena.

i Introduction

We shall now look at what we may call the production condition, that is, the idea that social phenomena must be produced by a society or social group.

So far we have given an account of a phenomenon which 'inheres' in a social group, the phenomenon of collective practice. We defined a collective practice as a practice general throughout a particular group, in which it was group common knowledge that the practice was general, where each member knew he was a member of the group in question and not the only one.

Now, Durkheim was quite concerned to deny that just any widespread practice in a society or social group is ipso facto a social phenomenon.² Yet presumably many of the widespread practices he wishes ruled out (like eating, sleeping and drinking) are group common knowledge in the groups in which they occur, are indeed collective practices.

Though Durkheim will probably not have explicitly considered this issue, he would, I take it, say that it was not enough for a practice to be 'collective' in our sense in order for it to qualify immediately as a 'social fact' in his.

It is clear that Weber agrees with Durkheim in this: the social, as each conceives it, is not to be found in just any general practice of a human group. Moreover, I think that Weber, also, would deny that collective practices as we have defined them here involve social elements as a matter of logical necessity. At least his official doctrine would surely dictate this, since social action as Weber defines it is not involved in collective practices in general, as a

¹ Cf. the quotation from Alfred Schutz in Section 3(i) of this chapter.

² See Rules, p.1 (English); p.3 (French)

matter of definition. For Durkheim, many such practices are ruled out, it would seem, by the notion that the society or social group in which his social facts occur should produce them.¹ However, we have yet to make the notion of production by a social group or society at all precise.

Durkheim himself does not give an explicit account of what he means by production by a society. In what follows I develop an account of what is essentially a type of production by a social group relating to practices; our neo-Durkheimian notion of a social phenomenon will be couched in terms of this. The account I work with is probably narrower than one which would cope with everything indicative of Durkheim's own thoughts on the subject, however, as I shall argue, there are advantages in a narrower account, as in an account of 'Durkheim's' social phenomena themselves which is of relatively narrow scope.

ii. A general account of production by society.

One might give a general account of what it is for a collective practice to be produced by its 'substrate' social group roughly as follows:

A collective practice Pr in a social group S is produced
by S if and only if S (or some inherent feature of S)
causes Pr.

This, it seems to me, is more or less what the most general account of production by a particular social group or society must be like.

I find this straightforward account clearly preferable to^a less straightforward kind of formulation which has been associated with it. Consider the following claim of Mauss and Fauconnet in their encyclopedia article 'Sociologie'

¹ Insofar as Durkheim and Weber would both judge our 'collective practices' to be 'subsocietal' my own intuitions on socialness would judge them to be wrong. See our concluding chapter.

'...it seems clear that the group, crowd, or society has its own nature, but it causes certain ways of feeling, thinking, and acting, in individuals, and that these individuals would not have had the same tendencies, habits or prejudices if they had lived in other human groups'. (my translation) (p.141).

What is the relation of the occurrence of a practice in different societies to its being produced by a given society? Have Mauss and Fauconnet given here what is in effect an explanation of what it is for one practice to be produced by a given society? I shall argue, briefly, that they have not. Consider first the following claim: a collective practice 'Pr' is produced by its social group, S, if and only if, had conforming members of S not lived in S (but rather in some other group) they would not have conformed to Pr.

If this is taken to mean that Pr must be peculiar to S if it is produced by S, then surely we must find it unacceptable, because, for one, it is too narrow. Thus suppose two societies, S1 and S2, in both of which there is a certain collective practice Pr. Suppose that S1 and S2 have never interacted. Now according to this notion, the collective practice Pr of S1 cannot be produced by S1 because of the existence of Pr in S2. Intuitively this is not what we want; certainly Mauss and Fauconnet, followers of Durkheim, did not really want this result. For these sociologists, and their master, a striking and exciting fact for social science, perhaps the basis of the subject, was the existence of the same social institution in different, even widely dissimilar, societies. If we are trying to demarcate the province of social science by references to social facts, and to define social facts as phenomena, typically collective practices, produced by societies, then this account of production cannot satisfy us.

It is, moreover, not an intuitively sufficient condition for production by a society S that a practice Pr be peculiar to S. Suppose Pr were peculiar to S. This does not mean that it is, in any sense, produced by S. Perhaps it suggests strongly that it is in some sense so produced. Nonetheless we cannot deduce from its uniqueness its being brought about in such a way.

The latter criticism applies also to the idea (perhaps closer to what Mauss and Fauconnet had in mind) that we can give an account of production by society this way: 'had conforming members not lived in S...they might not have conformed to Pr'. Here practices shared by some societies are not ruled out any longer. Universal practices, or perhaps rather necessarily universal practices are. Yet here again the notion of production is not brought sufficiently straightforwardly into the picture. It does not follow from the fact that a practice only occurs in some societies, that that practice is in a clear or obvious sense produced by the societies in which it does occur. Nor, conversely, does it obviously follow that an in-some-sense-necessarily universal practice is not also a practice produced by particular societies in their members.

Now Mauss and Fauconnet, following Durkheim, certainly wished to rule out as social facts practices which are general in a group because they derive from the 'organic and psychic nature of the individual', that is, ^for humans in general (op.cit). Examples given are 'the phenomena of sensation, representation, reaction, or inhibition', and these authors write: '...they are the same, whatever group the individual belongs to. If an isolated individual were conceivable, one could say that they would be what ~~that~~^{they} are even outside any society'. Organico-psychic phenomena, then, are conceived of as absolutely unconnected with an individual's existence as a being in a society; they are the class of reflex 'actions', 'instinctive' tendencies, and the like. Now, one can presumably argue that if a practice is an organico-psychic one, it will be universal with regard to societies: every member of every human society will conform to it. Hence, if a practice is not universal, it will not be organic. However, universality does not seem to entail organicness. There could surely be practices which are universal, even in some sense necessarily universal, in human groups, but were not in this class of organic phenomena. It is even possible that isolated individuals would conform to them too. At this point it looks as if the core notion here is, in any case, 'not organico-psychic' rather

than 'not universal'. (Steven Lukes is overly kind to Mauss and Fauconnet when he credits them with 'isolating' the non-organicness of social facts by characterising them as ways of feeling, etc., 'which individuals would not have had "if they had lived in other human groups" ' (Lukes, p.14).)

Someone might now hazard a partial definition of 'produced by society' as follows: a practice Pr is produced by its social group or society if it is general (either in its own or in all social groups) yet it is not biologically or psychologically determined.

This again will not do; since this negative characterisation of the practice in terms of what it is not determined by simply does not tell us what it is determined by, thus leaving things too open. Perhaps the required generality of the practice may seem to close things up, making this a matter of 'sociological law', but in fact it does no such thing. Until we know why these practices are general we just do not know whether this is a matter of production by society, at any rate. So, according to Durkheim's basic conception, we just do not know whether this is or is not a matter of sociological law. As examples which seem to counter this we can cite examples of 'means-end' regularities. Weber for one made special mention of this particular species of practice: one which involves a 'rational' adaptation of action to the end aimed at, and which one might expect to find was quite general in many, if not all, social groups.¹

The unsatisfactoriness in general of a negative characterisation of production by society like that in terms of organicness has, I hope, been demonstrated. Similarly, no reference to the incidence of a practice across societies seems to work. I take it, then, that the most satisfactory and maximally general account of a practice's production by a social group will be the straightforwardly causal one presented at the beginning of this section.

¹ Thus Weber: 'a uniformity of action may be said to be "determined by the exploitation of the opportunities of his situation in the self-interest of the actor" '. 'Economic action' is for Weber the prime example in this category. See The Theory of Social and Economic Organization, p.121, also pp.122-4.

However, this rough general account lacks explanatory value. In particular it gives us little idea of what kind of mechanisms of production we can expect. In what follows I develop a more specific account, an account of a species of production by a social group. It is, I contend, an important species, and perhaps, in spite of some obscurities in its nature, nonetheless the most intelligible kind.

iii. A 'special' account of production by a social group: examples

Some collective practices appear to be themselves productive of collective practices. In particular, consider collective attitudes of approval and disapproval.

Suppose someone were aware that in his social circle people collectively disapproved of boasting. He might well find this a good reason for not boasting, and hence not boast. This phenomenon would probably be widespread. That is, many people, knowing of the generalised, collective disapproval, would refrain from boasting, even though perhaps they themselves might like to indulge in a boast or two. Thus, there may arise a collective practice of refraining from boasting as a result of the collective attitude of disapproval.

In fact, as is well known, our 'socialisation' begins in childhood. Approvals and disapprovals are made manifest by parents, teachers, and peers at school. (Clearly, different groups may have different standards. So one may learn, say, to boast at home and in the playground but not in front of one's teachers.) At an early age a person may well come to act in accordance with certain collective attitudes rather automatically, so that the model of one's actions being directed by current collective practices in one's society may seem to falsify the facts of the case. However, as long as one is sensitive to one's present environment to some extent, as we may surely assume most people are, this may be considered to play some role in one's behaviour. Thus though one may have ceased to think of boasting as a real possibility, one may be aware that were one to suddenly think of this, the thought would have to be suppressed.

And this will surely act as an extra, if not perhaps currently essential, 'force' in the situation.

A lot of collective practices may be seen as standing ~~on~~ ^{or} 'duty-ascriptions' [^] 'permissions' of a kind. These permissions and ascriptions of duty may attach to distinct categories of persons or to people in the group in general. A collective understanding or state of 'agreement' that, say, people of 'equal' social status may chide each other publicly in a certain way, but an 'inferior' may never do this to a 'superior', ~~and this~~ obviously would provide a foundation for the persistence of a collective practice of chiding among people who are 'equals' and of some people's ('inferiors') refraining from chiding others ('superiors'). Should common knowledge of the 'basic' understanding be shaken, for instance, should people begin to suspect that a belief in the irrelevance of social status to interpersonal behaviour is gaining ground, the nature of 'chiding' practices themselves may be expected to change. Another example: a woman marries and is now 'a wife'. At some points in the history of what we call the institution of marriage a wife may be 'expected' to put her husband's career above her own; more strongly it may be considered her duty to do so, and so, she may do her duty. Durkheim expresses his consciousness of the 'productive force' of such collective practices when he writes:

'When I fulfil my obligations as brother, husband, or citizen, when I execute my contracts, I perform duties which are defined externally to myself in law and custom. Even if they conform to my own sentiments...I did not create them; I merely inherited them through education' (p.1, English).

Collective practices may be self-amplifying or self-perpetuating. Thus someone might notice that few people boast, and, though inclined on some occasions to boast about his achievements, may suppress this because he sees that literally 'it isn't done'. He may perhaps come from a culture where boasting is actually de rigeur, but the absence of boasting in his new environment may be sufficiently striking to give him pause: 'they don't seem to do that here' - 'its obvious that it isn't done'. He may not wish to stand out as a stranger, or more

generally to appear odd, and if this motivation is prevalent almost any well-entrenched collective practice will be self-perpetuating-if it needs to be. By this I mean that some collective practices will also be a matter of spontaneous impulses or biological or natural compulsion, and these hardly require special motivations in order to be repeated. Thus humans will generally 'eat, sleep, think' without spur from their environment of collective practices, without encouragement from collective approvals or collective conformity.

On the basis of examples such as the foregoing we may isolate the idea of a particular species of production by a social group. We are concerned at the outset with the production of collective practices, and our idea was to see how these might be produced by an inherent feature of a social group. A clear type of such a case would be that in which one collective practice of G is produced by another such practice. We may now present our model of production by society. The root idea is this:-

a collective practice, Pr, is produced, by a social group, S, if an inherent feature of S provides one of the reasons for which those who conform to Pr do conform.

We elaborate on this idea shortly. From here on we shall concentrate on one instantiation of this model: where the collective practice is produced in the relevant way by something of the same kind: a collective practice .

iv. A contrast case.

Before saying more about reasons as 'productive forces', I note a 'contrast' case which, though it involves two collective practices in a causal relationship, does not involve the special kind of production by society I mean to isolate. This indicates that the notion I wish to focus on is a 'special' one.

Consider the following case: Members of society S agree to treat their water with chemical C, believing that this will be universally beneficial. The unintended result is that everyone in S takes to loudly and explicitly fantasising about asparagus all the time; the chemical used, unknown to members of S, had these special fantasy producing properties. It thus becomes a collective practice in S to fantasise about asparagus. Now the society was not,

if you like, the immediate cause of the change in collective practices. The change in the water, due to the drug, was what really did it. Nonetheless, at least this is true: an inherent feature of the society S, its collective practice of treating the water with C, is part of the causal history of a change in the collective practices of S. Thus it seems that one might want to say that - in a sense at least - the change of practice here was produced by the society, S, in which it took place.

Introduction of this case may recall Karl Popper's claim that 'the main task of the social sciences is the task of analysing the unintended social repercussions of intentional human actions'.¹ Popper notes the suggestion of Polanyi that it was Marx who first conceived of social theory as the study of the unwanted social repercussions of nearly all our actions.² The above statements of the views of Marx and Popper seem to me quite unhelpful to our own enterprise of saying in an illuminating, maximally general way what a social science is. To make much sense, these statements must, I imagine, be taken as follows: social theory studies social phenomena. Many of the latter are (surprisingly perhaps) merely the unintended consequences of human actions. Such a statement does not appear to be aimed at, or to give, any account of the social nature of a phenomenon. It most certainly does not appear to be intended as or to be suitable as an account of production by a social group.

Our asparagus fantasy case however does stand as a case of something intuitively speaking produced by a society - of a collective practice, indeed, so, produced. It indicates that an account of production by society broader than the one I give could have application to real cases.

I should be happy to agree that an investigation and account of the different kinds of 'production by social groups' that there are, could be a

¹ 'The Autonomy of Sociology' in Mill, ed. Schneewind, p.437. Extract from The Open Society and Its Enemies, Vol.II, Ch.14.

² (Ibid p.433)

useful piece of work. Meanwhile, whatever the results of such an enquiry, I claim that the kind of production by a social group that my account will capture is an important and clearly distinguishable kind.

Once this notion is further clarified, I hope to have presented a totally unmysterious account of a phenomenon's inhering in and being produced by a society or social group. This will I hope, lead to an unmysterious, while recognisably Durkheimian, notion of a social phenomenon.

v. Reasons as productive forces.

In our special account of the production by a social group of a collective practice we are giving an important role to 'reasons'. To recap, we are saying that a social group S produces a collective^{practive}/Pr if

an inherent feature of S provides one of the reasons

for which each member who conforms to Pr does conform.

In fact, we shall later argue that a certain kind of reason, which we shall call a 'basic reason', is involved in the phenomena in which we are interested. Leaving aside this complication for now, I say something here about how 'reasons' play their role as 'productive forces', linking individuals and social groups.

On one interpretation of such statements, saying something of the form:

'(the fact that) F was part of X's reason for Q-ing' is to say at least:

(1) F

(2) X's belief that F caused him to Q (or was part of the cause of his Q-ing).

Though F (or the fact that F) may be referred to ('loosely' if you like) as a 'reason' here, one's understanding is not that F caused X's Q-ing, but rather

X's belief that F did.¹

Here, however, we are interested in beliefs (as reasons) which have a particular causal relation to the 'external' world. We are interested in situations where a social group S's having an inherent feature F provides one of a person X's reasons for Q-ing. We are interested, in other words, in situations in which, at least:

(1) a social group S has inherent feature F

(2) S's having F is all or part of the cause of X's belief that S has F.

(3) X's belief that S has F causes him to Q.

But if (2) and (3) are both true, then S's having F is, albeit indirectly, a cause of X's Q-ing. If Q-ing is conforming to collective practice Pr, if F is a collective practice Pr' of S, and if we can in (2) and (3) substitute for 'X', salva veritate, names for most Pr-conforming members of S, in turn, then S's having collective practice Pr' is, surely, a cause of the existence of collective practice Pr.

Pr' is not, indeed, necessarily a cause of Pr's collectivity, but is rather a cause of each of the conforming instances (or dispositions to conformity) which make it the case that Pr is a general practice of S. Being a cause of Pr qua general practice we must allow that it is part cause at least of Pr's

¹ On 'reasons as causes' there is, of course, a large philosophical literature. Donald Davidson is, perhaps, the best known proponent of a view of 'reasons as causes', see his 'Actions, Reasons, and Causes'. Opponents of some ideas of reasons as causes include Peter Winch (ISS) and A.I. Melden (Free Action). There is no room to take up the debated issues here. It must suffice for me to take a certain stand, which I hope will have sufficient plausibility in itself for my readers. Supposing beliefs are argued not to be causes, I might perhaps protect my picture of production by society as long as some quasi-casual role is accepted for them.

Saying that the ascription of a reason is (or may be) a casual claim is not to say anything about whether for instance one has any special authority with respect to one's own reasons, or indeed anything as to the nature and possibility of confirmation of such claims. The truth values of these claims may be hard to retrieve, but this doesn't make it clear that they cannot have such values.

existence as a collective practice. For by definition Pr could not be collective were it not general.

We note that as long as a feature F of the relevant kind provides a certain number (not necessarily all) of the instances of conformity to Pr, S's having F is going to be a cause of Pr's generality, and hence of Pr's being a collective practice of S. It is possible, in other words, that some number of people conform to Pr just because they feel like it, or in ignorance of any collective practice which would provide a reason for conformity, and so on, yet that Pr still be ('as a whole') produced by S in the relevant way.

To sum up, then, we have noted that a collective practice Pr of a social group S will be produced at least in part by a collective practice Pr' of S if the fact that S has Pr' provides one of the reasons underlying conformity to Pr, in the case of at least a certain critical number of conforming members. When this is the case it is intelligible to claim that the social group S itself produces p, or so we have already argued (see section 2(ii), 2, of this chapter).

vi. 'Basic reasons'.

I believe that a preferable account of 'production by a social group' along the lines sketched above will involve the requirement that a social group provide a particular kind of reason only, and will specify that kind. This kind of reason will appropriately be given a name like 'basic reason', and I shall here call the kind of reasons that I think must be involved basic reasons.¹

My belief results primarily from a consideration of certain cases which intuitively seem not to involve the kind of thing we should wish to call 'production by a social group' and which contrast with other cases which are

¹ Cp.D.Davidson's 'primary reasons'. The concept of a basic reason here will be different from and more complex than Davidson's (I think). Davidson (in 'Actions, Reasons and Causes') never says what kind of reasons his 'primary reasons' contrast with and hence never really explains what a primary reason is.

more clearly acceptable.

Consider first an 'acceptable case'. I look out of my window and see that 'everyone' is wearing a raincoat today. I was going to wear my fur coat, but I don't want to appear overdressed or in general out-of-step with the rest. I therefore don my own raincoat and go out in it. Here my donning my raincoat appears to be brought about in the right way for it to be said to be 'produced by society', or more specifically, by a collective practice of a social group (say the people of my village).

Consider now the following case in David Lewis's Convention(pp.119-120).

'I may wear my raincoat because other do, thinking that they are probably wearing raincoats because they have heard a forecast of rain...their wearing of raincoats makes me wear one because I take their actions as evidence: evidence of their expectations of rain and therefore (indirectly, through my standing beliefs about the likely causes of their expectations) evidence that it is likely to rain. It might happen one day that everyone in town wears his raincoat because he sees the others wearing theirs and infers - reasonably enough, perhaps, but falsely - that they probably heard a forecast of rain.'

In the last situation described by Lewis it is clear that we could say that each raincoat wearer wears his raincoat because the others wear theirs. We can also quite properly say of one of these raincoat wearers that his reason for wearing (or perhaps better, continuing to wear) his raincoat was that the others were all wearing theirs, and we can also say that the fact that others are wearing raincoats is each one's reason for wearing his. In other words, it seems that we can use the language of reasons here perfectly well.

Now it is also clear that we can equally properly say here that the raincoat wearers do not wear their raincoats because the others do! We may avoid the appearance of paradox by saying, rather: they do not wear their raincoats just because the others do. Or we might use a word like 'real': their real reason for wearing raincoats is that they expect rain.(This of course does not explain what thing or things 'real' reasons may be).

Now, as far as Lewis's particular case has been described, it need not be a candidate for production by a social group as we have so far defined this. Even given that the population in question is a social group, it is not clear that any of the raincoat wearers wear their raincoats because something is a collective practice. In particular they may have no views on what is group common knowledge in the group, or perhaps believe that there is none. Thus, it might be argued, if this case seems intuitively to be something we do not want to count as produced by a social group (or as a Durkheimian social phenomenon) our definition already takes care of it. However, we can, I think, graft the missing feature on to this type of case, and still have a problem.

Suppose that on a warm day everyone dons his raincoat 'just in case', and then looks around to see whether or not he needs to go on wearing it or can carry it and be less wrapped up. Each one notes that everyone else wears a raincoat as well. He also reasonably takes this universal raincoat wearing to be a collective (albeit temporary) practice in his village. Each then reasons (perhaps wrongly): there is this collective practice, therefore no one has seen fit to remark to his neighbour that he, for one, is wearing his raincoat because he has a chill, but the forecast was rather good, wasn't it?. No one has, in other words, challenged another's raincoat wearing while explaining away his own. The fact that the practice is collective and is a relatively stable practice is taken as confirmation, then, that everyone has good reason to believe that there will be rain or that there might be. So everyone continues to wear his raincoat in part because raincoat-wearing-today is a collective practice in the village. (Compare those cases of joke-pointing to the top of a building, say, to see how many people you can get to look up.) Once again, we would, I think, want to say that these people do not keep on wearing their raincoats because there is a collective practice of so doing, but rather because they now more confidently expect rain. Or we might say, once again, 'their real reason for wearing raincoats is their confident expectation of rain'.

I suggest that it is intuitively plausible to rule that cases like Lewis's case, even those involving a belief that a certain practice is collective, are not cases of a practice produced by a social group. At the least cases need distinguishing here, being of importantly different types. As is our practice, we shall attempt to restrict the phenomenon under consideration so it involves only the non-Lewis type.

I take it that our intuitive understanding of how beliefs-as-reasons result in action (including change of belief) involves the assumption that beliefs are inefficacious without a certain kind of related want (and vice versa). For the sake of discussion the type of reasons thought essential to produce actions might be schematized along the lines of a 'deduction' as follows:

(a) I want to do A if C obtains,

(b) C obtains,

So (c) Let me do A.

It is the combination of the relevantly related desire (represented by (a)) and belief (represented by (b)) which is presumed necessary for the action or attempted action (represented by (c)) to result. Let us say that 'premisses' (a) and (b) represent a reason pair (for action A). Such a pair consists of a desire to do A under certain conditions, and a belief that those conditions obtain. We may cite either member of a reason pair as the reason for the action i.e. both wants and beliefs may be reasons, but for an agent's action to result from either member of the pair the other must be present. Thus only reason pairs result in action, intuitively.

Consider, now, the course of each man's likely reasoning in the Lewis type of case. There is first the following reasoning (A):

A (i) Everyone in G is wearing a raincoat and this is common knowledge etc.

so (ii) Probably most people in G have good reason to expect rain

so (iii) Probably there is good reason to expect rain

so (iv) Probably it will rain.

That is, each one's belief that (i) leads him, via this reasoning, to a belief that (iv).

Now we take it that in Lewis's case the reasoners have only the following relevant desire: to wear a raincoat if and only if it is raining or soon going to rain. (Cp. Lewis, last para p.119). We assume that in that case each person's belief in the truth of A(i) above (that everyone else is wearing a raincoat etc.) is not of the right type to lead directly to raincoat wearing, but one belief in A will: the belief in the truth of A(iv), that it is going to rain (probably). Thus the 'reason pair' resulting in action, and its result, would in this case be schematized as follows:

B (i) I want to wear a raincoat if it rains

(ii) (Probably) it will rain soon

(iii) Let me wear a raincoat.

In this case, then, the fact that everyone else wears a raincoat does not fill out a reason pair for each one's wearing a raincoat, just by people's awareness of it.

Can the simple appeal to reason pairs do our job? That is, may we say that (for production by a social group) the belief that S has F must be one of an effective, i.e. actually operative, reason pair? Though this may rule out certain cases we want excluded - like our Lewis-like case, there is at least one reason why it will not quite do as it stands for all such cases. Lewis notes of his case (Lewis, p.119) 'This...(does not involve)...preferences that are conditional on others' actions'. Now consider a case somewhat different from the Lewis case. The raincoat wearers here have the following desire call it 'D': to wear a raincoat if they are in a group in which raincoat wearing is a current collective practice. Now assume that D might itself be a result of beliefs about the reasons which usually underlie raincoat wearing and have the belief in the likelihood that such a collective practice is evidence that there will be rain. Now, given that it exists, D will constitute with the relevant belief a basic reason pair. This pair could, it seems, be operative in a situation superficially like Lewis's.

If such cases are possible, however, it seems that we must fix things up

so that they too are ruled out. Or so my intuitions suggest.

Now, if desire D were itself quite independent of any other want or belief it surely could serve as the right type of want, and its correlative belief as the right type of belief. If D 'flows from' a preference for conformity for its own sake, say, from a desire to conform to any collective practice of one's social group, things will hardly be altered, intuitively.

We can count a D which flowed from such a preference as essentially 'independent' for our purposes here: we can see an overall preference for conformity of the type proposed as amounting to the same thing as a desire to ϕ , if there is a collective practice of ϕ -ing, and/or to ψ , if there is a collective practice of ψ -ing, and so on. Hence we might think that in this case D is as it were, 'embedded in' as much as 'flowing from' the overall preference for conformity.

Now consider the following situation. I desire to avoid painful experiences, and one painful experience is a hostile glance. So I wish to avoid hostile glances. I believe that conformity to collective practices avoids hostile glances. Therefore I desire to conform to such practices. In other words, I have a universal preference for conformity, but it derives from another, if you like, more basic or deeper desire. I imagine that insofar as there are universal preference for conformity they will in fact often be derived in this manner. Now assume that my desire D is 'arrived at' in this manner, and I know of the collective practice of wearing raincoats, so I wear one. This I think we should want to count as a case of production by a social group. If we do not we will exclude an important and intuitively acceptable class of practices from our category.

I suggest then the following rather ad hoc way of defining a class of reasons (which we will call basic reasons) which go towards allowing a social group to produce a collective practice in the relevant way.

Let us say that a belief that p constitutes a basic reason for X's Q-ing if it fits into a reason pair of the kind given in the following schema:

- (1) I desire to Q, if p, (for no reason) (because my Q-ing when p will cause desirable situation S (or prevent undesirable situation S'))
- (2) p
- (3) Let me Q.

Here the bracketing in (1) indicates alternative options one of which must be fulfilled. We do not 'bracket' as an optional reason for the desire 'since the truth of p is good evidence for..., and my Q-ing when...will cause desirable situation S or avoid undesirable situation S'. Thus the unacceptable situations are ruled out.

I hope that this schema indicates some intuitive justification for saying that only beliefs in collective practices which are 'basic reasons' contribute to the 'production by social groups'. To mention just one feature distinguishing cases here, in the acceptable cases, agents see an inherent feature of their group as itself requiring action: it is not merely 'doing duty for' another kind of feature of the world, one which is not a feature of a social group at all.

In conclusion, even if our account of 'basic reasons' turns out to be inadequate, I trust that the desirability of some account of them has been demonstrated in this section. Our final account of a specific type of production by society can now be stated, assuming our account of basic reasons:

a collective practice Pr is produced by a social group S

if a collective practice of S provides a basic reason for

which conforming members conform to Pr.

vii. The reality of our account.

Could there be real cases of production by social groups precisely as we have defined it? A worry might occur to someone, as follows: it's all very well to accept that group common knowledge of general practices occurs. But can we

accept that people ever, let alone often, do things because(in part) there is group common knowledge?

Do people ever even do things because, say (let 'K' = 'knows', 'E' = 'everyone') KEKEKEKEKEKEp, rather than because KEKEp, or some very low order of other related knowledge? The real issue here is about the likelihood of the prevalence in p of knowledge that it is group common knowledge in p that--

Now, we should not consider it a necessary condition of having the belief characterised in the above paragraph that one be able to verbalise it precisely as above. That is, one need have no word or phrase equivalent to 'group common knowledge'. Let us consider, then, the positive conditions under which you might be said to have the knowledge of (and/or belief in) group common knowledge.

This doesn't mean that you have actually 'run up' the chain of bits of higher order knowledge; no one can run up it to its furthest point, since it has none.

If anything, it means, rather, that you know that there is such an infinite chain, but what must we know to have such knowledge?. It seems that, as Lewis suggests (pp 60-61, Convention) anyone who is party to a bit of group common knowledge will know that he is. If he had evidence from which he can conclude that everyone in p knows that p, that everyone in p knows that everyone in p knows that p (and so on, ad inf), then he surely has evidence from which to conclude that: KEp, KEKEp, and so on, ad infinitum.

It seems plausible to claim, then, that in order to have knowledge that there is group common knowledge, of some fact, one may only have to be a party to that knowledge in the sense that one is one of the members of the group in which that knowledge occurs. Perhaps one has to have the idea of an infinite series, and perhaps one could be a party to group common knowledge without this idea. Nonetheless it seems right to ascribe an implicit concept of such a series to 'normal adult humans', and if we can do this, then it seems that we can be entitled to ascribe beliefs in common knowledge and/or knowledge of common knowledge to such beings.

What then of preferences conditional on the existence of group common knowledge? It seems pretty clear that if someone wants to ϕ if everyone in G knows that everyone in G knows that p, then he will most likely want to ϕ if it is group common knowledge in G that p. I do not think it unrealistic to claim that many people if not all have preferences of the latter kind. (It would, presumably, be misguided to expect some specific 'phenomenology' of such wants). Such a preference clearly could arise once one is confronted with certain items of group common knowledge. Thus suppose that it comes to be group common knowledge in G that most members of G feel uncomfortable if confronted with another member's unhappiness. Let 'p' stand for the stressed sentence. Now, any member of G may find the fact that p a strong reason in itself for keeping silent about his miseries. He may well want to avoid anything that is embarrassing to his fellows, but he may wish to shun also, and in particular, whatever is commonly known to cause embarrassment. Thus the collectivity of a certain tendency to embarrassment, for instance, may surely add to someone's motives for not causing the embarrassment. Finally, I imagine that many people, of a conformist sort, may have a standing preference to conform to the collective practices of groups that they are members of (perhaps ceteris paribus).

I conclude, then, that doing things because there is group common knowledge that...or because there is a collective practice of a certain sort, is both a perfectly humanly possible phenomenon, and indeed a likely one.

6. A Durkheimian notion of social phenomenon.

i. The definition.

We recall Durkheim's own basic conception of a social fact, as we elicited it:

a phenomenon p is a social fact if and only if it inheres
in and is produced by a social group.

Building on our understanding of a particular range of phenomena which are social facts according to Durkheim's conception, we can now propose a more specific 'neo-Durkheimian' conception as follows:

a phenomenon p is a Durkheimian social phenomenon if and only if it is a collective practice Pr of a social group G such that one of the basic reasons for each conforming member's conformity to Pr is provided by some particular collective practice of G, Pr' .

ii. Some examples.

We may clarify some aspects of our definition by examples. First the example of conformists who wish to conform to all collective practices of their group, ceteris paribus, indicates that Pr' may be the very same practice as Pr . Thus collective practices can be self-perpetuating: their existence as collective practices may lead to the continuance of the practice.

Second, the following should, it seems to me, be counted as an example of a 'Dsp' (Durkheimian social phenomenon) on our account of it. Most people in a social group G take tea at 5, and this is a collective practice. Some do so because there is a collective expectation of tea-taking at 5; some do so because there is precisely the collective practice of taking tea at 5; others do so because they know of a collective tendency to disapprove of taking tea earlier or later. This example indicates how the definition should be read: it does not stipulate that there is one and only one practice Pr' which produces the practice Pr . It is rather that for each conforming individual there is some collective practice Pr' such that Pr' provides the reason for that individual's conformity to Pr . The simplest kind of example of a Dsp would look, rather, like this: most people in G take tea at 5 and this is a collective practice; each one conforms at least in part because this behaviour is expected and moreover the expectation is collective. Allowing for the more complex kind of example makes the concept of a Dsp more capacious, and probably more realistic, and I cannot see anything against it.

As an example of the cutting edge of our notion of a Dsp we might consider

the question of so-called 'crowd-phenomena'. Durkheim wanted some such phenomena to be included among his social facts, thus he writes:

'...in a crowd(*assemblée*) the great movements of enthusiasm, indignation, pity, which occur, do not originate in any one particular consciousness...' (p6.Fr)

On the other hand, in 'Sociologie', Fauconnet and Mauss argued that these were, rather, borderline phenomena. This is largely because they proposed to define sociology as the science of ^{social facts} and social facts as 'institutions'. (They define 'institution' as 'an ensemble of acts or ideas which are already instituted, which individuals find confronting them and which impose themselves on individuals'). Crowd phenomena, they write, are found in 'societies without institutions': social aggregates which are very unstable or ephemeral, like crowds.

What, then, does our notion say of crowd phenomena? This will, clearly, depend on the precise nature of the phenomena in question. One is inclined to talk in terms of 'influence' in these cases: each person in the crowd is influenced by the noise, enthusiasm, general behaviour of the rest, but does this mean that each person acts for certain reasons? Or that he acts for causes other than reasons? Is he perhaps 'subtly influenced' in ways bypassing reasons and reasoning altogether? Does he just start shouting, say, not because of his belief that others are openly shouting, not because of his desire to conform or to take advantage of a collective letting off of steam, and so on? Suppose the explanation of a given crowd phenomenon (say the resonance of indignation, its growth to fever pitch) is the desire to conform to open expressions of anger, or to exploit the opportunity they provide, then we will have a Dsp, or something very like one. I add the qualification, since there is the question whether such a crowd does in fact form a social group. A given group or even a small society might form a crowd: but must a crowd, exhibiting phenomena like the one supposed, form a social group? This cannot be decided by our concept of a Dsp, but insofar

as crowds may either as such or under some other description be social groups, it seems clear that so-called 'crowd phenomena' may sometimes be Durkheimian social facts under our definition. The apparent vagueness of our answer only indicates the real vagueness in the notion of crowd phenomena - once these become properly understood their status as Durkheimian social phenomena becomes clear.

iii. Conclusions.

Our Durkheimian notion marks a clear distinction among an otherwise broad range of phenomena, the collective practices - as we have called them - of a social group. Though distinctions can, surely, usefully be made among kinds of Dsps themselves, a distinctive kind of feature of social groups is captured by our notion, together with a distinctive kind of 'production by society'. Our notion also provides a definition for the term 'social practice' which makes sense intuitively.

We noted earlier the importance of collective practices in general in providing a 'shared world' for societal and group members. Such practices clearly enable the members of groups to interact more smoothly and more efficiently - myriad items of shared knowledge may be left unstated in any interaction among group members.¹

Many of these collective practices will, evidently, be Durkheimian social phenomena. Those that are Durkheimian social phenomena may be contrasted, in particular, with 'mere' means-ends regularities which are also collective, and with 'organo-psychic' phenomena like blinking and (perhaps) having the concept of a 'material object'. (Exactly which collective practices are also Dsps will be an empirical matter).

collective practices in general seem to me - pace Weber and Durkheim - already to be highly social phenomena, intuitively, but of the kinds of collective practice just mentioned, the Durkheimian social phenomena are surely

¹ Experimental attempts to uncover such items in a given social circle are reported in Harold Garfinkel's Studies in Ethnomethodology.

the most highly social from an intuitive point of view.

There is surely also going to be something more intriguing about Durkheimian social phenomena from an explanatory point of view. Once we know something is a 'mere' means-end regularity, collective or not, it is pretty clear why it occurs. Similarly for organico-psychic phenomena: they could not have been otherwise, for human beings, but though the concept of a Durkheimian social fact encapsulated an explanation, these facts still 'hang in the air' in an intriguing way.

Durkheim's advice to seek the explanation of one social fact in terms of others¹ was based, I believe, largely, on the intuition that social phenomena may be assumed to cohere together, to mesh well, and thus to mutually support each other. Perhaps for the reason that that was the only 'external' support that they could have.

Whether or not our neo-Durkheimian notion captures every intuition Durkheim had on social facts, I am sure that it captures at least some of them, and the advantage of drawing lines somewhat narrowly along the way - as when we restricted the kind of production by society that was to be involved - is that of a larger measure of precision, and less mystification.

Durkheim has had many detractors in recent years,² and there is indeed, in his writing, an awkward mix of a kind of crude Hegelianism with a brand of (also crude) positivism: the talk of group minds and thoughts, combined with the insistence that social facts are 'things'. But there surely is much gold in those hills. I hope that in this chapter I have mined some of it.

¹ See especially the Rules, Ch.4; social anthropologist Rodney Needham alludes approvingly to Durkheim's advice in his Structure and Sentiment, p.126. I have no room to go into the question of the explanation of social facts in any detail here.

² See, for instance, Filmer, Philipson, Silverman and Walsh, who see Durkheim as their arch-enemy in New Directions in Sociological Theory. (1972)

Note: This chapter, from pp. 206 to pp 298 (as marked), was not read by the examiners for the D.Phil., because of the word limit for theses in Philosophy.

CHAPTER FIVE

ABOUT CONVENTIONS

1. Introduction:

A suggestion about conventions.

It might be suggested that, if we wish perspicuously to characterise the social sciences, we would do better to use the familiar notion of a social convention, than to use a new fabrication like our neo-Durkheimian notion of a social phenomenon. In this chapter I assess this claim. An account of convention is needed, and to give this is no easy task. A large part of the chapter is spent evaluating the analysis in David K. Lewis's Convention. This, and some suggestions by some of Lewis's critics, are found wanting in various respects. I myself am drawn to an account different from all of these, but close to Max Weber's. On the basis of this work the suggestion that 'convention' is superior to our Durkheimian notion is assessed.

The phrase 'social convention' is a member of a much-used but obscure cluster of terms. The importance of 'social rules', 'social conventions', 'norms', and the like for the social scientist is often stressed both by social scientists themselves and by philosophers. Lukes lists 'institutionalised norms' among Durkheim's social facts, and, for all his stress on 'social action', Max Weber's

'fundamental sociological concepts' do include among others that of a convention.¹ Nonetheless, talk of 'conventions' and the like is not entirely perspicuous. A good place to assure oneself of the proliferation of senses given to the word 'norm' in sociology is Jack Gibbs's 'Norms: the problem of definition and classification'(1965). Different definitions of 'convention' as of the other terms, have been given, and thus we need to decide on our own preferred account (or accounts).

Now to start with, one might think, there is our ordinary everyday, nontechnical notion of convention, and this at least will be common ground to all thinkers on the subject. However, it is by no means easy to find a satisfactory analysis of that everyday notion, and in order to compare the concept of a convention with that of a social fact an analysis - as opposed to a mere set of examples - would be the best thing to have on hand. I shall therefore start by considering in detail an intricate and imposing account of conventions that has recently been proposed by David Lewis, in his book Convention, and in a more recent paper 'Languages and Language'.

Among other things, Lewis claims to have captured a concept of convention we all share; the account of convention that he gives is not supposed to be merely stipulative. Thus Lewis:

'...This book is my attempt at an analysis. I hope it is an analysis of our common, established concept of convention... At least, insofar as I had a concept of convention before I thought twice, this is either it or its legitimate heir, and what I call convention is an important phenomenon under any name.' (p.3--unless otherwise stated page numbers to quotations from Lewis refer to Lewis's Convention)

As an account of our everyday notion, Lewis's is open to a fair amount of criticism. There are also a number of internal criticisms to be made of his

¹ For Max Weber on convention, see the Theory of Social and Economic Organisation, pp.127-8. References can be multiplied. See, for instance, Harre and Secord, The Explanation of Social Behaviour, Dorothy Emmet, Rules, Roles and Relations, and recall the claims of Peter Winch in The Idea of Social Science, on the centrality of socially established rules.

account of what he claims, is, in any case, 'an important phenomenon under any name'. Nonetheless I believe that Lewis's analysis is a good place to start.

We shall consider carefully how Lewis's account stands both as an analysis of 'convention' as we (or at least I) mean the term, and as a putatively precise account of an important convention-like phenomenon.

2. David Lewis on convention: overview.

i. Conventions as (L-) regularities.

Lewis's analysandum in Convention is of the following form:

'a regularity R in the behaviour of members of a population P when they are agents in a recurrent situation S is a convention if and only if...'

At the outset, then, we note that Lewis sees conventions as regularities in the behaviour of agents. I take it that the kind of regularity Lewis is concerned with involves a pair of a situation type S (e.g. meeting a stranger) and an action type A (e.g. raising one's hat). Lewis is, more particularly, concerned with cases involving a 'recurrent situation S'. Let us say that a situation S is recurrent in P if more than one token or instance of the situation type S occurs among members of P (for instance if members of P often meet strangers). It seems, then, that for Lewis there will be 'a regularity R in the behaviour of members of P when they are agents in a recurrent situation S' if and only if there is in P a recurrent situation S such that in (almost) any instance of S among members of P, (almost) everyone does A. It is with regularities of precisely this sort that Lewis seems to be concerned.¹ Let us say that when, and only when, there is a regularity of this sort in a population P, we have an L-regularity in P.

¹ In the more recent 'Languages and Language' Lewis has changed the form of his definiendum so that it now omits explicit reference to situation type S and hence to S's recurrence. I gather from Lewis, however, (in conversation) that he did not intend there to be a change from the sense of the Convention definiendum.

Now, if Lewisian conventions are all L-regularities then conventions (according to our everyday notion) can exist which are not Lewisian conventions. First, conventions may exist with respect to 'situation types' which can have only one instance and hence cannot be recurrent. Second, though the situation could in principle recur, instances of it may in fact never occur, let alone recur; the situation need not, in particular, be recurrent. Third, it seems that the relevant action type A need not in fact have instances in P even when the relevant situation type S does occur, whether S is recurrent or can be, and is, only instantiated once- yet there may still be a convention in P to do A in S. Illustrations of all these points follow.

Example 1: Before the battle, you and I decide that you will wave your blue scarf if and when our side loses this battle. Our side is defeated. Before you can think about waving your blue scarf, you are stunned by a blow, carried off and never seen again.

Now, I think it would be natural here to say that you and I had adopted a certain temporary convention, namely to use a blue scarf to represent bad news on this occasion. Surely we also had the convention. (I say to a third party 'why don't you use the convention Joe and I once had?'). So, it seems, here is a convention which does not relate to an even potentially recurrent situation, but to a situation type which can have only one instance, viz. our side's losing this battle. Moreover here no one ever conforms to the convention in the sense of acting in accordance with it in the relevant situation, even though the situation does occur.

Example 2: Members of C, a small island community, agree to drive on the right though no one yet has a vehicle in which to drive. Then the cars do not arrive, but group members go on hoping, and discussing the matter in anticipation of the happy event. It turns out that the cars never arrive. Here S is a situation type which could recur, but never in fact occurs at all, and the relevant action A is never done.

Example 3: There are one hundred of us; we are going into battle. We agree that whenever anyone is wounded, he will wave one of his red gloves above his head, to alert the others. All of us are in fact wounded, but are too enfeebled to consider doing anything like waving our gloves about. No gloves are raised. Here, then, S is a situation type which does recur, but A is still never done. Yet surely we all had (and went on having) a certain convention with respect to S?

Variants of this case are possible, where physical inability to conform is not the whole issue. Example 3': As in 3 initially, and all of us are wounded, though not so badly. However the wounds were made of poison darts which produce hallucinations of well-being, till our energy is sapped. No one ever waves a red glove, though each one could have, had he so chosen. Example 3'': As in 3 initially, and all are wounded: some are wounded mortally, some lose their gloves, some are delirious and don't feel their wounds, some realise there is no point in waving his glove since everyone else is in such a bad state. Again, the relevant situation S occurs but A is never done.

It seems, then, not to be a necessary condition for the existence of a convention that there is a recurrent situation S such that in (almost) any instance of S among members of P, (almost) everyone does A. In other words, conventions need not be L-regularities.

The foregoing examples raise grave doubts, in fact, as to whether conventions are, in everyday thought, ever to be identified with L-regularities, that is, whether conventions are themselves ever L-regularities of certain sorts.

I return to this issue later, (Section 5(ii)). For now I note the following. Most naturally one talks of actions as involving 'conformity to' conventions, of doing things (or conforming to regularities) 'as a matter of convention' and perhaps 'by convention'. These may seem to be trivial facts about local locutions, but it is such facts which, I suppose, determine our intuitions on convention, and it is our intuitions on convention that are

~~at are~~ at issue just here.¹ Lewis's analysandum might best be thought of perhaps not as 'convention' but 'conventional regularity'. The latter is a phrase Lewis himself uses (see, for instance, p.43); to put things more clearly we might say this: Lewis can be regarded as proposing an analysis of the notion of an L-regularity-which-is-conformed-to-by-convention. A complex concept this, and clearly closely related in any event to that of convention; analysis of the former might well be expected to throw light on the latter, if indeed it is not an analysis of it, at one and the same time.

With this cautionary preamble, we proceed to examine what we shall speak of, for short, as Lewis's account of conventions.

ii. The basic preference structure underlying conventions: -
Lewis's final account.

Lewis's account of conventions centres around a special structure of preferences within a group, on which, he claims, the existence of conventions depends. His own preferred characterisation of this structure, however, is not altogether clear, since there appears to be an unnoticed shift in the course of his discussion (see below 3(iii)).

At the beginning of his book Convention, Lewis defines a class of situations involving a particular structure of preferences which he calls co-ordination problems. He wishes to claim, evidently, that the basic preference structure underlying conventions is that of a co-ordination problem. Roughly and in brief he intends to claim that conventions are established solutions to particular recurring co-ordination problems.

Lewis's account of co-ordination problems draws heavily on game-theoretical conceptions. (I shall say more about the theory of games and Lewis's use of it shortly). In spite of the ingenuity and care involved in the account,

¹ See, however, the rather strange use of 'regularity' (Convention, p.51) in which it looks as if 'convention' and 'regularity' (of a certain kind) are assumed to be intersubstitutable. The resulting sentence reads extremely oddly, to my mind, and well-nigh disproves the assumption: 'the convention of a language...is a regularity restricting.etc.' (my stress).

it is still not entirely perspicuous, as I shall show in the next section. Moreover, Lewis's final stated account of the basic preference structure underlying conventions in Convention, and his account in the later 'Languages and Language' also, are fairly clearly different from that of the co-ordination problems he painstakingly characterised earlier! These two later versions differ, moreover, among themselves. So: what is Lewis's claim about the basic preference structure underlying conventions?. We can, apparently, avail ourselves of various options. We shall consider 'co-ordination problems' shortly. Meanwhile, the following captures fairly neatly the most recent account ('in Languages and Language' ^{and} / what is common to this and the final account in Convention. Let us label this, then, 'Lewis's final account':

Underlying a convention is a problem situation in which there are two or more agents, and two or more mutually exclusive ways each might act (doing A or doing B, say), such that given that all but one of the agents does A(B) each one prefers that the remaining one (who may be himself) does A(B); each agent has to decide what to do.¹

¹ Talk of 'doing A', 'doing B' (etc) as the options for each agent here is a useful simplification. In order that it square with the possibilities in actual cases 'it must be understood that A may stand for an unnaturally complex action-description' (Lewis, p.44, q.v.). For, instance, we are cut off in the middle of our mutually intriguing telephone conversation; I made the original call. For us 'doing A' might be 'calling back if and only if you were the caller'; 'doing B' might be 'waiting to be called if and only if you were the original caller'. Thus if I call back and you wait for the call, this may be an instance of our both 'doing A'. At other times A and B may stand for simple natural action descriptions like 'driving on the right', 'driving on the left'. (Cp. Lewis pp.10-11 on this point). Looking at the situation in terms of game-theory makes this particular point rather clear - see the last section of this overview.

iii. Lewis's argument as a whole.

I see the presentation in Lewis's book as a whole as unfolding the following line of argument, which it will be useful to present in summary here.

1. There is an important class of problem situations which have a common structure; the structure is precisely definable in terms of game-theoretical 'payoff' matrices of a certain specifiable type.¹ These situations will be called 'co-ordination problems'.²

2. The main distinctive feature of the matrix structure of co-ordination problems is virtually deducible from the existence of obvious properties of the situation in question.³ The distinctive feature at issue is the existence of two or more 'proper co-ordination equilibria'.⁴ In such situations, once each agent has reason to believe that everyone else will do their part in a certain p.c.e., he will have a decisive reason to do his part (if he is rational).

3. By force of precedent, and other things also, doing one's part in such a combination of actions can become a generalised regularity in behaviour in a human population; each member can develop a standing disposition to do his part whenever the relevant situation arises.

4. Such generalised regularities occur widely in human populations, and it will often be 'common knowledge'⁵ that they and their underlying preference structure obtain.

¹ The final section of this overview provides a note on game theory.

² See section 3, below, passim.

³ See Section 3(v) in which I dispute this especially.

⁴ See Section 3, below, (ii)(A) and (vi).

⁵ See our Section 4 (iii).

5. The existence of such regularities in behaviour together with the common knowledge is a precondition of language use,¹ and

6. 'conventions', as we ordinarily use the word, are regularities-plus-common-knowledge of the specified type.²

7. '...it is redundant to speak of an arbitrary convention' and the arbitrariness of conventions is (with a qualification concerning common knowledge) their instances reaching one of at least two proper co-ordination equilibria in the situation underlying them.

Whether or not we agree with 5,6, or 7, the other propositions are interesting both in general and for our purposes. For we should be interested not only in whether Lewis has captured our everyday concept of convention, but also in his description of what may be a new concept of his own, developed to capture 'an important phenomenon under any name'. Central to this description is his definition of co-ordination problems.

We shall in the next section turn to the basis of Lewis's account of the preference-structure underlying conventions, which is of independent interest in itself - his work on co-ordination problems. (We shall in fact be critical of the details of this discussion).

Before plunging into details, it will be useful to set out Lewis's full account of conventions as presented in 'Languages and Language' (some details are omitted):

¹ See our Section 4, (vii).

² See our Sections 4 and 5, passim.

'a regularity R, in action or in action and belief, is a convention in a population P if and only if, within P, the following six conditions hold. (Or at least they almost hold. A few exceptions to the "everyone"s can be tolerated.)

- (1) Everyone conforms to R
- (2) Everyone believes that the others conform to R.
- (3) The belief that the others conform to R gives everyone a good and decisive reason to conform to R himself...
- (4) There is a general preference for general conformity to R rather than conformity by all but anyone. (This is not to deny that some state of widespread nonconformity to R might be even more preferred)...
- (5) R is not the only possible regularity meeting the last two conditions...
- (6) The various facts listed in conditions (1) to (5) are matters of common (or mutual) knowledge: they are known to everyone, it is known to everyone that they are known to everyone, and so on'. (pp.5-6)¹

His final account in Convention (p.78) is only slightly different from this; in both the basic structure of preferences involved is as given in my account above, with one slight divergence which will be mentioned shortly.²

We have now seen where Lewis ends up. Before attempting to assess his 'final' version(s), we consider how he gets there.

¹ Lewis also mentions an alternative account of this condition in terms of disbelief which we discussed earlier.

² I shall not discuss at any length the change from 'regularity in action' to '... or in action and belief', between Convention and 'Languages and Language'. (See the latter, pp.10-12 for an explanation).

iv. A note on game theory.

Lewis's initial account of the preference structure underlying conventions makes use of some terms and procedures from the mathematical theory of games, together with some related terms of his own. From its beginnings on, students of 'game' theory have stressed its relevance to real situation, not only to what we naturally call 'games' (like chess) but also to a range of economic and social situations. In this section I shall give my own brief outline of some elements of game theory that are related to our subsequent discussion.¹

(1) Game theory deals with situations in which there are a number of 'players' each of which have a number of alternative 'moves' open to them (not necessarily the same moves for each one). It thus applies to situations in which a number of agents have a number of alternative actions open to them.

(2) The situations at issue are situations which might be called 'situations of other-dependent payoff' (my own phrase). For the following may serve as a maximally general definition of such situations: at least one move of one of the players is such that the value of its outcome to him will vary depending on what move at least one of the other players makes. (We can see the possibility of a spectrum of cases).

At one point Lewis writes 'games—problems of interdependent decision' in discussing Schelling's work.²

¹ Some loci classici on game theory: the theory was invented by John Von Neumann, see the Theory of Games and Economic Behaviour (1953) by Von Neumann and O. Morgenstern. More accessible to non-mathematicians is R.D. Luce and H. Raiffa, Games and Decisions (1957). Thomas C. Schelling Strategy of Conflict (1960) stresses the study of 'co-ordination games' David Lewis acknowledges a debt to Schelling (p.3).

² Lewis p.13. The phrase 'interdependent decision' originated with Schelling, see The Strategy of Conflict, p.16, pp.83 ff, and 89n. Schelling's own account of what such a situation is is somewhat unsatisfactory. See R. Nozick's The Normative Theory of Individual Choice, Princeton, Ph.D. thesis, 1963 (unpublished).

If one uses my general account of game theoretic situations this characterisation - 'problems of interdependent decision' - will not be universally appropriate. Moreover, it is at least moot whether all situations of interest to game theorists and discussed by them in the literature, are aptly characterised in terms of this phrase. I have in mind in particular the famous 'Prisoner's Dilemma' case in which each player has a 'dominant' strategy, that is, a choice of action such that no matter what the other player does he himself will get a better payoff than on any alternative strategy he might adopt. For this kind of case the phrase 'problem of interdependent decision' seems inappropriate.

It is true, however - and important in the present context - that in many situations of other-dependent payoff we can say things like this: I should take into consideration what you are likely to do in myself deciding what to do. Therefore your decision about what to do would be relevant to mine, if I could know what yours would be. This may be true of you also. The reason for the dependence of my decision on yours will, however, be the dependence of my payoff on your decision (action) and vice versa. Thus the notion of an other-dependent payoff seems to me to be both more general and more basic than that of 'interdependent decision'.

The task of characterising a 'game-theoretic situation' adequately is evidently somewhat problematic, and my own characterisation may involve problems of its own. One thing I would not count as an objection is that relatively trivial or unproblematic situations will be regarded as 'game-theoretic' on this account. These will just be the trivial or unproblematic cases of the type dealt with by the theory. ^Conceivably my account is too wide in other ways; however, with this caveat it may serve us here.

(3) 'Each player has preferences over the outcomes which meet the axioms of utility theory' (Luce and Raiffa, p.49). In particular, and roughly, this entails that the players' preferences among outcomes can be represented on a

scale of arbitrarily chosen units so as to show (a) which outcomes are preferred to which other outcomes and (b) how much, relatively, the outcomes are preferred to each other. Preference is assumed to be transitive.

(4) A device for representing such situations, in particular for two person games, is a 'payoff matrix'. In relatively brief compass, we can indicate what the range of possible actions for each agent is, and show exactly how the outcome of a given action by a given agent varies in value depending on what others do.

Example:

		<u>Col</u>	
		A	B
<u>Row</u>	I	0 0	1 1
	II	1 1	0 0

Key: 'Col' and 'Row' are labels for the players. (Cp. Schelling's 'Row' and 'Column', Lewis's 'Row-Chooser' and 'Column-Chooser'; my terms at least have the advantage of brevity). (In future diagrams I shall not write in the labels for the players; they will be assumed to be as here.)

'A' and 'B' stand for Col's two available actions here; 'I' and 'II' for Row's. In each 'cell' of the matrix (e.g. the shaded area) the lower left figure indicates Row's payoff, the upper right indicates Col's.

It is clear that whether Row gets the payoffs represented by '1' or '0' for doing I, will depend on whether Col does A or B, and so on.

(5) No interpersonal comparison of preferences is assumed; for various points of classification, however, the device of 'linear rescaling' may be used (cp. Lewis p.13,14) (see the following point).

(6) Probably the most discussed kind of game is the so-called 'zero-sum' game, in which the agents' payoffs in each square can be represented as summing to zero. (In fact it is the two-person zero-sum game which has been the central case studied).(See, for instance, Luce and Raiffa.p.385).

Now, take the matrix

	1	5
1	-3	
	5	1
-3		1

Is this game 'zero-sum', that is, do the agents' payoffs sum to zero in every square? Not in this representation of the payoffs, but remember that the numbers simply represent the relative values of the outcomes for each one, using arbitrarily chosen units. Suppose, then, we subtract 2 from each of Row's payoffs as represented here. We get the matrix

	1	5
-1	-5	
	5	1
-5		-1

in which the agents' payoffs are represented as summing to zero in each cell. This matrix shows essentially the same game as before.

(7) Now, clearly the two person 'zero-sum' game is a game of conflict, intuitively; in particular two conditions are fulfilled: (i) if Row prefers a cell A to a cell B, Col will prefer the cell B to the cell A; (ii) if Row likes A twice as much as he likes B, Col will like B twice as much as he likes A, and so on. Thus we can say that the agents' rankings of the different outcomes are perfectly correlated inversely.¹

The kinds of games studied are not restricted to cases of such conflict. Clearly, a whole spectrum of cases is possible. Schelling argues for the utility of studying (pure) 'co-ordination games' in which (Schelling, p.84) 'the players win or lose together: they...rank all possible outcomes ~~identically~~, in their separate preference scales'. (If they do this, then, as Lewis notes, 'after suitable linear rescaling' the agents' payoffs will be equal in every square). Then there are the 'mixed cases' in which there is no perfect correlation of preferences direct or inverse, as in:

	0	12
0	12	
	6	-1
12	1	

(8) The theory is concerned with players who are concerned to maximise their 'payoff' from what they do. We may interpret this as follows: they wish to do as well for themselves as they can, according to their own appreciation of

¹ This intuitively attractive account of the 'conflicted' character of zero-sum games is only valid for two-person games, as we shall see.

what 'doing well' amounts to.¹ These players do not have to be 'selfish' in any intuitive sense, though they may be.²

(9) Roughly speaking, the aim of game theory as 'classically' conceived is to say what choices the players 'should' make in given, well-defined classes of cases (for instance, the two-person zero-sum case). In other words, it is to demarcate 'solutions' for different classes of cases.

(10) To return, briefly, to the matter of complex 'action' descriptions, noted earlier; we might now consider this issue in terms of a game matrix.

Consider the following:

	A	B
I	0 1	2 1
II	2 2	1 0

(Imagine A=doing the washing up; B=going to bed; I=finishing writing a chapter in the study, and II=marking exam papers by the fire. Row most prefers marking exams while having Col enjoy her expostulations as he washes up, and so on).

We could here say that Col and Row most prefer that both do AII, where 'doing AII' stands for 'doing one's part in the combination of actions in which Col washes up and Row marks exams.' More naturally we can say each one prefers that both do their part in the action-combination AII (or in AII, for short). We shall often

¹ Such players will sometimes be called 'rational' ; but see Luce and Raiffa p.5, for instance, on the plethora of definitions of 'rational' in game-theoretical discussions, and also below, on Lewis's own apparent usage.

² See Luce and Raiffa, p.120, on various possible interpretations of 'payoffs' in a game.

use the latter type of formulation in what follows.

3. Lewis on co-ordination problems: a critique.

i. Introductory.

Lewis's first, rough definition of 'convention' as stated in Convention (p.42) is couched in terms of co-ordination problems.

'A regularity R in the behaviour of members of a population P when they are agents in a recurrent situation S is a convention if and only if, in any instance of S , ^{among} $\wedge$ members of P ,

- (1) everyone conforms to R ;
- (2) everyone expects everyone else to conform to R ;
- (3) everyone prefers to conform to R on condition that the others do, since S is a co-ordination problem and uniform conformity to R is a co-ordination equilibrium in S .'

(There is a slip in clause (3) as printed which should be rectified:

(3) should read: '... S is a co-ordination problem and uniform conformity to R is a proper co-ordination equilibrium in S .)'¹

Lewis's account of co-ordination problems is, in a sense, the focus of his whole account of conventions. 'Co-ordination problems' are (apparently) carefully defined in game-theoretical terms at the beginning of the book and before the first, rough, definition is given.

A large part of the attraction of Lewis's work lies, to my mind, in his attempt to harness the rather precise 'mathematical' theory of games to the old philosophical end of conceptual analysis.

One might be as or more attracted by Lewis's claim to have harnessed game theory in the 'analysis' of an intuitively important and prevalent 'natural' class of real situations.

In either case, as Lewis himself writes, 'my theory of convention had its source in the theory of games of pure co-ordination - a neglected branch of the

¹ In view of the trouble Lewis takes to introduce his notion of a proper co-ordination equilibrium (of which more, shortly) this slip is odd. When I pointed it out to Lewis (in conversation, spring 1977) he agreed that this was indeed a slip - which had not previously been brought to his attention.

general theory of games of Von Neumann and Morgenstern' (p.3). Though Lewis goes on to say that 'in the end, the theory of games is scaffolding, I can restate my analysis of convention without it', his analysis of convention is developed from game-theoretical beginnings, and these may I think equally aptly be thought of as foundations.

I shall now spend some while on the project of pinpointing problems in Lewis's scaffolding - or foundations. If there was ever any interest in such a scaffolding at all - as I believe there was - it should be worth considering its defects in some detail. This is even more to the point since a number of Lewis's problematic claims about co-ordination problems have recently been accepted uncritically by some philosophers who have made use of his work.¹

ii. Lewis's definition of 'co-ordination problem'.

Lewis's definition of 'co-ordination problem' is itself somewhat problematic: the meaning of a crucial technical term 'proper co-ordination equilibrium' - is not made clear; two distinct accounts appear possible from what Lewis gives us. Moreover there appears to be a discrepancy between Lewis's definition of 'co-ordination problem' and his final account of the preference structure underlying convention, however one interprets the ambiguous term. Since these points have not to my knowledge been noted by previous commentators or by reviewers,² it seems worth making clear what the situation is, here.

Lewis's preferred account (definition) of co-ordination problems is as follows:

¹ In particular Stephen Schiffer in Meaning (esp.p.143) and Edna Ullman-Margalit in The Emergence of Norms (p.80).

² Nor by Lewis himself till I pointed them out to him, in conversation, May, 1977.

they are

'situations of interdependent decision by two or more agents in which coincidence of interests predominates and in which there are two or more proper co-ordination equilibria' (p.24)

Two technical terms must be interpreted before the above account can be understood;

viz: 'proper co-ordination equilibrium' and 'coincidence of interests predominates'

A. 'Proper co-ordination equilibria'.

Unfortunately, Lewis does not explicitly define 'proper co-ordination equilibrium' (which I shall henceforth call 'p.c.e.' for short), in spite of its centrality to his scheme. However, he defines some related terms. Let us, therefore, consider these.

(1) An 'equilibrium' simpliciter is a combination of agents' actions such that no one could have done better by acting differently himself, given the others' choices (Lewis, p.8). (The term 'equilibrium' is standard in game-theory. Cf. Luce and Raiffa, p.62).

(2) A 'co-ordination equilibrium' is

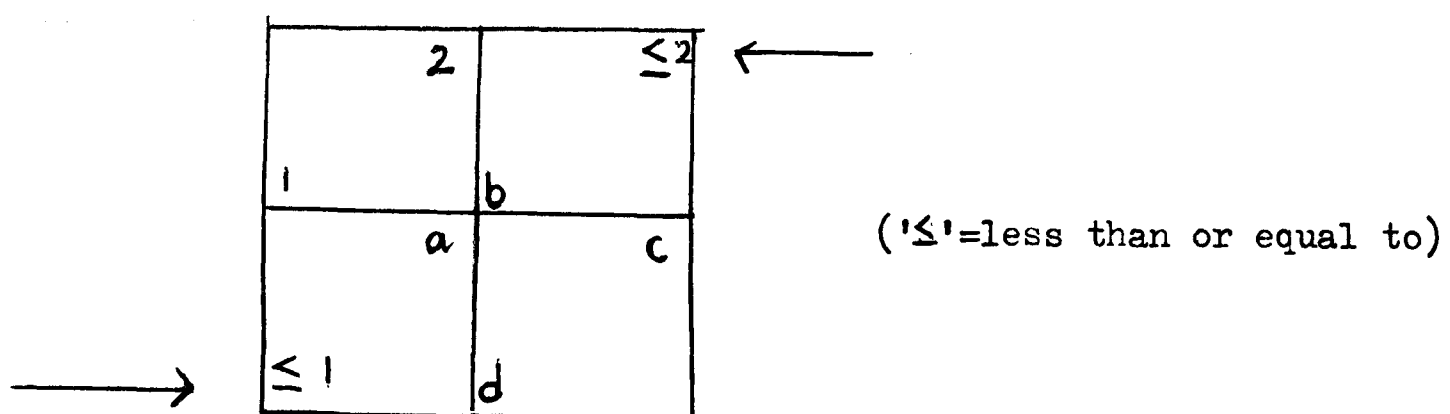
'a combination in which no one would have been better off had any one agent alone acted otherwise, either himself or someone else' (p.14).

(3) A 'proper equilibrium' is a combination of actions such that

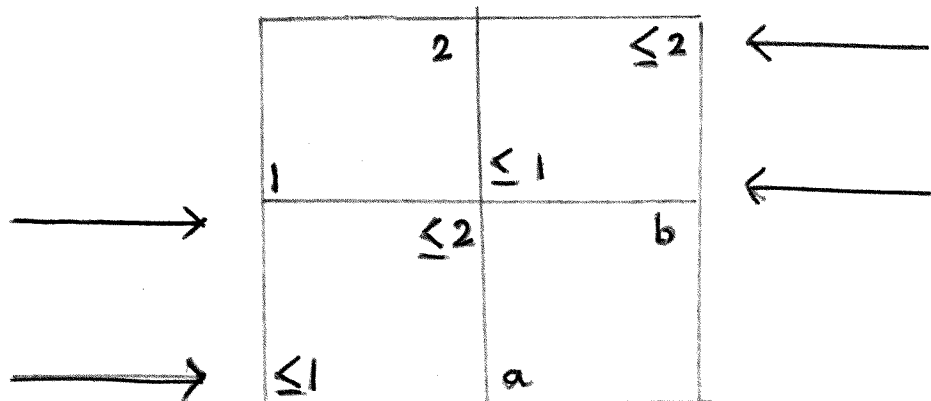
'each agent likes it better than any other combination he could have reached, given the others' choices ' (p.22).

We may diagram these three possible types of action-combinations as follows. (Here the action combination in question is represented in each case by the top left hand cell of the matrix. Each arrow points to a restriction on payoffs that is necessary for the upper left hand square to represent an equilibrium (or whatever). The payoffs represented by letters a, b, c, etc., can be anything one pleases, and the upper left hand cell will still represent an equilibrium (or whatever)).

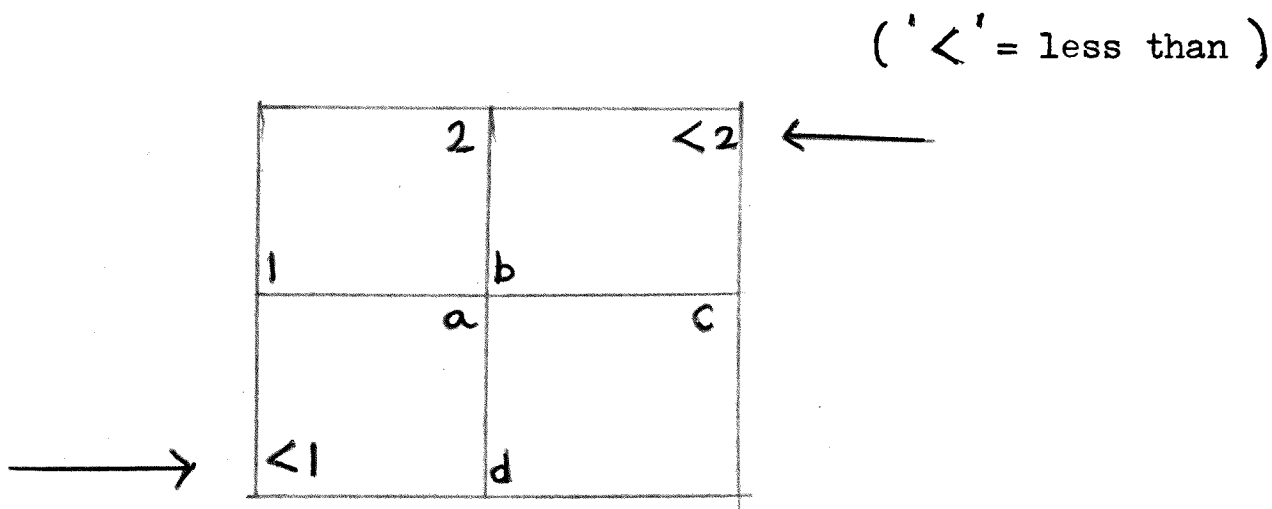
Equilibrium



Co-ordination Equilibrium



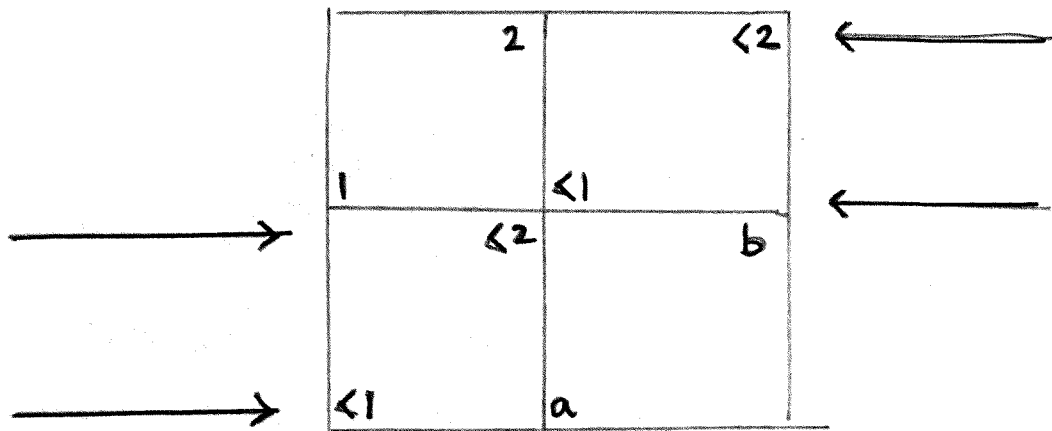
Proper Equilibrium



Now, given these definitions, it may seem that a p.c.e. will be a combination of the different agents' actions such that (a) no one would be better off had any one agent alone acted otherwise, and (b) each likes it better than any combination he could have reached given the others' choices. In other words, a p.c.e. would be a co-ordination equilibrium which is also a proper equilibrium.¹

¹ R. Grandy interprets Lewis this way in his review of Convention, J.Phil. 74 (1977) pp.130-1. Recently Lewis 'interpreted' himself in this way, also, in conversation with the present author, Spring 1977.

Another, not implausible, way of taking Lewis here would be this: in a p.c.e. each one prefers this combination to any other in which some one agent, himself or another, had acted otherwise, In diagram form:



Here, the qualifier 'proper' converts everything restricted merely to $\leq n$ in a co-ordination equilibrium to $< n$.

In this interpretation, we see the restriction on 'properness' working in an analogous way on 'mere' equilibria and on 'mere' co-ordination equilibria - strengthening in a global way the restrictions imposed by the initial condition.

This was the interpretation which first occurred to me.¹ And Edna Ullman-Margalit seems inclined to make it.² An ambiguity in the implicit definition of 'p.c.e.' has now, I trust been demonstrated.

Now, if we understood 'p.c.e.' according to the first ('weak') interpretation above, it does not follow that, in a p.c.e., each one will prefer any lone nonconformist to conform. Each could, apparently, be neutral with

¹ See my 'About Conventions' (1974) p.72.

² In The Emergence of Norms (1976): 'a proper co-ordination equilibrium... is a co-ordination equilibrium in the strong sense' (p.81). In fact her explicit account of a p.c.e. is surprising, and unjustifiably weak: it must surely involve an inadvertent slip. Thus she writes that a p.c.e. is 'a co-ordination equilibrium such that at least one agent would have been worse off had any one alone acted differently'(ibid). Why 'at least one'? No reason is given for not having 'each agent' here.

regard to the other's conformity and some nonconforming state. Such neutrality is not possible in a p.c.e. in the second ('strong') sense. Nor is it possible in the structure posited in Lewis's final account of the preferences underlying convention. Thus Lewis's fourth condition in 'Languages and Language' is, (in part)

'There is a general preference for general conformity to (regularity) R rather than slightly-less-than-general conformity - in particular, rather than conformity by all but any one' (p.5).

We might put the point this way: were we to cull an account of 'p.c.e.' from 'Languages and Language', we should define it using the strong sense of p.c.e. noted earlier.

If Lewis did mean 'p.c.e.' to be interpreted in the weak sense, has he clearly changed his mind on the structure of preferences underlying convention by 'Languages and Language'? We recall the condition - as yet unexamined here - in the definition of 'co-ordination problem' that 'coincidence of interests predominates'.¹ This does not occur in 'Languages and Language': has it perhaps been incorporated into the implicit strong definition of 'p.c.e.'? We now turn to this question.

B. Predominant coincidence of interests.

The suggestion that the 'coincidence clause' in the definition of 'co-ordination problem' has been elided with the requirement of two weak p.c.e. in the final account might seem to be supported by Lewis's claim, of his fourth condition in 'Languages and Language' that

'it serves to distinguish cases of convention, in which there is predominant coincidence of interest, from cases of deadlocked conflict' (p.5).

Let us first consider what sort of coincidence of interest the condition in question does require, or rather, what sort of coincidence of interests occurs when there is one p.c.e. in the strong sense.

First, if the condition is fulfilled, there will have to be at least a certain (relatively small) degree of coincidence of interests in the situation.

¹ For brevity, I henceforth refer to this as the 'coincidence clause'.

For each agent will have a certain conditional preference concerning his own action (preferring to do A if everyone else does), such that every other agent will at least endorse that preference of his. Another way of putting it: each one will have a certain partial ranking of outcomes, one above another, that the others endorse. (If there is a p.c.e. in the weak sense, agents need not agree on any ranking of two particular outcomes one above the other).

Second - a corollary of the first point - it appears that situations where Lewis's fourth condition holds will never be zero-sum games. This is at least quickly obvious in the two-person case: since each player must share a preference or ranking of two outcomes one above the other, they clearly cannot have that perfect inverse ranking of outcomes which characterises games of the two-person zero-sum type. Now, games of the latter type are, intuitively, cases of 'pure' or 'deadlocked' conflict. Hence Lewis's fourth condition does, clearly, rule this kind of conflict out of the situation in two-person 'games'. So too, however, does the original, weaker condition of one or more p.c.e. in the weak sense.

N-person zero-sum games evidently cannot be so purely 'conflicted' as the two-person sort. There is no necessary conflict between any two players in the n-person game. However, the players are never all in unison over any ranking of outcomes one above another. Nor can it even be the case that the players all either share a given person's ranking of A over B, or are at least neutral between A and B. To this extent, then, they are, as a group, in conflict, but if Lewis's fourth condition holds among players they cannot, it seems, be conflicted to this extent. So, the fulfilment of the condition in n-person games means they are not zero-sum. This is also true, once again, in situations with one or more p.c.e. in the weak sense. Both weak and strong p.c.e. requirements, then, rule out the conflict inherent in zero-sum games, but the weak p.c.e. requirement allows for slightly more conflict in a situation, as our first point above made clear.

Looking back to the definition of 'co-ordination problem', in Convention, we find the requirement that in a co-ordination problem 'coincidence of interests predominates'. Does this 'coincidence clause' in the definition of 'co-ordination problem' have at least the following function : to ensure whenever there are two p.c.e's, in the weak sense, there will in effect be two p.c.e's in the strong sense, and hence there will be the particular kind of coincidence of interests two strong p.c.e's involve?

It turns out that we cannot take it that the 'coincidence' clause in Convention has the function in question. In the Convention section on co-ordination problems Lewis gives an explicit account of what he there means by 'coincidence of interest predominates'; he there uses the phrase in a special, technical sense. (In 'Languages and Language' he is fairly clearly no longer using the term in this sense). Lewis makes it clear that he does not want to require, in co-ordination problems, 'perfect coincidence of interests' (p.14). Then, without argumentation, he stipulates that they are to be defined in terms of predominant coincidence of interest, that is, as situations in which

'the differences between different agents' payoffs in any one square (perhaps after suitable linear rescaling) are small compared to some of the differences between payoffs in different squares' (p.14).

Now, regarding our present query, we can construct a game matrix showing a situation with two p.c.e's in the weak sense only, which fulfils the stated requirement of the coincidence clause.

	A	B	C
1	5 5	4 5	0 0
2	5 4	0 0	5 4
3	0 0	4 5	5 5

Here the 2 p.c.e.(weak sense) are A1 and C3. Coincidence of interests predominates (by the definition given), but it is not the case either that Row prefers A1 to B1, or that Col prefers A1 to A2. Hence it is not the case that each agent prefers A1, a p.c.e., to any combination in which any one, himself or another, acts otherwise; the same goes for the other p.c.e., C3.

The coincidence clause in the definition of 'co-ordination problem', then, does not entail that given a p.c.e. (weak sense) we also have (in that same combination of actions) a p.c.e. in the strong sense. We may note also that if we do have 2 p.c.e. on the strong interpretation we do not then necessarily have 'coincidence of interest predominating' in the sense of Lewis's definition. See the following matrix:

		A	B	C	D
1		3	0	0	0
2		0	3	0	0
3		0	0	-15	15
4		0	0	15	-15

Y

Here A1, B2 are proper co-ordination equilibria on the strong interpretation, but it is not the case that the differences between X's and Y's payoffs in any one square are 'small compared to some of the differences ^{between payoffs} in different squares'. It follows that we do not necessarily have 'co-incidence of interests predominating' if we have 2 p.c.e. in the weak sense .

Finally, if we posit 'co-incidence of interests' in a situation, do we automatically have one or more p.c.e. on either interpretation? Apparently not (see matrix below)

	16	16
16	16	1
16	1	

It is clear, then, that the 'coincidence clause' does not have the effect of bringing the two possible definitions of 'p.c.e.' together; the coincidence clause appears to have no bearing on either p.c.e. requirement, nor they on it.

iii. Shifts in Lewis's thinking, and some queries arising therefrom.

A. If Lewis started intending a weak interpretation of 'p.c.e.' in the definition of 'co-ordination problem', by 'Languages and Language' he has implicitly changed his definition to one using a notion of p.c.e. in the strong sense. If in fact he always intended a strong interpretation of 'p.c.e.' he has still changed his view of the structure of preferences underlying conventions, for he has now dropped the further requirement that 'coincidence of interests predominates'. Either way, then, he has changed his view, since the original account of co-ordination problems contained the very specific 'coincidence clause' we have noted.

Lewis has also clearly adopted a requirement of what are in effect two p.c.e. in the strong sense by the time he reaches his final definition of 'convention' in his book. Thus 'we require that all share the conditional preference of each for his conformity to R' (p.69). If there has been a change here, it is neither noted nor argued for. Of the requirement just noted, Lewis writes: 'if S is a self-contained problem of interdependent decision, this... requirement makes uniform conformity to R a proper co-ordination equilibrium'(69). This may show that Lewis at this point thinks that a p.c.e. in the strong sense is what was at issue all along, but of course the last remark is consistent with a weak interpretation of 'p.c.e.' Moreover, though in this section Lewis explicitly intends only to alter his conditions on convention in ways irrelevant to our present discussion - the structure of co-ordination problems - (see.p.68), he makes what amounts to a fairly major change in the area of the overall coincidence of interests that is required in the situations at the base of conventions. For consider that at this point in Convention the original

'coincidence clause' has gone, while a new clause has appeared: in the problem situation to which the convention applies, every agent must have 'approximately the same preferences regarding all possible combinations of actions' (p.78. see also p.69). This is perhaps supposed to be a somewhat vaguer statement of the original coincidence requirement, but in fact it looks quite a bit stronger. Indeed, Lewis himself seems to regard it as entailing that every equilibrium in the situation must be a proper co-ordination equilibrium in the strong sense (see p.69). Since the original coincidence clause does not entail this, it will certainly be stronger than that clause if it has this entailment. (It will also be stronger than Lewis previously thought he needed. On p.22 he actually wrote 'there is no need to stipulate that all equilibria in a co-ordination problem must be proper' and he gave an example of a case he considered 'essentially similar to our clear examples of co-ordination problems, despite the impropriety (of one equilibrium)' (22)).

Since defining 'co-ordination problem', then, Lewis certainly changed things in Convention with a new clause about overall coincidence of interests which is different from the original 'coincidence clause' and apparently much stronger. Another change - depending on where he started - may be that from a weak to a strong p.c.e. requirement. That he ends with a strong requirement is clear. 'Languages and Language', differs, finally, from the last Convention statement by a final dropping of any clause about overall coincidence of interests.

B. The strong interpretation of 'p.c.e.' requires in effect, a specific area and type of 'coincidence of interests' among agents. By 'Languages and Language' Lewis conjectures that the requirement that a p.c.e. in this sense underlies a convention may be all that is needed in respect of the 'coincidence of interests' involved in conventions (p.16). Apparently abandoning his old, technical definition of 'predominant coincidence of interests', and the requirement in terms of it, and abandoning the later general-agreement-in-preference clause in Convention, he writes that the requirement of an underlying

strong p.c.e. 'serves to distinguish cases of convention, in which there is a predominant coincidence of interest, from cases of deadlocked conflict' (ibid). We can see that the strong p.c.e. does serve this distinguishing purpose; but so, we have noted, would the weak p.c.e. requirement have done. It is still unclear which requirement Lewis meant originally to posit, and some obvious questions arise. Would the 'coincidence of interests' already required through the weak definition of 'p.c.e.' be sufficient for Lewis's purposes?. Or was the stronger definition of 'p.c.e.' necessary? And is the requirement of two p.c.e. (on either definition of 'p.c.e.') necessary or sufficient specification of the preference structure underlying convention? If so, is any further 'coincidence of interests' clause necessary?

The answer to these questions, in the context of discussion of our everyday notion of convention, clearly hangs on our intuitions about when we do and when we do not have a case of convention; we shall turn to this issue later.

We can also ask these questions, however, in the context of what may be an important new concept of Lewis's. We can only answer the questions in this context if and when we have a better grasp of the phenomena, conventions proper or not, that Lewis wishes his concept to capture. It is to the latter issue that we now turn.

iv. Lewis's examples: two distinguishable structures.

Before defining 'co-ordination problem' Lewis presents some examples. Indeed, the definition is intended to arise out of the examples. In a footnote (p.24) Lewis explains his procedure as that of introducing a class by means of examples and taking the 'defining features' of the class to be 'those distinctive features of the examples which seem important for an understanding of their character'¹ (Here we should bear in mind that these examples may have been chosen with another set in mind: examples of conventions).

¹ Lewis cites M. Slote's suggestion that this is a common procedure; see Slote's 'The Theory of Important Criteria' (1966), J.Phil, 63, pp.211-224.

Before looking at Lewis's discussion of his examples, I note that among the examples, as presented, there seems to be the possibility of at least two rather different kinds of case. Since Lewis never brings out the difference I have in mind I shall do so briefly here. Insofar as there are two different kinds of case one might think that the definition of co-ordination problem should opt for capturing one kind only, if it is to isolate a relatively natural kind of phenomenon. Moreover, we shall see that established solutions in one kind of case will not be conventions according to the final definition that Lewis (willy-nilly) ends up with.

First, let us take the 'telephone case' (p.6., see also Schelling p.94):

'Suppose you and I are talking on the telephone and we are unexpectedly cut off after three minutes. We both want the connection restored immediately, which it will be if and only if one of us calls back while the other waits. It matters little to either of us whether he is the one to call back or the one to wait. We must each choose whether to call back, each according to his expectation of the other's choice, in order to call back if and only if the other waits.'

Let us diagram this case. Lewis says that 'it matters little...' so we might suppose that it does matter somewhat. Thus suppose each one would prefer not to call back, ceteris paribus, since it will cost him some money to do so. The matrix could look like this:

	Wait	Call back
Wait	0	9
Call back	10	0

If 'it mattered not at all...' the matrix could look like this:

	Wait	Call back
Wait	0	10
Call back	10	0

(Note that Lewis states, pp.9-10, that he does indeed wish to allow for these two kinds of possibilities. Cp. His Figs 1 and 2).

Second, consider the 'staghunt' case (from, Lewis says, Rousseau's Discours sur l'inégalité):

'Suppose we are in a wilderness without food. Separately we can catch rabbits and eat badly. Together we can catch stags and eat well, but if even one of us deserts the stag hunt to catch a rabbit, the stag will get away; so the other stag hunters will not eat unless they desert too. Each must choose whether to stay with the stag hunt or desert according to his expectations about the others, staying if and only if no one else will desert' (p.7).

Now, if starving is the option if we don't at least get a rabbit for ourselves, the matrix could presumably look something like this:

Diagram I

	Separate	Staghunt
Separate	50	0
Staghunt	50	100

Or the 50s might well be 80s, say. Note that in a real case unless we can feel for some reason virtually certain of what the others are going to do it looks

¹ This represents, in fact, some analogous 2-person case.

as if the obvious decision for each is to go separately to catch a rabbit, since that way he minimizes his possible loss and won't starve. So given lack of definite expectations, agents here presumably should separate unless they are great risk-takers. In the telephone case there is no such obvious 'maximin' (maximizing your minimum possible outcome) strategy available. There could be, however, if say, you spent some money whenever you made a call, whether you reached the party you were calling or not. In any case it is clear that real agents may wish - in either case - to find 'a better way' than the combination of maximin strategies. So it is not the case that the stag case is in no intuitive way a problem situation.

A more crucial aspect of the 'staghunt' case as diagrammed here, something implicit in the matrix and which clearly distinguishes it from the rest, is that the agents have no particular desire that the others hunt rabbits 'with' them, given that they themselves do hunt rabbits. True, they desire to hunt rabbits if the others do, but they don't care whether or not the others hunt rabbits if they do. They most desire that everyone staghunts, that is, that they hunt stags if the others do, and that the others hunt stags if they themselves do. Compare this with the telephone case. There each one desires to re-establish the connection. More specifically each desires either that each one calls back if and only if he was the original caller; or that each one calls back if and only if he was not the original caller.

Now, we could interpret the stag case as one in which everyone prefers that no one starve. Then the matrix could be redrawn, something like this:

Diagram II

	70	0
70	50	100
0	100	

In such a case 'co-ordination over separation' clearly is something each one desires in itself. Yet Lewis's description of the case gave no reason to suppose it was of this kind. Indeed for all he says his hunters could be partial misanthropes whose matrix looks like this:

Diagram III

	50	0
50	80	100
0	100	

Interestingly, Lewis's final definition of 'co-ordination problem' rules this matrix out whichever interpretation of 'p.c.e.' we adopt . It does not, however, force a decision between I and II. The point about there being an obvious minimax strategy still stands, indeed stands firmer, with diagram II, but substituting II for I the case would then resemble cases like the telephone case in what is a salient and may be an 'important' respect, and II, not I, involves a proper co-ordination equilibrium in the strong and not merely the weak sense. These considerations might, evidently, suggest the use of the strong definition of p.c.e. in a definition of 'co-ordination problem', but Lewis says nothing about them.

So much, then, for two distinguishable types of structure apparent in Lewis's examples. It is hard to be sure whether Lewis meant the staghunt case to have a weak p.c.e. structure or not; however, his description can surely be interpreted in that way.

v. Lewis's arguments for his definition of 'co-ordination problem'.

We now turn to Lewis's discussion of co-ordination problems and the elements that are to be counted definitive of them, the 'common character of co-ordination problems' (p.8).

A. Description.

At the outset of his discussion Lewis says that 'the outcome of any action an agent might choose depends on the actions of the other agents'.(8) If we take 'outcome' to be equivalent to 'payoff' here, we may prefer, at least initially, not to be quite so restrictive even for co-ordination problems.

Consider that in, say, the telephone case each agent (presumably) has the option of committing suicide rather than working on the (say) rather hopeless task of re-establishing a connection. The outcome of this action (committing suicide) for each one, does not depend on the actions of the other agents. Assuming that it makes sense to assign payoffs to the act of suicide then we could get a matrix for the whole situation like one of these:

	call back	wait	suicide
wait	1	0	0
call back	0	1	0
suicide	0	0	0

	call back	wait	suicide
wait	10	0	1
call back	0	10	1
suicide	0	0	1

	call back	wait	suicide
wait	1	0	-1
call back	0	1	-1
suicide	-1	-1	-5

and so on.

So we might at this point want to say more tentatively, that co-ordination problems are clearly what I call situations of other-dependent payoff, situations in which at least one action of one of the agents is such that his payoff from that action is a function of what the other players do, and in fact in most of the example cases at least two actions of each of the agents is such that his payoff, and so on. (Not so, however, on one interpretation of the staghunt case!) Neither this latter feature, nor of course the former, are distinguishing marks of co-ordination problem situations.

Lewis notes that in all his examples there are equilibria; but that co-ordination problems are not distinguished by equilibria (p.13) since these can occur in situations of pure conflict - and, Lewis implies, co-ordination problems are not situations of pure conflict. Lewis suggests that his co-ordination problems are among those 'at or near the pure co-ordination end of Schelling's spectrum' (14) (see above). He does not 'want to require perfect coincidence of interests' (14), and we have just noted that in almost all his examples he leaves room for the players' rankings of all the possible outcomes to differ at least somewhat.

Lewis then introduces his 'coincidence of interests predominates' clause. We have already considered some aspects of this clause. It is worth noting some more here.

The restriction does allow differences between agents' payoffs in a given square, as is intended. By definition, it ensures something along the lines that these differences will be relatively small.¹ However, it is worth stressing here that one thing it does not ensure is that 'it matters little' to either which of some of the various combination is chosen. Nor, most interesting and most important, does anything in Lewis's definition of 'co-ordination problem' anywhere ensure that this is so. There may be reasons for not requiring this. However, we may note that it is a stated feature of almost all of Lewis's examples. (It maybe false of the stag case; however, compared with starvation it presumably does matter relatively little whether the people eat well or badly. In one other case (of some contented oligopolists') it is not clear that an 'it matters little' could not have been inserted.)

¹ It does not, however, rule out a situation like the following:

	A	B
I	50 900	1000 1000
II	1000 1000	-1000 -1000

where Col likes AI, BI, and AII more or less equally, while Row will much prefer BI or AII to AI (while agreeing on BII's worst ranking). (This situation will in fact count as a co-ordination problem as Lewis defines it).

Lewis next notes that all the sample co-ordination problems have two or more different co-ordination equilibria (p.16) and here some 'a priori' reasoning comes in. It is reasoning to the conclusion, first, that a co-ordination problem should be defined as having two or more co-ordination equilibria (its two co-ordination equilibria are an important feature of it); and second, that it should have - by definition- at least two p.c.e. Following this reasoning might be expected to help throw light on the sense intended for 'p.c.e.'. However, the reasoning is often obscure and, I shall argue, often mistaken. Since some apparently erroneous claims have been accepted by others, I give Lewis's claims quite careful scrutiny in my next section.

B. Argumentation: the 'deduction' of two proper co-ordination equilibria.

First, Lewis claims that 'If there is no considerable conflict of interest, the task of reaching a unique co-ordination equilibrium is more or less trivial' (my stress)¹.

In my view this claim is simply false. In fact, it breaks down into a number of related claims. First, a general claim seems to be implied in the above quotation, namely, that if there is a unique co-ordination equilibrium (in a situation without considerable conflict of interest) then it will be the outcome most desired by each agent. This claim will be refuted shortly. Second, Lewis implies that there are varying degrees of triviality in situations with unique co-ordination equilibria - and he proceeds to discuss, in sequence, the general case and then a special case with 'dominance'. There are many puzzling things in this discussion. We proceed to follow his discussion with a critical eye.

¹ Cf. Stephen Schiffer, 'borrowing' Lewis's 'idea of arriving at an account of convention via 'the theory of games', agrees here-or at least goes along with Lewis (p.143). Thus Schiffer: 'situations occur in which there is only one co-ordination equilibrium. In such cases co-ordination is not a problem'.

To begin, we consider Lewis's claim about the general case. He seems to want to argue that situations with a unique co-ordination equilibrium (or 'c.e.' for short) are always trivial at least in the following way: the unique c.e. will be reached:

'if the nature of the situation is clear enough so that everyone makes the best choice given his expectations, everyone expects everyone else to make the best choice given his expectations, and so on' (p.16).

This claim is not very clear. For one thing depending on whether we take Lewis to be talking only about what I shall call a 'one-off' case (see below), or about all possible cases with a unique co-ordination equilibrium, we get rather different possible claims.

By a 'one-off' situation here I mean something rather specific: a situation (like one of Lewis's examples) in which no one has or can come by any (justified) expectations about what the other agents involved will do, in advance of his own decision about what to do, except what he can glean by considering that there is group common knowledge among the agents concerned of their shared rationality and of the matrix. (I leave open for a moment the sense of 'rational' intended here). We might say that, in such a situation, no one has any external expectations about what will be done. Typically this will be in cases where the situation has never occurred before among these agents, but a case's being of this kind is not sufficient for its being 'one-off' in the sense I have in mind. (For instance, though these agents may never have been in this situation, it may be common knowledge among them that certain other people have, and have successfully resolved the question of what to do in a particular way. This may justify expectations of each other's acting in the relevant way). It is clear that in a 'one-off' case as defined, agents have to decide what to do without resort to communication. (Lewis himself does not mention communication in these pages).

Consider, then, that we are talking about a one-off case. Then Lewis's claim about 'ordinary' unique co-ordination equilibria could be construed something like this:

(1) the best (most desired) combination for each agent individually will be that equilibrium (this is the general claim noted earlier); therefore(?)

(2) given only common knowledge of the matrix and of their own rationality, each one will expect everyone else to do his part in the co-ordination equilibrium, and hence, being rational,

(3) each will do his own part, and thus the unique co-ordination equilibrium will be reached.

That something like this was meant here is indicated by remarks Lewis makes later (p.70), thus:

'Recall the discussion...of the triviality of any situation with a unique co-ordination equilibrium and predominantly co-incident interests. We are now in a better position to describe this triviality: common knowledge of rationality is all it takes for an agent to have reason to do his part of the one co-ordination equilibrium. He has no need to appeal to precedents or any other source of further mutual expectations' (my italics).

(Lewis's discussions of his concept of 'common knowlege' and of types of sources of 'mutual expectations' have come after the passages on matrix structure that we are considering). I take 'have reason' here as 'have a sufficient (decisive) reason'.¹

Now the interpretation of Lewis's claim in terms of 'one-off' situations is not the only one possible. Lewis could be taken to be making a claim about all possible (relatively unconflicted) cases with a unique co-ordination equilibrium, and not merely one-off cases. To phrase this interpretation as similarly to the former as possible, it would run:

(1) the best combination for each agent individually will be the unique c.e. Therefore(?)

(2) Whatever 'external' bases for expectations each player has in advance

¹ In general, in Convention 'rational agent' seems to mean something like 'agent whose actions and beliefs are based on reasons which justify them (and with whom justifying reasons lead to actions and beliefs)'. We shall construe 'rational' and its cognates here accordingly. I take it that Lewis-rational agents will, also, strive to maximize their payoffs. (On 'rational' in game-theory see Luce and Raiffa, p.5, op.cit.).

of making his move about what the others will do (precedents, agreements, and so on) these will always be irrelevant to his decision, whatever they are, since

(3) common knowledge of rationality and of the matrix alone will always give each agent decisive and sufficient and indeed overriding reason to expect everyone else to do his part in the c.e., and hence

(4) Each will do his part, and thus the c.e. be reached.

In brief, the claim^{is} that all cases can and should work exactly as if they were one-off cases: the crucial basis for expectation and action will be the same.¹

These two interpretations of Lewis are not equivalent; I now argue against the claims made in each of them.

First I introduce some terms of my own. These will be useful now and later also. Let an absolute best point be an outcome such that every other outcome is worse for every agent. Let a best point be an outcome such that no other outcome is preferred by any agent. (Thus there may be more than one best point but not more than one absolute best point in a given situation).

Now, suppose that in a one-off situation with an absolute best point there is group common knowledge of rationality and of the matrix. Then the best combination for each will (by definition) be the point in question and will be obvious to all. Will it be true that agents in such a situation will always expect the others to do their parts in the absolute best point combination, and that each will do his part? (Let us assume that all options are equally accessible, or that difficulties in reaching them are already incorporated in the matrix.)

I suggest that the absolute best point may well not be the obvious thing to aim for or to expect others to aim for under the assumptions stated. If there are grave possible losses involved in the relevant course of action compared

¹ Lewis himself indicated, in conversation, that this is what he had in mind. (Spring 1977).

with others, the agents surely will have reason to be apprehensive of it. Similarly the possible losses for others may make each one less secure in any expectation that the others will do their part in the absolute best point combination, less secure, that is, that common knowledge of rationality and of the matrix, plus their own preferences, will dictate that course of action. Thus consider a matrix such as:

	A	B	C	D
1	5	-500	-500	4
2	-50	3	3	3
3	-500	3	3	3
4	4	3	3	4

Here, 4, D, seem to be the options if any that players should choose in real one-off case with lack of communication. It is an option which minimises each one's possible loss (or in other words 'Maximizes his security level') and is guaranteed to get him relatively near to his personal absolute best positive outcome.¹ There is an absolute best point in this situation; it is, not 4D but AI.

It is certainly false that in a situation with definite grounded expectations about what the others will do one nevertheless always has a decisive and sufficient reason for doing one's part in an absolute best point combination, even given common knowledge of rationality and of the matrix. Consider the

¹ Von Neumann considered maximising one's security level the fundamental aim of a rational agent in a zero-sum 2-person game. It would seem to be a reasonable principle in this case also.

previous matrix. Suppose that for some reason Row reasonably believes that Col, for some reason, will do C. Clearly Row will then be foolish to choose I, wise to choose 2,3, or 4. In fact it seems that in this matrix it will always be wise to choose 4 (or D) except when the other's doing A (or I) is pretty well guaranteed.

Given certain definite expectations about what the other player will do, then, it may well be foolish to perform the action leading to an absolute best point. Note that considerations of maximizing one's security level need not play a role here: if one knows what another will do and wishes merely to maximise one's payoff on that assumption, one may have not to do one's part in an absolute best point. Note also that even if my expectation is that the others will do their parts in an absolute best point, it may still be insufficiently strong to override considerations of risk!

Let us now turn to Lewis's co-ordination equilibria. Lewis might be thought to imply that a unique co-ordination equilibrium will be an absolute best point, when he speaks of the (apparently obvious) 'task' of reaching a unique co-ordination equilibrium. This seems to presuppose that one is obviously going to want to reach such an equilibrium if possible. We may refer also to Lewis's argument, on page 17, for a claim about 'any finite two-person game of pure co-ordination with a unique equilibrium'. One of the premisses is '(RI,CI) must be the unique equilibrium, and $P(RI,CI)$ must exceed every other payoff in the game'. (Here P = payoff). Now, while it may be true that in a pure co-ordination game a unique (co-ordination) equilibrium must be an absolute best point, this is not true for games of slightly-less-than-pure co-ordination, as I shall show.

Lewis himself notes - p.8 - that it is not the case that an equilibrium combination, simpliciter, 'must produce an outcome that is best for even one of its agents'. This is also true of Lewis's co-ordination equilibria; it is true of them in general and of unique co-ordination equilibria in particular.

Thus consider the following matrix:

	A	B	C	D
1	6 6	0 0	0 0	-1000 -1000
2	0 1	0 0	-1000 -1000	0 0
3	-1000 -1000	1 0	90 89	89 90
4	-11 -10	-10 -11	0 0	0 0

Here A1 alone is a co-ordination equilibrium. Yet it is clearly neither a best point nor an absolute best point. Indeed there are none such in this whole situation. Nor is A1 an outcome that would be best for even one of the agents.

Now, in view of the vast superiority of C3 and D3 for both, compared with A1, it is hard to see why either should be interested in A1 in any special way. Without any guarantee about the other's future action, the combination of the two least risky choices of these agents is B4: this is what they would do if they wished to maximize their security level.

In the case we are considering someone might feel that common knowledge of rationality and of the matrix does lead to guarantees about what the other actor will do. It will be noted that Col knows Row prefers C3 or D3 to anything else, and that Row knows Col prefers C3 and D3 also. So perhaps Col can infer that Row will place himself in Col's hands. Then, since Col prefers C3 he will infer that he himself should do C. However, it seems to me that there are really no such guarantees available in view of the maximin 'solution' that is available, and goes contrary to the idea of choice of 3 or C.

If one thing is clear about this situation it is that A and 1 are not the obvious things to do in a one-off case; and the same goes, a fortiori, in a situation where there are expectations that, say, C and 3 respectively are what

the other person will do, derived from facts other than rationality and the matrix, e.g. from facts about precedents. Yet A1 is a unique co-ordination equilibrium.

What we have above then, is a situation in which we have a unique c.e. which is (1) not an absolute best point, hence not the most desired combination for each agent individually, (2) not such that in a one-off case each player will, given common knowledge of rationality and the matrix, expect the others to do their parts in it, and (3) not such that in a one-off case each should do his part in it if rational. There is, hence (4) no reason to suppose that rational players will end up doing their part in this c.e. in a one-off case, and (5) there is even less reason, if that is possible, to suppose that a case with an external basis for expectation should always be treated as it was alleged a one-off case should be treated. This case therefore, provides a counterexample to both of the possible Lewisian claims concerning the 'triviality in general' of situations with a unique c.e.

We noted that - perhaps contrary to Lewis's drift - unique co-ordination equilibria do not have to be absolute best points. Nevertheless they may be. If there is a unique co-ordination equilibrium in a situation and also an absolute best point, these must be the same point. For all absolute best points must be co-ordination equilibria, though not vice versa. However, we found that, as with co-ordination equilibria in general, the absolute best point combination is not always the obvious thing for rational agents to do their part in in a situation with, or without, any external expectations. Our conclusions about absolute best points in general, and about unique co-ordination equilibria in general, also go for those absolute best points which are also unique co-ordination equilibria, as is illustrated by the following matrix.

	1	2	3
A	5	-500	4
B	-500	3	3
C	2	3	2
	4	3	4

Considering the risks involved choosing A and B, Row will presumably not find the choice of A obvious; similarly for Col's 1 and 2, in a one-off situation, and given an expectation that say, Row will do B, Col will hardly choose to do his part in the unique co-ordination equilibrium.

Our conclusions on situations with a unique co-ordination equilibrium, then, so far contradict Lewis's claims. It is not clear that situations with a unique c.e. are generally trivial in any intuitive sense.¹

Lewis next draws our attention to a special class of situations with a unique c.e. which are 'still more trivial (and more deserving of exclusion)' (16-17).

Lewis has in mind here situations in which each agent has a (strictly) dominant choice - a choice they prefer no matter what the others do. These, he notes,

'can have only one equilibrium (and a fortiori only one co-ordination equilibrium), namely the combination of dominant choices.'

This remark of Lewis's does not really make the nature of the triviality of such cases at all clear. As we have seen, the existence of a unique co-ordination equilibrium in a situation does not in any obvious way make the situation intuitively trivial, yet Lewis could here be taken to imply that it

¹ We discuss triviality in more detail below.

is the fact that situations where everyone has a dominant choice must have a unique co-ordination equilibrium that makes them trivial. A few remarks on kinds of intuitive triviality, and in particular the contribution of dominance, seem to be in order.

Let me make a beginning of such an account here. We can, I think divide types of triviality of matrices into at least two kinds, which I shall call the 'choice-trivial' kinds and the 'goal trivial' kinds. Goal triviality is a function of the obviousness of where one would like to end up, if possible. A maximally goal-trivial payoff structure will be one with an absolute best point, that is, one where (a) it is obvious to each agent which combinations of actions he most prefers, and (b) that combination is the same for each one.

One special reason one might feel maximally goal-trivial situations are trivial simpliciter is that should the agents be allowed 'preplay' communication, and in the absence of any special context, rational agents would have no problem agreeing on what each would do. However, in the absence of such preplay communication, as we have shown, things may look extremely problematic for each agent as he tries to decide what to do.¹

Choice-triviality concerns what the agent should do. Maximal choice triviality of a payoff situation seems to arise when there is some choice of action for each agent such that it is obvious that he should make that choice whatever bases for expectation about others' likely actions he has.

¹ We may note here that it has sometimes been argued that games of pure co-ordination in general (where all points have the same relative value for all agents) are trivial, and given preplay communication it might be argued that these situations should be trivial as long as there is (as there must be!) at least a best point. I suppose 'rational agents' should simply toss a coin between best points in such a situation, though in practice, choosing between equally good (though different) options can seem highly problematic.

(A structure may determine a choice-trivial situation in what we have called a one-off case even though it is not a maximally choice-trivial payoff structure).¹

Maximal triviality in an intuitive sense might seem to occur when a structure is (a) maximally choice trivial and (b) maximally goal trivial and (c) the intersection of the choices each agent should make is an absolute best point.²

We have already agreed, in effect, that a unique co-ordination equilibrium does not ensure that a situation is either maximally choice trivial, ~~or~~ even choice-trivial in a one-off case. Nor is it necessarily maximally goal-trivial. Hence it certainly need not be maximally trivial in an intuitive sense.

We may now consider the case of situations where each agent has a dominant choice. In what way is such dominance related to (intuitive) triviality? (Let us call the intersection of each player's dominant choice, where he has one, a point of universal dominance or u.d.) Situations with a point of u.d. need not be maximally trivial intuitively. First, even the choice-triviality of such situations is not always undubitable. Second, the goal triviality of such situations will vary.

¹ A possible example, to be discussed shortly, is the following:

	10		0
10		0	
	0		5
0		5	

² A further factor which we have not yet noted here, and which we shall generally for simplicity's sake ignore, in what follows, is what we might call 'differential triviality' or 'risk triviality', where there is not much to choose between some/any of the outcomes for some/any of the agents. Thus a choice-trivial situation lacking maximal goal triviality may still be relatively highly trivial intuitively. For the next best point for all may be what is reached by the best choice, and it may be nearly as good as the best point.

Regarding the first point, we can cite a case Lewis himself notes, that of the so-called 'Prisoner's Dilemma'.¹ Here the combination of dominant choices is, as Lewis says, a nonco-ordination equilibrium. Perhaps more to the point, its apparent 'choice-triviality' through dominance conflicts with the goals of the agents in an extreme way. Each player gets only his third best payoff (out of four) and relatively speaking this could be very bad .

In view of the conflict here, between 'dominant choices' ^{and goals} it might be argued (and has been by some) that we cannot speak of choice-triviality here, for it is really not obvious what the players should do.²

Cases with a u.d. point can however be more intuitively trivial than Lewis's Prisoner's Dilemma type cases. Prisoner's Dilemma cases may for instance be differentially quite trivial, and/or each might get his second best outcome at the intersection of dominant choices.³

¹ Lewis give this matrix

	-8	-10
-8	-1	
	-1	-2
-10	-2	

see Lewis pp.16-17.

² Cp. J.W.N.Watkins, 'Imperfect Rationality', in R. Borger and F.Cioffi, eds: 'I say that there is no optimal solution, that the idea of optimality breaks down here' (p.205). Luce and Raiffa, who believe that it is clear that agents should pursue their dominant strategy, observe ruefully that 'two "irrational" players will always fare better than two rational ones' (p.112).

³ See the following matrix:

	5	4	1
5	6	1	
	6	0	0
4	0	0	
	1	0	0
1	0	0	

What now, of cases of u.d., with a unique co-ordination equilibrium? First, if a situation has both of these features the u.d. point and the unique c.e. must be the same point. (For an example change the top left cell in the last matrix (in the footnote) to '7,7').

Now it seems clear that the high intuitive triviality of the matrix in such cases has nothing to do with the existence of a unique c.e. as such. I believe that in order to have a unique equilibrium in a situation of u.d. the co-ordination equilibrium must also be an absolute best point. If this is so then we have what seems to be (well-nigh) maximal triviality intuitively. For (a) it is obvious what each agent should do (choose his dominant strategy) and (b) the intersection of obvious choices reaches an absolute best point. However, while such situations are maximally trivial intuitively, their triviality seems not to stem from the existence of a unique co-ordination equilibrium and can be perspicuously characterized without reference to the existence of such a point.

Lewis effectively excludes cases with a unique co-ordination equilibrium together with a point of universal dominance, by introducing the requirement of at least two co-ordination equilibria in a co-ordination problem.

He points out, however, that trivial cases still remain, and this is undoubtedly true. His description of the trivial cases he cites is, I feel rather misleading, partly because it draws on assumptions we have now shown to be false. So let us consider these examples here.

First, the matrices (Lewis's figure 14, p.21):

	C1	C2
R1	1	1
R2	0	0

	C1	C2	C3
R1	1	1	.2
R2	1	1	.5
R3	0	0	0

Of these, Lewis writes:

'there is still no need for either agent to base his choice upon his expectation of the other's choice. There is no need for them to try for the same equilibrium - no need for co-ordination - since if they try for different equilibria, some equilibrium will nevertheless be reached' (p.21).

In fact, there is no need for either agent to try for an equilibrium, or to reason in terms of that concept, in either example.

In the left-hand matrix: one agent (Row) has a dominant choice, R1; and Row clearly should make that choice. Row does not in fact have to have complete knowledge of the matrix - only of his own payoffs relative to the different combinations of actions. Given Row's rationality and knowledge of his own payoffs alone, he can achieve a best point in complete ignorance of Col's desires or rationality, or of what Col will do. Row will know that he can do all this, just by considering his own payoffs and their circumstances. Thus this is choice-trivial in one-off and 'external expectations' cases. Moreover, whatever Row actually chooses, Col's 1 or 2 will both get him the same outcome. If Row is rational, both will reach a best point, if Row is irrational, neither will. All Col needs is to know his own payoffs, and he can see that it doesn't matter what he does, he might as well toss a coin. Thus neither Row nor Col needs to 'try for an equilibrium' as Lewis puts it. Row just needs to choose his own dominant strategy, Col to act at random; neither need know what the equilibrium points are, in order to see what he should do.¹

In the second case the existence of equilibria is again irrelevant. As long as each tries for a personal 'best point' (all of which are in fact

¹ (In this case it is of course true that if Row - on whom everything hinges - were to follow the maxim 'try for an equilibrium' he would effectively act correctly in terms of maximizing his payoff, and so would Col. - since he can act on any maxim).

equilibria) each will indeed reach a best point.¹ In fact in this second case each player has one or more strategies which is guaranteed in all circumstances to get him as good a result as any other strategy: R2 in Row's case, C1 and C2 in Col's case; (R2 is in fact a dominant strategy for Row). Each should surely follow his dominant or what we may call his 'quasi-dominant' strategy (where this choice and one other are equally good and dominate all others) and this strategy is obvious without a knowledge of the other agent's desires or rationality, or of what he will do. In the first case also, each has a quasi-dominant strategy. Yet the situations are still interestingly different, and the first may seem to be more trivial than this one. For here each one will only get a best point if the other acts rationally: no one, as Row can in the previous case, can guarantee himself a personal best point (or reach an equilibrium) just by acting rationally himself. Hence there is an asymmetry between these cases. Both of which, we may agree with Lewis, are intuitively quite trivial.

Lewis says of the examples just discussed:

'These cases exhibit another kind of triviality, akin to the triviality of a case with a unique co-ordination equilibrium'.

It is indeed true that in these particular cases (though not necessarily pace in all cases with a Lewis/unique c.e.) what each agent should do is obvious, and is the same, in both 'one-off' and other situations, and in so doing they each reach a best point at least. Hence these cases are highly trivial intuitively.

At this point Lewis introduces his requirement that there be two proper co-ordination equilibria. He writes - of a particular matrix, but presumably intending a general claim - 'the two proper co-ordination equilibria are sufficient to keep the problem non-trivial' (p.22) (The matrix involves two p.c.e. in the strong sense). Sufficient, but perhaps not necessary, Lewis notes:

¹ It is a consequence of my discussion here as a whole that the idea that anyone should ever try for an equilibrium per se, or for a co-ordination equilibrium, is not trivially obvious. See the section below on 'solving co-ordination problems' for a discussion of Lewis's proper co-ordination equilibria in this respect. It is true that the famous Von Neumann 'solution' to 2-person zero-sum games is in terms of equilibria of 'mixed strategies' but not all games are 2-person zero-sum, and even when they are not all equilibria are particularly desirable, even relatively desirable.

'we might prefer a weaker restriction that would not rule out matrices like those of Fig.16'(viz:

	C1	C2	C3
R1	1	0	0
R2	0	1	1
R3	0	1	1

Lewis does not state why we might prefer not to rule out the Fig.16 matrices. Clearly, however, the figure 16 matrices already lack one kind of triviality: that of universal (strict) dominance.. Lewis argues with respect to the previous matrix that 'the same situation can be represented by the new matrix in Figure 17, which does have 2 p.c.e.'s '

	C1	C2'
R1	1	0
R2'	0	1

This looks intuitively correct. For in effect the players will never have to choose between R2 and R3, C2 and C3, respectively. Each can adopt a random choice between these two, since either will do equally well in all circumstances. However, there are situations lacking intuitive triviality and also lacking a matrix obviously'reducible' to one with 2 p.c.e.'s (for either sense of 'p.c.e.').

For instance:

	3	0	0
3	0	0	
	0	3	4
0	4	2	
	0	0	0
0	0	0	

There is no u.d. point here in the sense defined, nor is the matrix obviously acceptably reducible to some matrix with 2 p.c.e. Thus if it is only some form of high intuitive triviality we wish to rule out, we will surely have to leave 'unreducible' matrices like the above in, and so the 2 p.c.e. requirement (in either the strong or the weak sense) is too strong.

Is Lewis's final requirement in fact too weak - are two p.c.e. in fact sufficient to keep the problem 'nontrivial' as Lewis claims?

Consider the following case:

	1	2
A	1000	0
	1000	0
B	0	.1
	0	.1

Here there are two p.c.e. in the strong sense - A1 and B2. A1 is an absolute best point, far superior to B2. All in all, intuitively speaking, if there is anything the players should do it is A1 respectively. Each has little to lose if anything goes wrong; no strategy of minimizing one's possible loss is left untouched.

This situation seems well-nigh¹ maximally trivial intuitively in a one-off situation.¹

This matrix is not, it is true, trivial irrespective of external expectations. Given a run of B2 for whatever reason, or a promise by Col (who has just gone mad) to do 2, it may be reasonable for Row to do B and Col to do 2. At least what each should do is now no longer obvious (even if Col knows his promise was a fake). However, this type of matrix is surely somewhat problematic for Lewis.

With regard to co-ordination problems, we have noted already that the feature of its 'mattering little' which of two or more alternatives are arrived at by all is a salient one, in at least the majority of the examples. Our sympathy with Lewis's whole idea that 'trivial' situations should be ruled not to be co-ordination problems could be based not so much on the fact that the examples are intuitively problem situations, but rather that, in a real one-off case, of a number of completely exclusive alternatives that would serve as solutions to the problem none is pre-eminently the best or obvious choice. However, merely requiring two p.c.e. - with or without the 'coincidence of interests predominates' clause - does not, it seems, exclude the possible existence of a point which is intuitively the obvious one for each agent to aim at, without prior agreement or external expectations of any kind, given common knowledge of rationality and of the matrix (assume there is no maxim strategy pulling against this choice).

¹ If one were to rule out as non-trivial any situation in which a player depends for getting his absolute best payoff on the other players' not acting rashly, this would have to be ruled out. But then some cases considered clearly trivial like:

1	1
0	0

are of this kind. Moreover common knowledge of rationality, which has been an assumption of our discussion (and of Lewis's, apparently) surely means that each player can depend on the others' not acting rashly.

A possible change with respect to this issue suggests itself swiftly; require that neither p.c.e. is an absolute best point. But, while this would rule our last example out, we may well not wish to rule cases with barely absolute best points out, as in the following:

		1	2
A		10	0
	10		2
B		2	11
	0		11

(This case is, indeed, far from trivial). I shall not at this juncture discuss further possible changes: suffice it to say here that 2 p.c.e., even in the strong sense, even given perfect coincidence of interests, maybe an insufficient or inadequate requirement if we do wish to rule out intuitive triviality in a one-off case and/or if we wish to capture a feature of clearly-similar co-ordination problems.¹

What, now, of the clause 'coincidence of interests predominates' in the definition of a co-ordination problem? We have noted that for Lewis it is important to be realistic enough not to require perfect coincidence of interest. This seems right, given the examples. Apart from this, Lewis does not attempt to justify his precisely defined 'coincidence' clause in the course of his game-theoretic argumentation nor is it clear that it picks out anything very clearly salient in the examples.

¹ With regard to the analysis of convention, Lewis believes that he does not need directly to rule out the 1000-1-0 matrix as part of the situation underlying convention. His reason: such situations are ruled out indirectly in his final definition of convention, by his 'common knowledge' clause. (See his discussion pp.73-4). I return to this issue later.

vi. Conclusions on the structure of co-ordination problems.

Lewis's 'deduction' of two proper co-ordination equilibria in the structure of co-ordination problems contains a number of flaws. His 'program' of progressively ruling out trivial situations suffers from the fact that he also appears not to want to rule out nontrivial cases, whereas many nontrivial cases are ruled out as he proceeds. In spite of this, even when we interpret 'proper co-ordination equilibrium' in the strong sense, intuitively trivial situations remain, given Lewis's final restriction on the structure of co-ordination problems. Lewis rightly, it would seem, does not want to insist on perfect coincidence of interest among agents in co-ordination problems. However, generalisations over 'mixed' cases in game theory are notoriously hard to come by.

Looking back, now, at Lewis's examples of co-ordination problems, it does look as if (a) they all at least have two or more proper co-ordination equilibria (in the weak sense). In fact (at least) most of them appear to involve two or more proper co-ordination equilibria in the strong sense - it is in everyone's interest that as many as possible conform, given that most do.

Lewis may, then, have hit upon the most striking salient feature of a class of (intuitively quite similar) problems with his two-or-more-p.c.e. requirement. As I have argued¹, from this point of view a stipulation that there be two-or-more p.c.e. in the strong sense could be justified in this way, to fix ideas.

4. Lewis on 'convention': a critical commentary.

i. Conventions as 'solutions' to co-ordination problems.

A. Proper co-ordination equilibria as 'solutions': a caveat.

At the beginning of the section of Convention¹ entitled 'solving co-ordination problems', Lewis writes:

'Agents confronted by a co-ordination problem may or may not succeed in each acting so that they reach one of the possible co-ordination equilibria' (24).

¹ Ch.1, S.3.

This provokes the question: in what way, if any, is a proper co-ordination equilibrium (surely Lewis meant 'proper' to be understood here) something rational agents will always want to reach if possible? We have already reached a negative conclusion about equilibria in general and co-ordination equilibria in particular: these need not be absolute best points, so at least one agent may prefer to end up with some other combination of actions. Clearly, some p.c.e.'s will be absolute best points. They will then be combinations of actions that everyone would most like to reach. Yet there may in fact be points in a Lewisian co-ordination problem which are better for all the agents involved, than any proper co-ordination equilibrium, as in the following matrix:

	1	2	3	4
A	6 6	5 5	5 5	5 5
B	5 5	7 7	6 6	6 6
C	5 5	6 6	90 89	89 90
D	5 5	6 6	7 6	7 6

Here there are two p.c.e., at A1 and B2, but both C3 and C4 (neither of them p.c.e.) are superior to either p.c.e., for each agent.

Further, we cannot argue that, their ideal apart, the agents should always do their part in a p.c.e., when they are in a co-ordination problem, since in a matrix like the above there is no reason to suppose rational agents should do this. (If anything, it would seem that Row should do C, and Col choose 3, and that neither should do his part in either p.c.e.).

Ignoring cases like the preceding for a moment, the fact remains that agents will not always be neutral between p.c.e. in a co-ordination problem, just imagine situations where one (and therefore only one) of the p.c.e. is an absolute best point. The other will then not be a point that all want to reach if possible. Nor need it be second best as the following matrix shows:

	1	2	3
A	50	49	2
B	2	-100	3
C	49	0	-100

(Here the two p.c.e. are A1 and B3; there is perfect coincidence of interests among the players). Hence it is wrong to suppose that agents in a co-ordination problem, as defined by Lewis, will always prefer some proper co-ordination equilibrium to any other point. With this caveat, we may proceed.

B. Conventions as 'solutions' to co-ordination problems.

Lewis reaches his first, rough definition of 'convention' by way of considering how rational agents in a co-ordination problem may end up doing their part in one proper co-ordination equilibrium in a co-ordination problem.

Let us consider an example. X and Y are driving in opposite directions on a winding two-lane highway (perhaps they have got to the point where each knows the other is approaching, but neither has yet seen the other). Each wants to drive on the right if the other does, but on the left if the other drives on the left. Each shares the other's preferences in this respect. It matters little to either in which lane he drives, as long as both drive in the same way. Each must choose how to drive.

In this case it is quite reasonable to assume that the only alternative to each one's driving on the left, or each one's driving on the right, is one's driving on the left and the other's driving on the right. It is also reasonable to assume that both prefer either of the former alternatives to the latter one (indeed they far prefer them to this one). This seems to give us a case of the following kind:- there are two proper co-ordination equilibria (strong sense) and, apart from that, only options which yield lower payoffs than either p.c.e. for each player. In such a case it is clear why the agents involved should positively prefer to 'act so that they reach one of the possible(proper) co-ordination equilibria.'¹

Now suppose X in our example is a rational agent, and is virtually certain that Y will drive on the left; then, given his own preferences, he has a decisive reason for driving on the left himself. So X will drive on the left. The same goes for Y, mutatis mutandis. So they will indeed reach a proper co-ordination equilibrium given the relevant (and sufficiently strong) expectations about each other's action.

One way of acquiring justified expectations about action in a co-ordination problem involves expectations of a different sort. Lewis distinguishes between first-order expectations about things (ordinary expectations about them) and higher-order expectations:

'An ($\underline{n} + 1$)th-order expectation about something ($\underline{n} =$ or greater than 1) is an ordinary expectation about someone else's $\underline{n}$ th-order expectation about it' (28).

Suppose X has reason to believe that Y wishes to drive on the left if X does. Then, if

¹ One might consider defining 'co-ordination problem' so that such situations were all, by definition, of this form. Sometimes Lewis himself seems implicitly to do this. (See for instance, p.73, where he imagines a (proper) co-ordination equilibrium which is 'only slightly preferred to some of the outcomes which are not (proper) co-ordination equilibria'). He never explicitly makes such a restriction, however, and our critique has proceeded and will proceed accordingly.

X also has reason to believe that Y expects X to drive on the left, X will have reason to expect Y to drive on the left. X's justified expectation about Y's action, then, may be produced in part via X's second-order expectation about his own action. Lewis generalizes his point as follows:

'Circumstances that will help to solve a co-ordination problem... are circumstances in which the agents become justified in forming mutual expectations belonging to a concordant system' (33).

How, then, might concordant mutual expectations of first and high orders be formed? One highly effective method is an explicit verbal agreement. Suppose X and Y are the only drivers on a private estate, and they agree to drive on the left, in case their paths cross. Each thereby gives evidence of his preference for driving the way the other drives, and of his own intention to drive on the left. 'By observing this evidence, we form first-order expectations about each other's preferences and actions. By observing each other observing it, we may also form concordant second-order expectations' (34). And so on, perhaps, for higher orders of expectation.

But Lewis wishes to stress that agreement is not the only way of producing a system of concordant mutual expectations to solve a co-ordination problem. Given that X and Y have by some means acquired concordant expectations concerning one another's preferences, what seems to be crucial is some mechanism for singling out one of the possible solutions so that it is conspicuously unique 'in some way the subjects will notice, expect each other to notice, and so on' (35). Agreement is just one such mechanism. One thing Lewis wishes to stress is that salience of a solution may occur without talk or gestural communication of any kind.

A possibility that Lewis does not draw our attention to is the following. In a case like that of X and Y each might in fact somewhat prefer driving on the left to driving on the right, so that each one's driving on the left will here provide an absolute best point. Here is a possible source of salience without

communication: the 'matrix' of preferences themselves.

Lewis himself stresses precedent as a source of salience-without-communication. He depicts the following type of case of co-ordination by precedent: knowledge of a number of past cases of a given type of problem situation S among members of a population P gives each of us (who are members of P) reason to believe that in (almost) all past cases of a given type of problem situation S, involving members of P, the problem has been solved by doing (an action of type) A. Doing A thus becomes the salient solution when we are agents in the relevant problem in P; so we do A, and achieve co-ordination by precedent. Once this starts happening,

'we have a metastable self-perpetuating system of preferences, expectations, and actions, capable of persisting indefinitely... This is the phenomenon I shall call convention' (42).

He then presents his first, rough definition of convention (quoted in Section 3,(i) above).

ii. On conventions started by agreement: a possible amendment.

We noted at the outset of this chapter a problem with Lewis's type of account in connection with our everyday notion of convention. That is, Lewis's definiendum involves only what we called L-regularities.¹ We noted also that we actually speak of the existence of conventions where there is no corresponding L-regularity. A good example here is where you and I have an agreement to do something in certain circumstances, say to drive on the left on our island, but the circumstances do not arise, say our vehicles never arrive.

Lewis does consider agreement in relation to convention (see especially pp.83-8). Thus he says, for instance that '... the exchange of declarations of intent in a face-to-face meeting would be sufficient to start a convention' (p.85). According to my intuitions, then, and possibly according to Lewis's also, full generality with respect to our everyday concept requires an account of convention which allows that conventions may obtain in a population in the absence of a related L-regularity.

¹ See Section 2, (i) above.

It seems that it would not require much tinkering with Lewis's account to rectify matters in this respect. However, a simple change we might make points up the need for an amendment of another kind.

Suppose we rephrase (and re-interpret) Lewis's definiendum somewhat as follows:

'A regularity in action in a situation of kind S is a
convention in population P...'

no longer stipulating S's recurrence. As it turns out, we need not change Lewis's definiens. Thus consider the first condition ('in any instance of S among members of P, everyone conforms to R'). One might think that Lewis required, for consistency, an additional 'and S occurs fairly often' as a counterpart to the phrase 'a recurrent situation S' in the original definiendum. But I expect he intended us to construe 'everyone conforms to R' as 'everyone conforms to R (where R is a regularity-in-the-behaviour-of-a-member-of-P-in-a-recurrent-situation-S)'. If we amend the definiendum as suggested, the first condition will be construed in a 'lawlike' way: so that it supports counterfactual and subjunctive conditional claims. The remaining conditions will be construed similarly.

Now, at this point one might feel that the elimination of any reference to actual doings in the definiens presents a problem. I take it that Lewis intends us to see the phenomenon of convention in at least the following way: conformity to a certain regularity is the result of expectations and preferences of the kind mentioned in the second and third conditions of the definiens. But justified expectation of others' conformity may have a variety of sources. In particular, Lewis has stressed widespread precedent as giving rise to expectations of conformity and hence to (further) conformity. However, the 'conformity condition', as we now construe it, says nothing about the existence of actual doings of some action A in situation S, and thus no longer gives us - even by

implication as opposed to stipulation - any specific basis for the expectations involved in convention. Thus the question arises whether something about the specific source of expectations should be written in to Lewis's definition of convention.¹

We may wonder whether there are sources of salience and hence of mutual expectations in a co-ordination problem which cannot in themselves give rise to convention - at least according to our own everyday notion. It seems that salience from within the 'matrix' of preferences themselves is one such source. Consider a case: it happens that cars and/or roads where X and Y drive are so constructed that it is easier to drive on the right than on the left. There are no agreements or precedents about driving lanes, but driving on the right is the salient solution to their problem by virtue of being the easiest solution to conform to, and hence by virtue of its being obvious that each would prefer that that option were chosen. So X and Y reasonably expect the other to drive on the right. Suppose they fulfill Lewis's other conditions for being a convention-population. Do they have a convention? I am certainly inclined to say of X and Y that they have no convention about driving lanes: we will then want this case ruled out if we want to capture 'our' concept of convention. So - if we start from Lewis's definition, and particularly if we amend it as suggested, it seems we will want some restriction on the source of salience and hence of expectations of conformity in convention-populations.²

¹ Cf. S. Schiffer's account of convention in Meaning (pp. 136-55). I am indebted to Schiffer for indicating the importance of this question.

² Such a restriction could be put in terms of some general characterization of the type of source required (for instance, that there has been some form of public adoption of a solution; compare Schiffer p. 130, op. cit.) Or perhaps we should rather list the permissible sources: for instance, precedent and agreement. In addition, a public decision by someone authorised to decide such issues should probably be included (perhaps something like this is what Schiffer has in mind when, in giving such a list, he included agreement, precedent, and stipulation).

Moreover, if Lewis does not wish to allow in 'vacuous regularities' of the kind created by agreement, it seems that his own restricted conception of conventions will be better articulated if he explicitly stipulates that precedent is the only permissible source of expectations of conformity to a convention.

Having noted the case for amending the initial rough definition in a way Lewis himself does not consider, we should turn to the major refinement Lewis does make to his original definiens: it must not only be true, but also common knowledge in the relevant population P that the (original) conditions for the existence of convention be satisfied.

iii. Lewis's 'common knowledge' condition.

In Convention, as we have noted, Lewis does not define common knowledge in terms of an infinite set of conditions, as we have, but rather in terms of a finite set.¹

Lewis's dealings with the notion are simplest in 'Languages and Language'.

There he says:

'The various facts...are matters of common (or mutual) knowledge: they are known to everyone, and it is known to everyone that they are known to everyone, and so on' (p.6).²

¹ Lewis first defines a notion of indication: a state of affairs A 'indicates to someone x that ___ if and only if, if x had reason to believe...A held, x would thereby have reason to believe that ___' (pp.52-3). It is 'common knowledge' in a population P that ___ if and only if some state of affairs A holds such that (1) Everyone in P has reason to believe that A holds, (2) A indicates to everyone in P that everyone in P has reason to believe that A holds, (3) A indicates to everyone in P that ___' (p.56).

² Lewis adds that the knowledge mentioned may be 'merely potential: knowledge that would be available if one bothered to think hard enough'. In my view such 'potential knowledge' probably is knowledge simpliciter (see our discussion of knowledge in Chapter 4). Lewis also contemplates a negative version of the quoted account which is surely too weak for most purposes (as we argued in Chapter 4).

As I noted in the last chapter, given my own notion of group common knowledge it trivially follows that everyone in a population P where there is such knowledge has the concept of a member of P. For that is involved in having the knowledge that everyone in P... It is true in such a case that 'P' might be simply 'the population consisting of Tom, Dick, and Harry' or even '...of him, him and him', and so on. In other words the 'concept of a member of P' here does not have to involve what would naturally be thought of as the concept of a certain property or feature members have in common. Nonetheless, something we can call the 'concept of a member of P' is surely an essential prerequisite of Lewis's 'Languages and Language' account of common knowledge, as of my own.

Now at one point in Convention Lewis claims that a convention might hold in a population with 'no general concept' of a member of P (see pp.64-8). I am not quite clear what Lewis means by 'general concept' here. If he means to claim that a convention-population, P, need not have a word for 'member of P', he is surely right. Otherwise I am inclined to think he is wrong.

Now Lewis makes his point in Convention by reference to an imaginary population of strange creatures:

'... a convention R regarding action in S might hold in a population P of creatures...(who) learn from experience not by coming to believe generalizations, but by acquiring propensities to come up with the right particular beliefs regarding any new case that is presented in sufficient detail'(p.67).

In particular, given an instance of S in this population, each creature:

'knows how each of his fellows would act therein...and he knows that they do so by convention; that is, given any of the defining conditions of convention as applied to a given agent in a given situation, he knows that the condition is satisfied. But he cannot think of more than any one instance - the given one - at a time. He has no general concept of an instance of S, of a member of P, or of an action in conformity to R'(p.67).

How are we, though, to understand these creatures? Note the condition that these creatures 'come up with the right particular beliefs regarding any new case that is presented in sufficient detail' (my stress). How much detail is sufficient, and sufficient for what? Ex hypothesi these creatures do not recognise members of P as members of P. Sufficient to get the right response from

P-members? This will surely make the account *vacuous*. Perhaps 'sufficient to be recognised as members of P by those (not members of P, in fact) who do have the concept of a member of P'. But what is in that case sufficient? And if one can say what is in that case sufficient, how does it help us to understand or describe these creatures who ex hypothesi lack the ('general') concept of a member of P?

It might be suggested that Lewis is here contemplating a definition of 'convention' in terms of what we have earlier called 'individual' rather than 'group' common knowledge. The arguments against doing this have been rehearsed with respect to collective practice in general, in the last chapter. In any case Lewis does not stipulate that there is, in the relevant population P, individual common knowledge that anything. The members simply (after 'exposure' to precedents) come to have the right particular expectations of the individuals in various subsets of the set of P-members, and it was not these mini-populations who were, at the outset, supposed to have a convention, but the population P itself. [It seems to me unhelpful to define or to try to explain the notion of convention in these (obscure) terms, and better to use a definition in terms of group common knowledge, and to say that unless these creatures have the concept of a member of P we cannot say they have a convention (or indeed understand what they 'have'). In what follows it will not affect matters if, for the sake of argument, we assume the common knowledge of Convention is defined in these terms I prefer.

Lewis's common knowledge requirement is quite sweeping, for it is not just the regularity R whose generality in P must be common knowledge, but the reasons underlying it also. Lewis gives a number of justifications for this requirement.

Lewis claims on behalf of his requirement that, for one,¹ it usefully deals with certain 'odd cases' which would have qualified as conventions under the original definition (p.59). The cases he mentions concern agents who fulfil the

¹ A second justification is dealt with in the next section.

original conditions but hold a certain false belief(f) about each other, e.g. each believes that the others would go on driving on the left no matter what they expected the rest to do. Lewis believes that these cases should be excluded. They are 'intuitively unlike clear cases of convention' (ibid). Lewis does not spell out examples of such cases, so it is difficult to be sure that one's intuitions agree with Lewis's. He notes that if the common knowledge condition is fulfilled, then there is some state of affairs which the agents have reason to believe holds, and which would enable them to conclude (given their own inductive standards and background knowledge) that f is false. Thus only if they are 'extremely irrational' will they persist in holding f. Let us leave aside, then, agents who would be extremely irrational to believe f. There may remain cases in which the original conditions are fulfilled, yet it is reasonable to believe f, in which common knowledge (albeit potential) of f's falsity isn't available. Can we construct such cases, and is it obvious that we should rule them out, as the common knowledge condition will (and without the aid of principle dismissing extreme irrationality)?

Examples of such cases are, I think, bound to be far-fetched. Each member of P is going to have to see the others as being very different from himself, and be in circumstances which make this a not unreasonable belief-while at the same time it is a false belief. The following might do as an example. There are two tribes, and it is common knowledge between them that each tribe has been given, as a gift from a friendly foreign power, a car together with a robot driving device which is programmed to drive in the left lane only. The head man in each tribe finds that he enjoys driving enormously, and opts to drive the car himself on all its journeys. He realizes that this is somewhat silly of him, since he could be letting the device get on with the job, but he decides to indulge himself. Having no reason to think that the other head man will take such great pleasure in driving, he reasonably supposes that the other tribe will use the robot device. What happens, then, is that both head men drive on the left, and so on, while believing that the driver of the other car is just a robot, with no expectations or preferences of any sort, and which would drive on the left

come what may. Taking as our population the drivers on the highway (assuming no one but the tribesmen drive), they seem to fulfil the original conditions for convention, but also to believe reasonably enough that the other members of the population do not.

My linguistic intuitions certainly have it that the drivers in the 'robot' case do not have a convention to drive on the left. The case may well be over-determined. But the common knowledge condition does give an acceptable result here.

Lewis claims that the common knowledge condition specifies an important feature of conventions(59). Recently he has argued, on behalf of the condition, that it'...ensures stability. If anyone tries to replicate another's reasoning, perhaps including that other's replication of his own reasoning...the result will reinforce rather than subvert his expectations of conformity to R' ('Languages and Language' p.6). Now in the Robot case, the false (but justified) beliefs about other drivers' inflexibility appears to make reason-replication quite redundant, or rather, it makes the attempt to replicate reasons quite unintelligible. But where such an attempt is intelligible, perhaps without limit in levels of replication, then the common knowledge condition does clearly ensure stability in the way Lewis describes.

The fulfilment of Lewis's common knowledge requirement is, therefore, an important feature of the cases in which it occurs. For if it occurs, then the agents concerned have common knowledge that they are agents in a particular co-ordination problem (see original condition 3). Given that they also have common knowledge ^{of one another's rationality, then common knowledge} that everyone expects everyone else to conform to R ensures that each one's indefinite (and intelligible) reason-replication will reinforce rather than subvert his expectation of conformity to R. Our last example involved agents in a co-ordination problem without common knowledge that they were in this situation. But examples of this sort are importantly different from those where there is such common knowledge: in the latter case further items of

common knowledge are 'stabilizing' additions. There is thus a specifiable reason to distinguish cases here, linguistic or other intuitions apart.¹

iv. The 'arbitrariness' of conventions.

In his Foreword to Convention, W.V.O. Quine writes:

'the keynote of conventionality is a certain indifference: the syllable "big" could have meant "small" for all we care... black ties could have been counted less formal than fancy ones... Such is the initial intuition, but the appropriate sense of indifference, or of "could have meant", needs a lot of refining.' (p.xii)

In this section I argue that Lewis does not clearly capture 'initial intuitions' about the arbitrariness of convention with his proposed conditions on convention; these seem to be too weak for this purpose. I later suggest in particular the irrelevance of Lewis's concept of convention to the 'could have meant' of language that Quine refers to.²

Conventions, according to Lewis, have 'alternatives' which are known to be such. To be more specific, for an L-regularity R to be a 'convention', on Lewis's view, it is necessary (1) that those who conform to R be able ~~not~~ to; it is also necessary (2) that there be some alternative regularity R' which each one would have wished to conform to had the others done so (thus doing his part in a proper co-ordination equilibrium). Further (3) each must be able (at least after practice) to conform to R' (Lewis, p.50). (4) It must be common knowledge in P that conditions (2) and (3) are fulfilled.

¹ The common knowledge condition may be amalgamated with the amendments suggested earlier. This is, as far as our everyday concept is concerned (1) we don't want to insist that S be recurrent or that there be common knowledge that S is recurrent. Yet (2) we do want some restriction of basis for expectations, to rule out cases like the uneven road case. We can, then, alter the definiens to include something like: (It is common knowledge that) (1) there is a precedent or agreement in P to conform to R (in S); (2) partly on the basis of common knowledge of this precedent or agreement, everyone expects everyone else to conform to R...(cp. Schiffer, p.154).

² See section (vi) below.

Now, Lewis sees here a further justification for his common knowledge condition (see p.73-6). He first indicates an intuition we may have about the applicability of 'convention'. Suppose the (proper) co-ordination equilibrium we would reach by conforming to R' is much worse than the one we reach by conforming to R, then

'do we really want to call R' a possible alternative convention?
And do we want to say that R' contributes to the arbitrariness and conventionality of R?' (73).

Lewis gives a tentative answer: 'perhaps not'. He goes on to say that his definition of convention, as it stands, is 'likely to disqualify this R' as an alternative to R'. For everyone's preference that everyone conforms to R, if all but one person does, may not be an item of common knowledge. He later writes, more strongly:

'by means of our common knowledge requirement...we can exclude (alternatives too unsatisfactory to contribute to the arbitrariness of a convention)...without doing so ad hoc' (76).

I would argue, however, that Lewis's common knowledge requirement is not an adequate way of ruling out 'alternatives' insufficiently satisfactory to contribute to the arbitrariness of a practice. For we can imagine cases in which albeit very unpromising alternatives are known to be available and to be 'preferred' in the manner (rather minimal, as we have seen) of all proper co-ordination equilibria. Perhaps there has been a lengthy public debate and discussion on the subject, so it is common knowledge that this proper co-ordination equilibrium exists. It can surely not be argued that alternatives brought to light in this manner would thereby no longer be sufficiently unsatisfactory to contribute to the arbitrariness of the practice actually adopted. In other words, I suggest that if any practices are, intuitively, sufficiently unsatisfactory to endow an alternative practice with 'arbitrariness', then they may still be sufficiently unsatisfactory even when their nature and existence is common knowledge. We can imagine the debaters noticing somewhat jocularly - 'Well, we can ignore that alternative; in fact, its obvious what we should do'. If they do indeed decide to do what is obvious, in full knowledge of the (rather ludicrously awful) alternative, I suggest that their ensuing

practice is no more 'arbitrary' than one decided upon without this (ultimately irrelevant) piece of knowledge. What may be crucial for convention is not the parties' knowledge or lack of knowledge of the existence of other p.c.e.-fulfilling regularities, but rather their perception that there are some viable alternatives, that is, alternative modes of behaviour which it might be reasonable to adopt. Thus if it is apparent to the parties concerned that 'they would each prefer to conform to R' if everyone else did' but it is also apparent that R' really would not do - 'we would be mad to adopt it' - then pace Lewis their agreement to conform to R - the only known alternative to R' - would not be the acceptance of a convention. This of course indicates that something is wrong with Lewis's final account. And though some restriction on the underlying payoff structures claimed for convention might help, it is not clear that this is the direction one should take in analysing convention.¹

v. Conventions as norms.

Lewis does not stress but does hold that conventions are 'norms', that is 'regularities to which we believe one ought to conform' (97). Lewis argues that this follows from his proposed account of convention. In most cases in which people decide to conform to a convention the agent automatically has two reasons

¹ I have in mind the suggestion that, say, we stipulate that (a) the set of p.c.e. in a co-ordination problem is such that no non-p.c.e. point is preferable to any p.c.e. point for any agent, and (b) the p.c.e. are identical in value for each agent. (As we have seen, one p.c.e. may be far worse than another). However as far as this suggestion goes (b) seems too strong intuitively with respect to convention. Perhaps Lewis would have done better (if more vaguely)-with regard to arbitrariness had he defined 'co-ordination problem' by first introducing a matrix that gave an indisputably apt payoff structure (say where (a) and (b), above, hold) and then saying something like 'a co-ordination problem is a situation so like the preceding as to involve no practical difference'. However, it is not my aim to redefine 'co-ordination problem' here.

for conformity: doing so answers, first, to his own preferences, and, second, to the preferences of most others involved. Moreover, Lewis claims, conventions are also, by definition, 'socially enforced norms': 'one is expected to conform and failure to conform tends to evoke unfavourable responses from others' (p.99).

Lewis's remarks here bring us back to an issue raised earlier—that of the exact nature of a proper co-ordination equilibrium, which was not explicitly defined. So far we have noted some reason in the nature of Lewis's examples for opting for the strong sense of 'p.c.e.'

Here we may have another. Does Lewis think it is important that in convention each one's conformity must answer not only to his own preferences but also to those of most others involved? If so, this will be a reason for preferring the strong rather than the weak p.c.e. requirement in the first place.¹

Their normative aspect will surely be an important feature of conventions, and the question arises: might our 'convention' not be most nearly defined by reference to this feature as opposed to the complex of features Lewis makes central.² We shall return to the normative aspect of conventions later in this chapter.

vi. Conformism

We have so far ignored an obvious cause for objection to Lewis's account, qua account of the everyday concept of convention; that is the suspicion that we ordinarily call certain regularities 'conventions' though they do not involve

¹ Other remarks of Lewis's indicate a preference for (or at least an assumption of) the stronger requirement. Thus consider his discussion, pp.45-6, of people who want to 'have a differently dressed minority to sneer at' at parties, as opposed to those who 'want everyone else to dress like the majority too'. Lewis claims that according to his first rough definition of 'convention' only the latter have a convention; they alone find everyone's dressing alike a proper co-ordination equilibrium. (The nasty party-goers, as Lewis notes, do not even find it a co-ordination equilibrium). In fact, it is not at all clear to me that a population of nasty party-goers who are nasty in Lewis's manner cannot have conventions about how one should dress.

² Cf. Weber, op.cit, p.127.

Lewis's type of preference structure.¹ In order to think clearly about this, we should first note an important feature of Lewis's conditions which may tend to go unnoticed - it is not something he himself draws attention to.

We should note that preferences satisfying Lewis's preference conditions seem to fall into two kinds: preferences for conformity as an end in itself, and preferences for conformity as a means to some further end, e.g. sustaining life. Thus often the members of a population will have some particular desire which demands general conformity to one or another of some set of regularities for its achievement. For instance, their wish to avoid death while driving requires that everyone drive in the same lane, no matter which. But it seems that a rather different type of desire may also generate conventions.

Consider the following, which we may call the conformity preference (c.p): X prefers any one more to conform to regularity R in an instance of S if most others in P conform to R then. Now suppose that it is true and common knowledge in P that (a) the c.p. operates throughout P; and (b) there is a widespread precedent in P for conforming to R in instances of S (perhaps this happened by chance, perhaps there was some unconscious copying of one person by another). Let R be the regularity of holding the fork in the right hand when eating from the fork. Now suppose some members of P are sitting down to dinner and wonder how to go about eating it. They don't mind what they do as long as it accords with what, if anything, the majority do. On the basis of precedent it is common knowledge that if there is anything that most people do, it is likely to be this. Thus it is common knowledge that everyone expects everyone else (or almost everyone else) to conform to R, and hence, given the c.p., common knowledge that everyone will conform to R in this instance of S. As time passes, and cases pile up, there will be common knowledge in P that, in any instance of S, members conform to R, and

¹ Cf. S. Schiffer: 'most of the conventions of etiquette do not seem to involve co-ordination problems' (Meaning p.152).

so on. Conformity to the regularity of holding the fork in the right hand when eating from the fork thus will, it appears, become a Lewisian convention among these people.

For such people, every situation presents a co-ordination problem: they want to co-ordinate their actions with those of others, so that they do 'what is done'. For these conformists, if most people are doing some thing, for no matter what reason, everyone prefers that as many as possible do it. We can see why what is done becomes a 'norm' for them, something which ought to be done. (For Lewis on conventions as norms, see pp.97-100). Moreover, we can see how these people's conventions have an extra kind of arbitrariness. Wherever and whenever most people act in a similar way, wherever widespread precedent or agreement is found, convention is derived from it. Thus we might find many apparently quite pointless conventions among them - they might all clap their hands three times an hour, for instance. Where more than mere conformity is at issue, the range of acceptable regularities will be restricted by whatever further factor is involved. To go back to conventions of etiquette, we can see that although they may solve no very specific prior problem, they may still be Lewisian conventions if for instance they are supported by the simple desire to do whatever others do. Whether particular conventions of etiquette, etc, are Lewisian conventions is of course, an empirical matter; at some point, however, one must raise the question whether there is a logical relation between Lewis's preference condition and conventions as ordinarily conceived. We shall shortly consider some alternative accounts of convention, where no such logical tie is stipulated. But first we should briefly consider the application of Lewis's concept as we have it to the important phenomenon of language use and the 'could have meant' of language.

vii. Conventions and language.

The claim about convention which most interests Lewis in Convention concerns language. He was initially concerned to show that our concept of convention 'permits language to be conventional' (2). Later he goes further, proposing a

definition of the notion of a population's actual language in terms of convention. He conjectures that his definition 'agrees with ordinary usage in clear cases, and draws a convenient line among unclear cases' (194). Before concluding our critical summary of Lewis's work on convention, I shall say something (briefly) about Lewis on 'actual languages'; what I say will be quite inadequate to a subject which merits much attention.

Lewis first gives an account of a 'possible language', 'in abstraction from any users it might happen to have' (161). The sentences of possible languages have semantic properties: in particular, an account of analyticity in a possible language is given in terms of possible worlds. Lewis next asks what makes a particular possible language the actual language of a population: what is it for a population to use a language in the relevant way? His claim, in Convention, is that for L to be its actual language a population must have a convention of truthfulness in L.

Lewis has recently modified the account of actual languages, and, consequently, the account of convention itself, to meet certain objections. Language use is now defined in terms of a convention of truthfulness and trust.¹

A possible objection involves the specification of the convention in terms of truthfulness (and trust). Might not some population of storytellers, for instance, whiling away their time in Valhalla, never try to tell the truth in their language - never try to tell the truth in language L, or credit others with doing so, and yet still utter and respond to sentences of L in such a way that we would naturally say L was their language? In Convention Lewis relegates such cases to the limbo of bizarre counterexamples. In 'Languages and Language' he has suggested ways of dealing with them, offering various suggestions, though not adjudicating between them. One suggestion introduce the notion of a 'serious

¹ The basic change in the definition of convention allows that conventions may be regularities in action or in action and belief. See Lewis ('Languages and Language'); see also Bennett, 'The Meaning-Norminalist Strategy' in *Foundations of Language*, 1973.

communication situation' and specifies the convention whereby P uses L as a convention of truthfulness and trust in serious communication situations. Lewis does not say whether he assumes that such situations must sometimes occur in P, if the members of P are to be said to use a language. One who thought this assumption was false might wonder whether truthfulness was then given an artificially central place in the account of actual languages. One would like to see this aspect of Lewis's theory elaborated and discussed in more detail.

Then there is the question whether the notion of an actual language is to be defined in terms of some species of Lewisian convention. A regularity of truthfulness and trust in speaking and hearing a given language is not necessarily a convention: whether it is depends on a number of factors. So even if truthfulness and trust in L is necessary to the use of L, conventionality of truthfulness and trust in L may not be, and there are reasons for doubting that it is - according to whatever intuitive concept we may have - or that it should be made so. In particular, surely the case of creatures who are 'unable to use any language other than the one that is built into them' (195) should not be dismissed briefly as a bizarre counterexample. It is not clear that this case is so different from 'language use as we know it' (ibid). To say that it is rather begs the question in favour of a Lewisian conception of language use. We might ask what is common to such a case and the case of those who can use several languages. A concept of language use covering both cases might be more useful than what Lewis gives us. Perhaps it might serve us when we talk about computers as well as humans; it would leave the conventionality of human language use an empirical matter.¹

¹ Cp. Chomsky's remark that a Martian's best hypothesis on meeting with a human linguistic community would be that of the innateness of the whole language (John Locke lecture, Oxford, 1969; referred to in Stich, 'Grammar, Psychology and Indeterminacy').

It is worth noting that the arbitrariness Lewis would attribute to all actual languages encompasses at least three kinds. One kind derives from the fact that words in particular possible languages need not be appropriate in any obvious way to their particular meanings. They need not be onomatopoeic, for instance; the notion of onomatopoeia hardly makes sense for all types of term, e.g. a copula. Secondly, given that language use allows us to fulfil (perhaps to have) certain purposes (e.g. to tell jokes, impart information) use of more than one possible language is likely to serve these purposes equally well, so that it matters little which one we happen to use, as long as we use one of a certain range. Thirdly, the arbitrariness of the language adopted may derive from the fact that not only would some other language have done as well, given our purposes, but we have the capacity (at least at some stage in our lives) to use it. The first two kinds of arbitrariness (not just the first) could remain for the languages of inflexible language users. The third kind could not. This kind of arbitrariness is essential to Lewisian conventions; but it is the first which seem to stem from the nature of languages and language in use in general rather than from facts about a certain kind of language-using creature, possibly not the only such creature.

One might, then, leave Lewis's discussion of language, accepting that he has provided a concept of convention which permits a population's language use to be a matter of convention; but doubting whether it is such that a population's language use must be conventional in his sense.

viii. Some conclusions on Lewis's concept.

We may now draw some general conclusions on Lewis's concept of convention, before turning to some rival proposals. Lewis does not set out to analyse all uses of 'there is a convention in population P to do A in S', or, if he does, he fails. For ^{one,} _^conventions may exist without the corresponding L-regularity obtaining in the relevant population. A number of ^{other} _^discrepancies have been cited too (and more will follow in the next section). However, ordinary language is not the only

yardstick by which Lewis may be judged.

Lewis's concept embodies an account of how a regularity R in action may persist throughout a certain population; the explanation may lie, roughly, in this: (a) each individual expects others to conform to R and (b) he prefers to conform to R if they do. Hence the conformity to R may persist indefinitely. An explanation of the fact that different regularities may appear and persist in different populations in the same situation is also implicit: more than one possible regularity is of the relevant kind, i.e., conformity by all is a proper co-ordination equilibrium (in the weak sense).

There is surely little doubt that Lewisian conventions are widespread throughout actual human groups. Indeed, wherever creatures of a certain kind converge for any length of time, we would expect Lewisian conventions to emerge: creatures, that is, with certain preferences for co-ordination, with freewill, and without any miraculous ability to co-ordinate their actions from moment to moment by sheer instinct, or any helpful, guiding natural features of the environment. We would expect, in brief, to find Lewisian conventions in human societies.

Lewis's notion of a convention is a complex concept, with many components. Though there are obscurities and mistaken claims in the course of Lewis's exposition, and though he does not seem to have captured the most general everyday concept of convention, he has (perhaps unwittingly) undoubtedly set a new standard of analytical reach and precision for the philosophy of social science. The number of people who have recently borrowed from him or been influenced by him, as well as criticising his work, provides some indication of this.¹

¹ These include S. Schiffer, E. Ullman-Margalit, Jonathan Bennett, Jane Heal, Christopher Peacocke, and the present author.

5. Other notions of convention.

i. Critics of Lewis

Critics of Lewis, among others, have offered or indicated alternative accounts of our notion of convention. In this section we consider two such accounts, by Tyler Burge and Steven Schiffer. Criticisms of these views will lead us to propose a rather different type of account of conventions in general, most akin, perhaps, to that given by Weber.

There is some consensus among published critics of Lewis that his definition of convention is too strong as far as ordinary construals of 'convention' go. Criticism has focused on the preference requirement, which is the core of Lewis's analysis.

Thus Tyler Burge (Philosophical Review, 1975) argues that even if a convention must, in some sense, 'have an alternative' it is not clear that the participants need know of this alternative, or prefer to conform to it if others do, or be psychologically able to adopt it should they ever in fact wish to.¹

Burge proposes his own account of the arbitrariness of conventions; roughly, conventions are not uniquely the best biologically possible means of fulfilling their social functions. Burge explains 'social function' as follows:

'F is a social function of regularity R in population P if (1) F results from the mutually expected conformity of almost everyone in P to R (my italics) and (2) F fulfills certain needs of most members of P' (p.243).

Burge does not explicitly present an account of conventions or conventional regularities themselves, but from what he says we seem to glean something like

¹ Burge discusses these disjuncts, in the above order, pp.250-1; pp.251-2; pp.254). Moreover it seems that Burge would deny that the belief that others conform to the conventional regularity need in fact, or from the point of view of 'ideal' rationality, play a role in any agent's conformity. For now I shall not comment on these claims, except to note that Burge presents a number of examples intended to support them. In fact I am inclined to agree with all of the claims, if not with the examples (see below).

the following idea: a regularity R in a population P is 'conventional' if (1) almost everyone in P conforms to R; (2) (almost) everyone expects (almost) everyone in P to conform to R, and this is (something like) group common knowledge; (3) R is arbitrary (in Burge's sense)¹. It is not clear to me whether Burge means to say (somewhat following Lewis) that every regularity which is a convention must have some social function (or indeed exactly what that is supposed to come down to. Must every regularity in action have a Burgeian social function?) If not every convention does have such a function, it is hard to be sure whether Burge's account of the arbitrariness of (all) conventions will work. Will he perhaps count (commonly known) regularities with no social function, but which are neither biologically nor psychologically necessary, as, automatically, conventions?² It looks as if this latter interpretation, at least, will make all 'mere' customs conventions.³ And perhaps sometimes these two, custom and convention, are not distinguished. This could be so in a very broad notion of 'convention' or 'conventional regularity'. However, I think Burge's account is too broad as analysis even if it is restricted to regularities fulfilling a need in a fairly clear and perhaps not ubiquitous way. To briefly argue this I use one of Burge's own examples:

¹ Or perhaps (1) and (2) can just be elided thus: (1/2) it is group common knowledge in P that (almost) everyone in P conforms to R.

² For the 'reason' that they cannot be the 'uniquely best' means of fulfilling the (nonexistent) function in question.

³ The same goes for an account in which every (regularity in) action does automatically fulfil some need ^{and so} has a social function in Burge's sense.

'...The convention of tipping one's hat to a passing stranger becomes a national trademark. The citizens are sentimentally attached to this mode of greeting and its associations...to the extent that each would rather fight for the traditional greeting, or give up greeting strangers altogether, than switch to another one, even if the others were to switch...' (p.252).

Suppose that it is mutual knowledge in Hattipland that everyone in Hattipland tips his hat to passersby. Suppose also that this hat-raising fulfils a deep need to counteract the looks of strangers in each citizen, and that a broad smile could have and would have fulfilled this need just as well. However, each one is convinced that he alone cares about this greeting, he assumes the others are quite indifferent towards it, that it is purely a matter of habit with them. Suppose further that each one only really cares that he himself be loyal to what has gone before, never mind what the others do. Nor would anything the others could do persuade him to change. I myself am dubious as to whether it is a convention among the confused citizens of Hattipland to tip hats to strangers. It may not rightly be called a custom either. At any rate hats are not just tipped as a matter of custom, but rather as a matter of loyalty to a tradition.¹

It seems to me that the motives for which the 'regularly' done acts are done may have something to do with whether a certain regularity may be called a convention, or at least with whether we may say it is conformed to by convention or as a matter of convention. Schiffer has given a clearly motivational account of conventions and to this I now turn.

Schiffer's analysis of convention is similar to Lewis's but lacks the conditions about preferences.² There is one other significant change (not remarked on as a change by Schiffer). For Schiffer the analysandum is not 'an L-regularity which is a convention' but, roughly, 'a convention prevails in a population'. The analysans involves what may be thought of as regularities, but not what we have called L-regularities: these may be vacuous (see later).

¹ I believe that the notions of 'custom' and 'tradition' are distinct. Cf. D. Armstrong. op.cit.

² Cf. Schiffer, op.cit. pp.136-155, esp.154.

One of Schiffer's criticisms of Lewis regarding 'so-called' conventions of etiquette' has already been noted. As noted, it is not clear that actual so-called conventions of etiquette do not conform to Lewis's requirements. This is an empirical matter. Moreover, I find no special plausibility in Schiffer's own account of such conventions. He puts it all in terms of expectations, as in this example:

' In some places...there prevails a convention for men to open doors for women...there is a precedent, or set of precedents, in a certain group for doing a certain sort of act (or activity) X in certain sorts of circumstances C; on the basis of this precedent everyone expects everyone else to do X when in circumstance C, everyone expects everyone to expect everyone to do X when in circumstances C, etc., and because it is expected, people do X when in circumstances C' (p.152).

What is the sense of 'expects/expected' here? My own sense of the facts, in such cases, is that the 'expectations' involved are what we might call normative. Though it may be true that people have 'ordinary' expectations on the basis of precedent, I suggest that such cases involve an important second kind of 'expectation' (which is motivationally crucial). I am inclined, then, to want to write, after 'women' above, something like: (because it is expected?) everyone normatively expects everyone else to do X when in circumstances C, everyone expects everyone normatively to expect...etc. and (because it is normatively expected?) people do X when in circumstances C.

What is a 'normative' expectation? The phenomenon I have in mind is not altogether easy to analyse. It is perhaps best conveyed by using a certain tone of voice (hard to express on paper), such as a mother might use in saying to her son 'I certainly expect you to come to your sister's wedding'. She is not merely saying that she believes (or knows) that he will in fact come to the wedding, but something more than this.¹ Both an 'ordinary' expectation and an attitude seem to be involved; in particular, in this example, the mother not only expects her son to attend (in the ordinary sense of 'expect') but also

¹ At least if we take 'conversational implications' as well as sense into account.

thinks that (ceteris paribus) he ought to attend the wedding . I believe that the 'flavour' of the case is best captured if we put things this way: the mother thinks that her son should (now) attend the wedding at least because (as he now knows) she expects him to come.

Let us, generalizing, take it that:

X normatively expects Y to do A in circumstances C if

and only if (a) X expects Y to do A in C and (b) X thinks

that Y ought to do A in C at least because X expects this.

I think we normally conceive of the 'conventions of etiquette' at least, as involving 'expectations' of some such kind. Given that we are talking about normative expectations, Schiffer's next remark is less relevant: 'What is not so immediately clear... is why one should do what one is expected to do'. He makes his own suggestion, including:

'Perhaps women expect men to open doors for them ~~because~~ because it is the ~~precedented~~ thing to do; men continue to open doors for women because they do not want to upset women by acting contrary to their expectations...' (ibid).

Perhaps, but it is not immediately clear why one should want or be thought to want people to act as one merely expects them to act. (Clearly one might be very pleased that, say, an homicidal madman did not act as one expected him to act). Why one should want people to act as one normatively expects them to act is an easier question. The phenomenon I have just described in terms of normative expectations, even if it doesn't posit motivational facts, makes queries about them relatively redundant. It is, surely, immediately clear why people should wish to act in accordance with commonly known generalised normative expectations. (Such expectations may be thought of as providing 'normative pressure' on a person (I have ~~taken~~ this phrase from Steven Lukes)). Obviously one will incur others' disapproval if one does not so act, and this just is (generally) unpleasant. Cases of etiquette where each one does what he does simply because of a well-known general normative expectation may provide a counterexample to

Lewis (see later). But let us here see how they bear on Schiffer. Surely, in cases where one does what one does on account of normative expectations, it may be false or at best misleading to say that one does what one does on account of expectations simpliciter: suppose, for instance, that, had there been only ordinary expectations in the case, one would have acted otherwise. Schiffer will, I suggest, have failed to have stated necessary conditions for convention if he is construed in terms of mere ordinary expectations and if the basic motivational fact involved can be (let alone must be) the 'normative' component of normative expectations.

Another issue arises for Schiffer: are motivational facts involved in the 'definition' of the (everyday) notion of convention as opposed ^{to conventional} regularity?¹ Motivational facts would not be involved if, in stating that a convention exists we are merely stating that certain conditions for certain motivations in acting obtain. I turn now to Schiffer's use of what we may call a 'conformity condition' and the question whether it is plausible to include such a condition, in general, in a definition of convention.

ii. Conventions and conformity.

Schiffer's account of conventions makes all conventions regularities of a sort. His definition is 'There prevails in G a convention to do an act...of type X when (or only when)...if...' His definiens, meanwhile, includes the clause '...almost everyone in G does X when...' (p.154). Now, Schiffer implies that the regularity (of doing X in S) may be vacuous. That is, there may be no instances of doing X in S. He states- what we agree with - that there may be a convention in a group G to do X in S even though no one in G has ever done X. (He cites a case in which the relevant situation never occurs).

The idea of a vacuous regularity, however, looks as if it involves a certain

¹ In contrast with Schiffer, Lewis seems deliberately to avoid actually making such claims in his various analyses of convention: his detailed accounts of conventional regularities, however, involve motivational claims, as perhaps they should.

lawlikeness. That is, it looks as if it implies something like: if ever S were to occur, members of G (for the most part) would do X (and moreover they would do X because...). Yet this last formulation is - I contend - problematic in the definition of convention. I suppose many 'lawlike' generalisations have their ceteris paribus clauses, but this one will surely have to have more than most. We have already considered some examples in which there seems to be a convention, where the relevant action is never done, although the relevant situation recurs (see Section 2(i)). We noted that these examples raised doubts about whether we would normally ever identify a convention with an L-regularity. (An L-regularity involves at least an actual set of occurrences of A in S). They also help to raise doubts about the identification of conventions with regularities of any sort.

In particular the question arises: can there be a convention in a population P to do A in S without there even being a 'vacuous' regularity in P of doing A in S, i.e., without it being true in any way that if S were to occur members of P would do A?

Consider now the following imaginary description. 'There's a convention in Lettertown that one writes a 'thankyou' note to one's host after a gracious dinner party. Few inhabitants of Lettertown, however, actually get round to writing thankyou notes after such parties. Many give the idea some thought, but somehow don't get around to sending notes. Sometimes a slight feeling of unease about their lack pervades them for a while. When-as very occasionally happens-thankyou notes are received, no one is put out. Everyone feels that the sender has acted in the proper fashion. The absence of notes, on the other hand, is disappointing, but is regarded with some resignation, since the general lack of conformity to the convention is well known.'

Is this description inconsistent? It strikes me as not only not self-contradictory, but as probably describing an actual state of affairs in some circles.

I hazard, then, that we may consider this as a case in which there is a

convention to do A in S, and in which there is a recurrence of S without A's generally being performed in S. Moreover it is quite clear that the people of Lettertown may correctly identify situation S (i.e. know that they have recently been to a fine dinner party) when it occurs, and may be perfectly capable of doing A in S (they need have no enormous psychological problems about writing, for instance), and generally have the means (pens, paper, money for stamps, and so on) to do so.

Under what conditions, if any, would the townspeople of Lettertown by and large conform to the regularity in question, i.e., do the required action in the situation at issue? I suggest that the following 'filling' to the story might enable us to answer the question. We return to the (rough) idea of generalised normative expectations.

It is group common knowledge among the people of Lettertown that everyone normatively expects everyone else to write a thankyou note after a party. Now, hardly anyone does conform, but none the less everyone has some reason to conform every time S occurs. This reason may never, in fact, outweigh other reasons for acting: the desire to avoid extra trouble, to leave for one's holiday forthwith, to save the money for a stamp, and so on. Nonetheless, the generalised, commonly known normative expectation is there and the following consequently seems to be true:

If, when he was in S, a person in Lettertown knew of no other reasons for acting he would still have, and know he had, a reason to write a thankyou letter and so if 'rational' in a certain sense he would write such a letter; that is, if he always did what the 'balance' of current reasons said he should do.

Now, it seems to me that if all the above is true(1) there is a convention in Lettertown that one write thankyou letters; (2) the existence of the convention will not depend on an L-regularity; (3) it is not best to describe the existence of the convention in terms of (albeit vacuous) regularities at all; (4) it is

the existence of commonly known generalised normative expectations, which is the salient basis for our ascription of a convention here.

iii. A nonconformist sketch of conventions,

I propose that^{something like} the following condition is sufficient for there to be a convention (in a social group G to do A in situation S):

(Condition C) there is a collective normative expectation in G that members of G do A in S.

Notes. 1. Let us say that, in any case, apart from everyday usage, when condition C is fulfilled in a social group there will be an 'E-convention' from ('etiquette') in the group. Let us stipulate that Condition C is both^a necessary and a sufficient condition for the existence of E-conventions.

2. When C is fulfilled, then if a member of G had no other reason for acting in S, he might act with the fulfilment of C as his reason for acting, and hence do A; if he were 'rational' (i.e. acted on the 'balance' of reasons) he would do so.

3. This account of E-conventions leads naturally to an account of a conventional regularity, or, rather, a regularity conformed to by E-convention.

A regularity R in a social group G will be conformed to by E-convention if and only if (1) there is an E-convention in G to conform to R, and (2) by and large members of G conform to R (at least in part) for the (basic) reason that (1).

4. This account of E-conventions and conventional regularities is somewhat close to Weber's account of convention simpliciter. Thus Weber:

'a system of order will be called convention so far as its validity is externally guaranteed by the probability that deviation from it within a given social group will result in a relatively general and practically significant reaction of disapproval'

and

'The term 'convention' will be employed to designate that part of the custom followed within a given social group which is recognised as 'binding' and protected against violation by sanctions of disapproval' (op.cit.p.127).

5. If I have stated sufficient conditions for the existence of a type of conventional regularity, then Lewis's account of such regularities is too narrow. Schiffer's account will be too narrow also.

I have only claimed to give a complete analysis of E-conventions. However, it does seem to me possible that a completely general account of the conventions of a social group can be given in terms of my condition C. At this point I must simply note my belief this this is so.

6. Conventions, Social facts, and the subject matter of Social Sciences:

Some claims in favour of conventions.

We have completed our discussion of the everyday notion of convention. Our results, unfortunately, throw doubt on every account of that notion that has recently been offered. Let us consider first, what thoughts might have supported an initial sense that the notion of convention is preferable to that of a ns. Durkheimian social fact in characterizing the (or a) fundamental subject matter of the social sciences. Three main points that spring to mind are:

(1) We all already have the concept of a convention; it is a well-established concept which may be assumed to have proved its worth over many years. Why make use of a very similar but untried 'invented' concept? In particular an old-established concept like that of a convention is likely to 'carve nature at the joints', and indeed is more likely to do this, and pick out a natural class of cases, than is a new fabrication.

(2) Conventions clearly differ from society to society - the English drive on the left, the Americans drive on the right, for instance. Conventions are, moreover, ubiquitous in a given society, they do much to determine the character of life in that society. They are, then, clearly important phenomena and phenomena which can usefully be investigated by the scientist.

(3) The existence of a certain convention in a population is surely a highly social state of affairs: perhaps nothing more 'social' exists.

I am sympathetic to all these proposals, however, the notion of convention is - as the whole of this chapter shows - very hard to analyse. Thus it hardly, at this point, looks like an important rival to our Durkheimian notion of social facts if what we want is an illuminating and precise account of a social scientific subject matter.

We do have to hand, however, the relatively precisely defined notion of a Lewisian convention on the one hand (let us take Lewis's most recent account in 'Languages and Language') and of our own E-conventions (and the related regularities). Let us now compare and contrast these notions with that of a Durkheimian social fact, and consider them in relation to the above three points and, more generally, in relation to our overall project. First, we should be clear about the relationship of these phenomena to each other. To recapitulate :

A Durkheimian social phenomenon is:

A collective practice Pr of a social group G such that one of the basic reasons for most conforming members' conformity to Pr is provided by some particular collective practice of G, Pr' (where Pr' may = Pr).

ES.

A regularity R in a social group G is conformed to by E-convention if and only if:

- (1) Normative expectation of conformity to R is a collective practice in G.
- (2) By and large members of G do conform to R, in part for the (basic) reason that (1).

Evidently, E-conventional regularities may be, but need not always be, Durkheimian social phenomena. They are such wherever R is a collective practice.

For present purposes a Lewisian convention may be defined as in essence:

A general regularity in action Pr, in a population P, such that

- (1) it is group common knowledge in P that Pr is general in P; and

where

- (2) there is a general belief in P that conformity to Pr is general in P, and this ^{is} group common knowledge; and

- (3) there is a general preference in P for conformity by all to Pr if all but one conforms to Pr, and this is group common knowledge.¹

As with E-conventions, the Lewisian variety may, but need not necessarily be Durkheimian social phenomena. They are so if, (a), (1) or (2) or both provide a basic reason for conformity to Pr; (b) Pr is in fact a collective practice, and the general practices which provide basic reasons for conformity to Pr are also collective; and (c) P is a social group.

We shall suppose for what follows that as a matter of logic (a) is fulfilled by an L-convention.

Let us now consider the claims which might be made in favour of convention, in relation to L-conventions. First, well-established though the concept of convention is, it is at least highly doubtful whether all and only the L-conventions in the world are the conventions (or conventional regularities). The notion of an L-convention, then, can hardly be considered well-established, even if the notion of a convention is. Thus in this connection it has no superior ancestry to the 'new' technical notion of a Dsp. ss.

With regard to 'carving nature at the joints' or capturing a natural kind of phenomenon, we have already noted that Lewisian conventions are themselves a motley: they can be sustained by accident plus sheer conformism - or they may be established solutions to an urgent and specific practical problem - like which side of the road to drive on, or how to signal that one is in distress. Durkheimian social phenomena are, in fact, something of a motley too, insofar as some may have an interesting property others lack. For instance some are, and some aren't, L-conventions. Some are, some aren't, E-conventions. So it is not clear that either notion is superior in this respect. My own intuitive feeling

¹ This is a shorter version of Lewis's 'Languages and Language' account.

is that both are at least superior to Weber's notion of social action in this respect, and so have a relatively high 'naturalness'. Further, insofar as the notion of a Dsp is both a more-inclusive notion and a relatively natural one, it is a more impressive offering in characterizing what social sciences are about. Strictly speaking, it does not include all possible Lewisian conventions as defined but those it does include are surely, in fact, the paradigm cases (more on this in section (7) below).

Next, conventions-whatever exactly they are - surely are important phenomena, ubiquitous and differing from one social group to another. And L-conventions do presumably share this feature, though as T.Burge has pointed out, L-conventions involve a lot of knowledge about the underlying rationale of practices, and the knowledge may not be present even if the rationale is, and the rationale may not be present even if a convention is. So L-conventions may be relatively rare. By parity of reasoning, Durkheimian social phenomena may be more widespread than ns. L-conventions. More importantly, questions of the causal origins, consequences, and interrelationships of phenomena of either type can, presumably, be usefully raised. And so phenomena of either kinds (or both) are prima facie meat for the social scientist.

Finally, from an intuitive point of view I should say Lewisian conventions are highly social phenomena. (Clearly, if Weber's type of 'social action' is indeed a social phenomenon, then so are Lewis's conventions. The existence of a Lewisian convention involves a plethora of social actions of Weber's type, for regularly over time such actions must be performed by most members of a certain population). However, it is by no means clear that Lewisian conventions are by their definition more 'social' or involve more 'social' aspects than our Durkheimian social phenomena. I would say, rather, that the reverse is true. (We return to this issue shortly).

So much, then, for the notion of an L-convention; it is not clearly superior to our Durkheimian notion in any of the ways proposed: age, naturalness, scientific interest, or socialness.

Turning briefly to E-conventions (or rather E-conventional regularities) these may do somewhat better. It is possible that, necessarily, all and only E-conventions are conventions (or 'social conventions'). However, we must leave that as at most a hopeful conjecture. Again, E-conventions have not yet been shown to be divisible into one or more sets of useful subcategories, hence the concept of an E-convention may be both better-established and more natural than that of a Durkheimian social phenomenon. Yet insofar as it is a narrower notion than that of a Dsp, which latter is relatively natural, it too may be a less impressive offering in characterizing the scope of social sciences. Finally while E-conventional regularities are highly social, Durkheimian social phenomena, it seems to me, are even more so. In particular, as far as our definition of E-conventional regularities goes, these need not themselves be collective practices though they must themselves be based on such practices.

As regards the range, homogeneity, scientific interest, and socialness of the phenomena it picks out our Durkheimian notion compares favourably with those of an L- or E-convention. This does not mean that the latter notions are unimportant or should be discarded: they too pick out important ranges of social phenomena. For - and this is the most general thesis of this work - the 'social', itself, is a motley. Part of the special interest of the Durkheimian notion is, however, its broad-range coupled with its precision, and its very precisely specifiable kind of socialness: these phenomena inhere in, and are produced by a particular social group.

ns.

7. Social groups revisited

i. Conventions and social groups

Someone might claim that a flaw in our Durkheimian notion of a social fact, as opposed to ~~that of~~ Lewis's concept of convention, is that the former involves the obscure notion of a social group, while the latter avoids it. Thus, it may be argued, Lewis's notion is preferable from the point of view of intelligibility.

I here tackle this contention - which I think may be rejected - on two fronts. First, I argue that it raises a completely new criticism of Lewis's account of 'convention', both as an account of our everyday notion, and more generally as well. Second, I argue that the notion of a social group, in spite of its evident recalcitrance to analysis, is in any case not so very obscure. First, then, the criticism of Lewis.

According to our everyday understanding of the term 'convention', I think one would judge that only social groups can have social or interpersonal conventions. It seems quite odd to say, on the one hand, that it is a convention in a population P to do A in S, yet, on the other, to deny that population P is a social group of some kind. Assuming, then, that this judgement is correct for whatever reason, it seems to provide a test for accounts of convention of the form 'a convention exists in a population P if and only if...' That is, we can ask: can there be populations which fulfil the stated conditions on convention yet which are still, intuitively, not (thereby) social groups? If so, the account fails this test. I believe that Lewis's account, which does not stipulate that 'population P' is a social group, does fail this test; I rest my case on the following example. Consider the following case.

1. Everyone in social group G is a conformist of a special kind: each prefers to do whatever most others in any logical class he belongs to do, ceteris paribus. More generally, within each logical class to which he belongs, each prefers general conformity to any regularity R rather than conformity by all but one. This is all group common knowledge in G.

2. Everyone in G reads the Daily News: this is common knowledge in G.

3. The Daily News's headline yesterday was: in G most redheaded people clap their hands twice on rising every morning. This headline was true.

Given the truth of these three assumptions, it will easily come about that:

4. All redheaded people in G (let us call this population population 'P') believe that most redheaded people in G clap their hands twice...etc., and this is group common knowledge in P.

5. The belief that others clap their hands twice gives each member of P a good and decisive reason to clap his or her hands twice, etc., (or rather to go on clapping his or her hands twice etc.,) and this too is group common knowledge in P.

6. Had they believed most redheaded people clapped their hands twelve times on rising this would have given members of P a good and decisive reason for clapping their hands twelve times, and this is group common knowledge in P.

7. It is group common knowledge in P that most members of P (will continue to) clap their hands twice on rising.

Population P in the above example fulfils Lewis's conditions on convention with respect to the regularity of clapping one's hands twice on rising, yet the population of redheads is surely not a social group (at least by virtue of the above facts). At the same time, and perhaps consequently, it is hardly clear intuitively that there is anything we should naturally call a convention in P.

The logical possibility of the situation described therefore calls into question the sufficiency of Lewis's conditions for convention in the everyday sense, and this may be because Lewisian convention-populations do not have to be social groups. The insufficiency of Lewis's conditions for convention in the ordinary sense, in this particular respect, makes a stipulative account in his terms look somewhat suspect. There is surely something rather unsatisfactory, at least, in a notion which is so wide as to count the population P in our example as having a convention, even if conventions can be allowed to result,

in some cases, from the motive of 'mere' conformism.¹

ii. The concept of a social group.

The concept of a social group, and the more specific collectivity concepts like that of a society, are surely not ~~in any case~~ so very mysterious or obscure. For we can at least describe cases - more or less elaborately - in which, by a set of simple steps concerning the actions and thoughts of individual persons, a state in which there is a social group comes about. For instance, suppose three strangers meet on a train, get into conversation, and begin to play cards. Someone coming across them as they play will, surely, be coming across a small social group. Though transient and small, such a group may, as Lewis points out, develop a set of temporary, local conventions to deal, say, with disagreements over the rules of their game.²

Now someone might ask: must we not consider how people can come to 'agree to play cards': does this not presuppose that they already are members of a wider social group whose members know what playing cards is, and know the same language?

Pace Winch, I do not see that this is so. Let us develop an outright fantasy.

¹ It may be that Lewis's account can be amended so as to (a) not stipulate that population P is a social group, but (b) entail that if a population P fulfils the conditions on convention it will be a social group. This may be a better way of amending his notion than simply to stipulate that P is a social group. It may be that an analogous point should be made about my own accounts of Durkheimian social facts and E-conventions. That is, it may be that they too could usefully be refined in such a way, that they ceased to stipulate that the practices in question were practices of a social group, while it was nonetheless entailed that if a population had the feature in question it would be a social group. Such refinement of my own notions, if necessary, and of Lewis's, must await further work.

² Cf. Lewis, p.104: Convention.

* End of section omitted by the D.Phil. examiners.
The following pages (to page 302) were read
as an appendix to Chapter 4.

Imagine three beings who, after drinking from a certain stream, begin to speak and think in English, and each finds himself desiring to play cards. Faced with each other, one asks, hesitating, unsure if he will be understood: 'shall we play rummy'? Another, almost equally tentatively, 'I don't know how to play rummy'. Finding they seem to speak the same language, they go on with the conversation. The nature of rummy is explained, some rudimentary cards are manufactured with pen and paper, and they proceed to play. I would say that this population comes to be a social group of a sort after not being one; nor, we can stipulate as part of our fantasy, does any other such group actually exist anywhere.

We have here described a transition from a nonsocial group to a social group without positing or presupposing the existence of any social groups prior to the existence of this one. It does seem to be essential that our imaginary beings had certain concepts, for instance, the concept of a card game. Does this mean that they have the concept of a social group? And if so, does this make our construction of a social group covertly imply at the outset that there are social groups? Even if we assume that these creatures do have the concept of a social group, we can surely deny that this assumption entails that there are social groups. The origin of a social group here then, appears to have been described without covert assumption of a pre-existing social group.

Naturally, it would be good to have a perspicuous account of the fine details which, intuitively, go towards 'producing' a social group in situations like the above.

Towards this end, I now present a very rough sketch of what I think is a promising kind of general account of a social group. This draws on the elements of collective practices already introduced, but is not quite as simple as the (too) simple account of social groups abandoned in Chapter 4. It depends, importantly, also, on the notion of a face-to-face situation (to be defined below). The basic idea is that social groups paradigmatically result from the interaction between individual human beings. Though this is so, one does not expect each

member of every social group to be acquainted personally with every other member. The following account at least captures these two facts, if they are such, about social groups.

First, I would define the crucial notion of a face-to-face situation somewhat as follows:

A 2-person face-to-face situation obtains with respect to individuals A and B, if and only if A and B are within perceptual range of each other and (a) A is aware of B's presence; (b) B is aware of A's presence; (c) it is commonly known between A and B that (a) and (b) obtain, on the basis of their 'physical' circumstances (e.g. direction of gaze, nearness, and so on); (d) it is group common knowledge in the population comprising A and B (call it 'P') that (a) and (b) obtain.

Perhaps when individuals A and B are in a face-to-face situation they already form a social group, of an elementary sort. I myself am, indeed, inclined to say this is so. I have space to discuss only one possible objection here. Two very hostile persons may face one another in this way, with the intention of fighting it out to the death. Nonetheless insofar as the elements noted are present, it seems to me that these people may intuitively be said to form a social unit, albeit of a highly primitive type. One reason for saying this is, it seems to me, that in a simple and natural way we can build up from face-to-face situations between two individuals to wider and wider structures, without introducing any concepts of a radically different type.¹

¹ It does not follow from this that in every 'war of all against all' the populations involved are social groups. For there need be no collective practices in these wider populations, first, nor need there be any group common knowledge in them. Second, given mini-populations of, say, one man striking another, may not be 'face-to-face groups'. (Say one shoots at another from behind a tree and kills him; or stuns him with a blow, and so on). However, it does follow that some such mini-populations may form 'social groups' of an elementary and primitive kind and this is surely acceptable (as we ruled in Chapter one).

Let us now consider the derivation of more complex structures from face-to-face situations, whether or not populations in these situations are judged to be social groups already. Let us go back to the mushroom-pickers of Chapter One. Imagine that, rather than avoiding one another's gaze (and hence avoiding mutual recognition) A and B come to form a 'face-to-face group'. A points to a mushroom and says 'Do you like mushrooms?' B nods vigorously. It is now, let us take it, group common knowledge in P^* (the population consisting of A and B) that A and B speak English. Imagine they are rather shy and bid each other good-day after this brief encounter. Then imagine many such incidents in the wood. Each mushroom picker soon begins reasonably to infer that everyone living around the wood (call the population of these people P) speaks English, and moreover, that everyone in P knows this, and so on. He infers this on the basis of numerous^s face-to-face situations of common knowledge among members of P. There comes to be, in effect, a collective practice analogue in P that members of P speak English. But in this context, I suggest, P will be a social group. We might, then, generalise:

a population P is a social group if (a) each member of P has been involved in one or more face-to-face situations with other members of P in which each member of the (face-to-face) mini-population involved ϕ -d, and where it was group common knowledge in that mini-population that each member of it ϕ -d; so that (b) on the basis of face-to-face encounters (i) it is group common knowledge in population P that most members of P ϕ ; and (ii) it is group common knowledge in P that (b)(i).

Clearly not all social groups are of this type, hence we cannot quite write 'if and only if' here. For instance take it that each mushroom-picker lives, not alone, but with a 'wife' and child. The wives and children never enter the

woods or meet other mushroom-pickers or wives. Each 'nuclear family' is a social group, with many collective practices based on face-to-face contacts. But perhaps all members of these families can be considered members of one all-embracing group: one reason might be that the wives and children are parties to collective practices of 'we woodlanders', learning about the others from the man who goes out and mixes with others. In such a case, the group is a group, ultimately, because of face-to-face interactions between many members.

With this brief sketch I must here leave the 'obscure' 'mysterious' notion of a collectivity. I hope that I have made it look sufficiently penetrable, so that it will not be regarded as a vital flaw in our Durkheimian notion of a social fact, or in any other notion, that it uses this concept, ~~of Lewis's concept of convention, that it does not.~~ However, it is also clear that further work can usefully be done in this area.¹

¹ Note that our partial definition of a social group here does not as stated entail that any Weberian social actions have been performed in these groups, though in our example such actions were performed. It is still not clear to me that such actions are in fact logically necessary to the existence or formation of social groups.

CHAPTER SIX

ON SOCIAL FACTS

1. A theory of 'socialness'

In this chapter we finally return to the only global 'theory' of social science clearly standing, that relatively trivial theory noted in Chapter One, which ran thus:

a social scientific theory is any scientific theory
whose subject matter is (a) intuitively social (b)
social to what is intuitively a sufficiently high degree.

We now have a good basis on which to say more about what it is for a phenomenon to be social -- the intuitions about socialness which have emerged in the foregoing chapters. I shall here hazard a general theory of the 'social' or of socialness, which takes these intuitions seriously and attempts to elucidate a corresponding principle. The theory will incorporate the intuition that there are degrees of socialness.

Let us review our intuitions. First, intuitively, 'social' things are such that no global theory couched in terms of one natural kind of 'social' thing can fail to be misleading in some way. The most promising 'piecemeal' theory was judged to be that in terms of our Durkheimian notion of a social fact. Weber's notion of a social action provided an unsatisfactory rival,

but at least did pick out an intuitively social kind of thing, and made it clear that to attempt a global theory in terms of our Durkheimian notion, or something very like it, was slightly risky. Hence one broad divide among social phenomena has been noted; this can roughly be characterized as that between the properties or features of populations or social groups (what we might call, for short, population-based phenomena) and the properties or features of individual persons (which we might call 'individual-based' phenomena). In particular there were on the one hand the items of meaningful behaviour of individual persons, which might be called 'social' for various reasons, and on the other hand the collective practices of social groups.

Let us first consider our judgements on kinds of actions individuals might perform. In discussing Peter Winch's claims, in Chapter 3, we noted some possible definitions for the phrase 'social action' such that the kind of socialness which, logically, had to be involved in such actions appeared to be very weak.

Recall in particular:

- (1) a 'social action' is any action of anyone who has at some time in his life had experience of a society with its socially established rules; and
- (2) a 'social action' is any action which presupposes the belief that others (other human beings, say) would, in the appropriate circumstances, do such-and-such.

We noted at the time that the actions social under definition (1) appeared to have a 'borrowed' kind of socialness - borrowed from 'societies' and their 'socially established rules'. Moreover their relation to societies and such rules is rather weak itself. The agent does not, in particular, have to have these in mind when he acts. Nor does the definition, as it stands, even posit any causal connection between the agent and his experience of societies. We shall consider possible reasons for the weak socialness of (2) shortly. One possible definition, noted in our discussion of suicide in Chapter 2, defines a phenomenon which is about as 'weakly' social as that falling under (1), thus consider:

(3) a 'social action' is an action of a type whose incidence in a given society is influenced or caused by particular states of collectivities.

Actions social in this sense, like those falling under (1), have a 'borrowed' kind of socialness - borrowed from 'particular states of collectivities'. Here, however, no present or previous experience of a society is posited for a producer of a social action in this sense. A causal relation between the agent and his society (if he has one) is what links actions and societal features in this notion.

Finally, Weber's notion of 'social action' appears to have a stronger kind of socialness than notions (1) to (3) above. Thus we ranked higher than the foregoing on the scale of socialness actions such that - in Weber's terms - their 'subjective meaning' includes reference to one or more other persons, including 'possible persons' - that is people the agent believes may exist.

In this connection we may note that insofar as one's Weberian social action is oriented to others one knows exist, it may thereby be more social intuitively than actions oriented to persons one merely believes may exist. Thus someone's putting on a pretty dress to please her fiancé, is probably more aptly called a 'social action' than her dressing well in case 'Mr. Right' should happen to be at the party. There may be no Mr. Right (she may have inexorably high standards).

The salient feature of the most social kind of action in this chain appears to be a fairly strong kind of 'mental connectedness' between one person and another, and perhaps in particular between one real person and another. The wider Weberian notion involving 'possible people' is probably a notion of a less social phenomenon than a more restricted one.

Returning now to notion (2) above, there is what can be construed as a reference to merely possible others - which I believe contributes to one's sense of its weak socialness. Moreover, and probably most important, the 'mental connection' here between agent and others, real or otherwise, is itself intuitively weaker: reference to the others is not --- as we noted in Chapter 3 --- part of the subjective meaning of the agent's act by logical necessity.

In summary, concerning kinds of 'meaningful behaviour', the most strongly social kind considered -- Weber's -- involved a clear mental connection between individuals; a more weakly social kind -- (2), above -- involved a weaker sort of mental connection. Two kinds of 'meaningful behaviour' considered very weakly social derived the 'socialness' they had from the agent's (rather weak) connection, causal in one case, 'experiential' in another, with societies and/or socially established rules and collectivity states.

Let me turn now to 'collective practices' and some of their kinds. First, our Durkheimian social phenomena, collective practices of social groups produced by other collective practices, seemed to have a very high degree of socialness. They certainly appeared to have a higher degree of socialness than that of the phenomena just considered. And if our hypothesising up till now is correct, we can already see why. Let us first consider an action done with a certain collective practice in mind.

Consider a mushroom-picker who coughs on seeing another mushroom-picker, in part because it is group common knowledge among the mushroom-pickers that, in their circle, that action is generally normatively expected in the circumstances. In performing this action he is performing a Weberian social action with respect to certain real other persons (not all personally known to him, assume). In fact he is mentally connected in the strongest way we have considered to a social group, whose members, including himself, are themselves closely mentally connected, at the least insofar as there is the group common knowledge he refers to. There is implicitly a reference to these people -- and their mental connection-- in the subjective meaning of his act.

Now, this is the position of at least two persons -- typically of many more -- when there is a Durkheimian social phenomenon. Each of the agents involved will then in fact be mentally connected with the others in a number of ways. First, each one knows that it is group common knowledge in his circle that, say, a certain practice is considered mandatory in that circle; second, each conforms to the practice because of this group common knowledge; third, each is party to the group common knowledge that the practice he conforms to is general in his group. Finally, since the agents are, ex hypothesi

members of a social group, there surely are, and have been, many 'mental connections' between them, including that surely very strong kind involved in the individual common knowledge one has in a face-to-face situation. The existence of Durkheimian social facts entails, in brief, an enormous amount of mental **connectedness** between real persons.

A 'mere' collective practice according to our working definition is itself intuitively a highly social phenomenon -- though not so highly social as a Durkheimian social fact. And the appearance and persistence of a social group, on which the existence of Durkheimian social phenomena depends, is, also, clearly a highly social phenomenon. None of these things, it is clear, is a kind of action, though these things do not come about ~~unless~~ agents act or at least think in certain ways, and know certain things.

Common to social groups, collective practices, and Durkheimian social facts, is the phenomenon of group common knowledge, and, ultimately, as we suggested in our final remarks on social groups, the phenomenon of individual common knowledge. We recall that it is no simple kind of knowledge that is at issue here: each individual's 'cognitive labour' must be thoroughly mixed with his fellows' before collective practices, and related phenomena, come about. In this case, there is not just 'mental connection' but 'interconnection' of a precisely specifiable sort: each individual must not only know certain facts about what others think and do and have first-order knowledge of. He must also know what they know (indirectly) about

his own knowledge, and so on, as we have argued, ad infinitum - and similarly for each one.

The high intuitive socialness of phenomena like our Durkheimian social facts may explain a tendency among those who see some form of human action as the fundamental social fact to nonetheless move (albeit sometimes unconsciously or inexplicitly) towards a type of action which is logically tied in some more or less natural way to facts about groups or societies.

Weber himself, whose notion of 'social action' is free from any stipulation of interaction or mutual cognitive relations between persons, has in fact a clear tendency to move in this direction and even to implicitly to define 'social action' as 'meaningfully' connected with Durkheimian social phenomena.¹ Recall in this connection his implicit stress on 'action oriented to a consensual order' (noted in Chapter 2). Consider also his remarks on the theoretical possibility of formulating a 'sociology of the relation of men to animals'. He cites as a ground for this possibility:

' Many animals "understand" commands, anger, love, hostility, and react to them in ways that are evidently often by no means purely instinctive and mechanical and in some sense (are) both consciously meaningful and affected by experience' (p.104).

This evokes a picture of face-to-face situations leading to a knowledge of collective (if one-sided) practices, with respect to the form of commands or expressions of anger. Weber seems to be suggesting, in effect, that it may be more than a picture.

¹ Weber is not alone in having this tendency. See Inkeles, *op.cit.*, p.71.

So much, then, for the intuitive judgements of relative 'socialness' made on cases. At this point a fairly simple systematizing principle suggests itself:

a phenomenon is a social phenomenon if and only if it involves one person's being mentally connected in some way with another person or persons.

As for degrees of socialness:

the degree of 'socialness' of a phenomenon is in general directly correlated with the amount and degree of mental connectedness between persons which it involves.

2. 'Mental connectedness'

A crucial notion, that of 'mental connectedness', has been left undefined so far, though illustrated with examples. In this section I suggest how this might be more precisely interpreted.

We might first define 'A is actively mentally connected with B' as 'A has B in mind in some way'. Examples of such mental connection are: A's knowledge of B, A's wish to hurt B, A's acting out of dislike for B, A's dreaming about B -- these are examples. Clearly the relation of 'active mental connection' is not symmetrical. A more general relation of mental connection simpliciter, which is symmetrical, may also be defined. Here if A is mentally connected with B, then vice versa. This mental connection only exists, we shall say, when one party is actively mentally connected with another.

These definitions allow that a real person may be mentally connected with a fictional one, or with a 'possible' person like 'Mr. Right' or someone who might pass by. But two 'possible' people cannot be mentally connected (except in fiction) since for that to happen one must have some active mental connection with the other and no unreal character can be in active mental connection with anything, since no unreal character can actually do (or think) anything (except in fiction).

We should now say something about degrees of mental connection. There are evidently various dimensions here. Knowing things, and thinking things about a person clearly connect you to him. Acting while thinking about him does also, but acting because he did so ~~and so~~, or in order to do so and so ~~to~~ him, or merely doing something to him seem to connect you closer to him. This may be because the latter phenomena already involve the former, but not vice-versa. At any rate the matter of degree seems settled intuitively (Weber's notion of social action incorporates the stronger kind of mental connection here).

Clearly A may be actively mentally connected to B and vice versa. Then there will be a higher degree of mental connectedness between these two than if the active mental connection had only gone one way.

If A knows something about B personally, then his mental connection to B will be closer than if he just knows something about all people in a population into which B falls. In the latter case he may be said to have (~~merely~~) potential knowledge of B and hence be less strongly actively mentally connected with him.

If B is only a 'possible' person then we have the lowest possible degree of mental connectedness. It is as if we have the 'connecting link', and A, but nothing at the other end. Perhaps this should really be called: mental quasi-connection. This relation will be asymmetrical.

A person A may be actively mentally connected to a group G. This seems to involve a greater degree of socialness than such connection with one individual. I think there are two possible explanations of this intuition about mental connection. (1) Insofar as the group in question has at least two members, you might be said to be mentally connected here, ultimately, with more persons than you would in active mental connection with a single individual. The active mental connection in the group case is only potentially with particular individuals, but perhaps there is always a trade-off of weight between numbers directly or indirectly involved and 'actuality' here, so that numbers win. (2) Mental connection to a group involves active mental connection to something itself partly constituted by mental connections -- and intricate, tight ones at that. So perhaps for this reason whenever we ^{have} such mental connection we have something more highly social than with mere mental connection between persons. That is, it is the tight, intricate mental connections between a group that raise the degree of socialness of the whole phenomenon. Consider again those connections exemplified in group and individual common knowledge. These surely occur typically, if not of logical necessity, in social groups. They also seem to exemplify, in their respective spheres, an extremely high degree of mental connectedness between persons.

3. Causal connections between persons.

I now briefly turn to the question of causal connections between persons, and between persons and collectivities. First, recall the notion of an action which is somehow causally influenced by the existence of a certain collectivity state. This is surely less 'social' intuitively than an action 'meaningfully oriented' to such a state. Our theory suggests a reason: 'meaningful' orientation to others involves a high degree of socialness, which this phenomenon lacks. Perhaps all the 'socialness' of this phenomenon derives from the role of a collectivity in it?

However, we should consider a sense of 'social action' not yet noted. It is possible that Mill, in 'On Liberty' should be taken to be adopting this. That is, a definition of 'social action' where it means 'action which affects other people'. This may be an intuitively acceptable definition of 'social action'. Even if so, surely the 'socialness' involved is less than that involved in, say, an action intended to affect other people. In order to capture the kind of socialness involved in this new case, and also to indicate its weakness, we might amend our theory somewhat as follows:

a phenomenon is a social phenomenon ^{and only if} if ~~it~~ involves one person's being connected either mentally or in some causal way with another person or persons.

And as for degrees we might leave things much as they are, but with an addition:

the degree of 'socialness' of a phenomenon is in general directly correlated with the amount and degree of mental connectedness between persons which it involves, and with the degree of causal connectedness between persons, where some mental connection always takes priority over any causal connection.

Clearly, however, more work is needed in order truly to delineate the finer points of the matter. I hope that I may be judged to have made a contentful and useful start.

4. On characterizing social sciences in general.

Though we have hazarded an account of what makes a phenomenon a social one, in general, the fact clearly remains that the realm of social things is a motley. No single concept capturing a relatively natural kind of phenomenon, I contend, can expect to provide a definitive global account of the range of social sciences. The all-embracing notion of socialness, of which we have in effect given an account in this chapter, may be the only one apt for a global theory of social science, yet it is not, intuitively, a concept of a 'natural' kind of thing. Any more natural concept is going to be obviously partial, or piecemeal. The general notion of a 'social' phenomenon discussed in this chapter is, unfortunately, hardly illuminating as to the nature of and prognosis for possible social sciences, though it has some appeal, in fact, in its openness and lack of prejudgement of issues. Those philosophers debating the question 'is a social science possible?' must take this motley nature of the 'social' into account. Of the notions considered here, we have found our Durkheimian concept of a social fact to be the most promising given the criteria adumbrated in Chapter I. Whether there can be a 'science' of such facts, whether there already is, or whether there is likely to be, are further questions. But even if there are affirmative answers to all these questions, the scope of the social sciences will not have been definitely decided. Social scientists, or rather enquirers about the world, had better ask what questions they will and answer them if they can. Then, if they like, they can attempt to formulate concepts which bring out in a perspicuously the precise way in which the subject matter they deal with is, indeed, social. As I hope I have made pellucidly clear, there are many such ways.

BIBLIOGRAPHY

This bibliography lists works referred to in the text, citing editions from which quotations in the text were taken; a selection of other works consulted is also included.

Abel, Theodore, 'The Operation called Verstehen', American Journal of Sociology, Vol. 54, 1958.

Aimard, Guy, Durkheim et la Science economique, Presses Universitaires de France, Paris, 1962.

Alston, W.P., Philosophy of Language, Prentice-Hall, Englewood Cliffs, N.J., 1964.

Andreski, S., Elements of Comparative Sociology, Weidenfeld and Nicolson, 1964.

Anscombe, G.E.M., Intention, Basil Blackwell, 1963.

Argyle, M., The Scientific Study of Social Behaviour, Methuen, 1957.

Armstrong, David M., 'The Nature of Tradition', in Liberty and Politics, O. Harris, ed., Pergamon Press, 1976.

Arrow, Kenneth J., 'Values and Collective Decision-Making', in Philosophy, Politics, and Society, 3rd Series, Laslett, P., and Runciman, W.G., eds., Basil Blackwell, 1967.

Ayer, A.J., 'Can There Be A Private Language', Proceedings of the Aristotelian Society, Supp. Vol. XXVII, 1954.

Bambrough, Renford, 'Universals and Family Resemblances', Proceedings of the Aristotelian Society, Supp.Vol. LXI, 1960-1

Bell, D.R., 'The Idea of a Social Science', Proceedings of the Aristotelian Society, Supp.Vol. XLI, 1967.

Bendix, R., 'Two Sociological Traditions', in Bendix R., and Roth, G., eds., Scholarship and Partisanship, 1971.

Bennett, Jonathan, 'Substance, Reality, and Primary Qualities', American Philosophical Quarterly, 1965.

Bennett, Jonathan, 'The Meaning-Nominalist Strategy', Foundations of Language, Vol. 10, 1973.

Benoit-Smullyan, E., 'The Sociology of Emile Durkheim and his School', in Barnes, H.E., ed., An Introduction to the History of Sociology, University of Chicago Press, Chicago, Illinois, 1948.

Berger, Peter L., and Luckman, Thomas, The Social Construction of Reality, Allen Lane, 1967.

Black, Max, 'Notes on the Meaning of 'Rule'', Theoria, Vol. 24, 1958.

- Black, Max, ed., The Social Theories of Talcott Parsons, Prentice-Hall Englewood Cliffs, N.J., 1961.
- Boonin, Leonard G., 'Man and Society: An Examination of Three Models', in Pennock, J. Roland, and Chapman, John W., eds., Voluntary Associations, Atherton Press, N.Y., 1969.
- Borger, Robert, and Cioffi, Frank, eds., Explanation in the Behavioural Sciences, Cambridge University Press, 1970.
- Brand, Myles, ed., The Nature of Human Action, Scott, Foresman, and Co., Glenview, Illinois, 1970.
- Braybrooke, David, ed., Philosophical Problems of the Social Sciences, MacMillan, 1965.
- Brodbeck, May, ed., Readings in the Philosophy of the Social Sciences, MacMillan, 1971.
- Bryant, C.A., 'In defense of sociology: a reply to some contemporary sociological criticisms', British Journal of Sociology, Vol.21, No.1., March 1970
- Burge, Tyler, 'On Knowledge and Convention', Philosophical Review, Vol.84, 1975.
- Chomsky, Noam, Aspects of the Theory of Syntax, M.I.T. Press, 1965.
- Cohan, Percy S., 'The Very Idea of a Social Science', in Lakatos and Musgrave, eds., Problems in the Philosophy of Science, North-Holland, Amsterdam, 1968.
- Cook, Alan F., The Linguistic Analogy and the Study of Social Action. An Examination of some Applications of Linguistic Philosophy and Generative Grammar. D.Phil. thesis, Oxford, 1975.
- Coope, Christopher, and others, A Wittgenstein Workbook, Basil Blackwell, 1970.
- Davidson, Donald, 'Actions, Reasons, and Causes', Journal of Philosophy, Vol. 60, 1963.
- Davidson, Donald, and Harman, Gilbert, Semantics of Natural Language, D.Reidel, Dordrecht, 1972.
- * Douglas, Jack D., The Social Meanings of Suicide, Princeton University Press, Princeton, N.J., 1967.
- Douglas, Jack D., ed., Understanding Everyday Life, Routledge, 1971.
- Douglas, Mary, 'The Abominations of Leviticus' in Douglas, Mary, Purity and Danger: An Analysis of Concepts of Pollution and Taboo, Penguin, 1970.
- * Dahrendorf, Ralf, Die angewandte Aufklärung: Gesellschaft und Soziologie in Amerika, R. Piper and Co. Verlag, Munich, 1963.

- Durkheim, Emile, Montesquieu and Rousseau: Forerunners of Sociology, tr. R. Manheim, University of Michigan Press, Ann Arbor, Michigan, 1960. (original in Latin: Quid Secundatus Politicæ Scientiæ Instituendæ Contulerit, Gounouilhou, Bordeaux, 1892.)
- Durkheim, Emile, The Rules of Sociological Method, tr. Solovay, S.A., and Mueller, J.H., Collier-MacMillan, 1938, (translation of the following entry.)
- Durkheim, Emile, Les Règles de la Méthode Sociologique, Presses Universitaires de France, Paris, 1968. (First published 1895).
- Durkheim, Emile, Suicide, tr. Spaulding J.A. and Simpson, G., Routledge, 1952. (First French edition: Le Suicide, 1897)
- Durkheim, Emile, 'La sociologia ed il suo dominio scientifico' Ri.It.S., Anno IV, 1900; tr. in Wolff, K., ed, and passages tr. in Giddens, ed.
- Durkheim, Emile, Sociology and Philosophy, tr. Pocock, D., Cohen, 1953. (tr. of French: Sociologie et Philosophie, 1924).
- Durkheim, Emile, and Mauss, Marcel, Primitive Classification, tr. and with an introduction by Reedham, Rodney, Cohen and West, 1970. (French: 'De quelques formes primitives de classification; contribution à l'étude de représentations collectives', L'année sociologique, Vol. VI 1903).
- Emmet, Dorothy, Function, Purpose, and Powers, MacMillan, 1958
- Emmet, Dorothy, Rules, Roles, and Relations, MacMillan, 1966
- Evans, Gareth, and McDowell, John, eds., Truth and Meaning: Essays in Semantics, Oxford University Press, 1976.
- Fain, Haskell, and Phillips Griffiths, A., 'On Falsely Believing that one doesn't Know', American Philosophical Quarterly, Monograph Series, No.6., Studies in the Philosophy of Mind, 1972.
- Fauconnet, P., and Mauss, M., 1901, 'Sociologie', in La Grande Encyclopédie, Vol. 30, Société Anonyme de la Grande Encyclopédie, Paris, 1901. Reprinted in Mauss, M., Oeuvres.
- Filmer, Paul, and others, New Directions in Sociological Theory, Collier-Mac Millan, 1972.
- Fletcher, R., The Making of Sociology, Michael Joseph, 1971.
- Floistad, Guttorn, 'Social Concepts of action: Habermas', Inquiry, Vol.13, 1970.

- Ganz, J.S., Rules: A Systematic Study. (Janua Linguarum, series minor, 96) Mouton, The Hague, 1971.
- Gardiner, Patrick, ed., Theories of History, Free Press, Glencoe, Illinois, 1959.
- Garfinkel, Harold, Studies in Ethnomethodology, Prentice-Hall, Englewood Cliffs, N.J., 1967.
- Garver, N., 'Rules', in Edwards, Paul, ed., Encyclopedia of Philosophy, MacMillan, 1967.
- Gellner, Ernest, 'Holism vs. Individualism' in History and Society' in Gardiner, Patrick, ed.
- Gettier, Edmond, 'Is Justified True Belief Knowledge?' Analysis, Vol. 23, 1963.
- Gerth, H.H., and Mills, C.W., From Max Weber: Essays in Sociology, Oxford University Press, 1946.
- Gibbs, Jack P., 'Norms', American Journal of Sociology, 1965.
- Giddens, Anthony, ed., Emile Durkheim: Selected Writings, Cambridge University Press, 1972.
- Giddens, Anthony, New Rules of Sociological Method: A Positive Critique of Interpretive Sociologies, Hutchinson, 1976.
- Gilbert, Margaret, 'About Conventions', Second-Order, Vol. 3, 1974.
- Gilbert, Margaret, and Berger, Fred R., 'On an argument for the impossibility of social science prediction', American Philosophical Quarterly, Monograph Series, No. 9, Studies in Epistemology, 1975.
- Goffman, Erving, Interaction Ritual, Doubleday-Anchor, Garden City, N.Y.
- Goffman, Erving, The Presentation of Self in Everyday Life, Edinburgh University Press, 1956.
- Goffman, Erving, Behaviour in Public Places, Free Press, Glencoe, Illinois, 1963.
- Goffman, Erving, Stigma, Prentice-Hall, Englewood Cliffs, N.J., 1963.
- Goh, S.T., 'The Logic of Explanation in Anthropology', Inquiry, Vol. 13, 1970.
- Grandy, Richard E., Review of David Lewis's Convention, Journal of Philosophy, Vol. 74, 1977.
- Grice, H.P. 'Meaning', Philosophical Review, Vol. 66, 1957.
- Grice, H.P., 'Utterer's meaning and intention' Philosophical Review, Vol. 78, 1969.
- Gumb, Raymond D., Rule-governed Linguistic Behaviour, Mouton, The Hague, 1972.

- Hacker, P.M.S., Insight and Illusion, Oxford University Press, 1972
- Harman, Gilbert, Thought, Princeton University Press, Princeton, N.J., 1973.
- Hart, H.L.A., The Concept of Law, Oxford University Press, 1961.
- Harre, Rom, 'The conditions for a social psychology of childhood' in Richards, ed., The integration of the child into the social world, 1972.
- Harre, Rom, 'Some remarks on "rule" as a scientific concept', in Mischel, Theodore, ed., Understanding Other Persons, Oxford University Press, 1974.
- Harre, Rom, and Secord, P.F., The Explanation of Social Behaviour, * Oxford University Press, 1972.
- Hempel, Carl G., 'Rational Action', Proceedings and Addresses of the American Philosophical Association, 1962 (in Bobbs-Merrill reprint series).
- Hempel, Carl G., 'Fundamentals of Taxonomy' in Hempel, C.G. Aspects of Scientific Explanation, Free Press, Glencoe, Illinois, 1965.
- Herr, Lucien, Review of Durkheim's Rules of Sociological Method, in Review Universitaire, 2, 1, 1894.
- Hollis, Martin, Models for Man: Philosophic Thoughts on Social Action, Cambridge University Press, 1977.
- Homans, G.C., 'Bringing Men Back in' American Sociological Review, Vol. XXIX, 1974.
- Homans, G.C., Social Behaviour: Its Elementary Forms, Harcourt, Brace, and World, New York, 1961.
- Homans, G.C., The Nature of Social Science, Harcourt, Brace, and World, New York, 1967.
- Honore, A.M. 'Groups, Law, and Obedience' in Simpson, ed., Oxford Essays in Jurisprudence (Second Series), Oxford University Press, 1973.
- Hymes, Dell, ed., Language in Culture and Society, Harper and Row, New York, 1964.
- Inkeles, Alex, What is Sociology? An Introduction to the Discipline and the Profession, Prentice-Hall, Englewood Cliffs, N.J., 1964.
- *Heal, Jane, 'Common Knowledge' Philosophical Quarterly, Vol. 28, 1978.

- Jarvie, I.C., 'Understanding and Explanation in Sociology and Social Anthropology' in Berger and Cioffi, eds., Cambridge University Press, 1970.
- Jarvie, I.C., 'Reply' (to Peter Winch) in Berger and Cioffi, eds., 1970.
- Jamieson, Dale, 'David Lewis on Convention' Canadian Journal of Philosophy, Vol. 5, No.1., 1975.
- Johnson, Barclay D., 'Durkheim's One Cause of Suicide' American Sociological Review, 1965.
- Jones, O.R., 'Can one Believe what one Knows?' Philosophical Review, Vol, LXXXIV, 1975.
- Kant, Immanuel, The Critique of Pure Reason, tr. Norman Kemp-Smith, MacMillan, 1929.
- Kaplan, David, 'Quantifying In', Synthese, Vol.19, 1968-9.
- Kaufmann, Felix, Methodology of the Social Sciences, Oxford University Press, 1944.
- Kripke, Saul A., 'Naming and Necessity' in Davidson and Harman, eds., 1972
- Kripke, Saul A., 'Identity and Necessity' in Identity and Individuation, ed. Milton K. Munitz, New York University Press, 1971
- Kripke, Saul A., 'Speaker's Reference and Semantic Reference' in Midwest Studies in Philosophy, Vol II, eds Peter A. French and others, University of Minnesota Press, Minnesota, 1977.
- Kripke, Saul A., 'Wittgenstein on Rules and Private Language: an Exposition', forthcoming.
- Kummer, Hans, Primate Societies, Aldine Press, 1971.
- Lacombe, Roger Etienne, La Méthode Sociologique de Durkheim: étude critique, Paris, 1926.
- Lakatos, I, and Musgrave, A., eds., Problems in the Philosophy of Science, North-Holland, Amsterdam, 1968.
- Laslett, P, and others, eds., Philosophy, Politics, and Society, Series 1 to 4, Basil Blackwell, 1956, 1962, 1967, and 1972.
- Leibniz, Gottfried Wilhelm, Discourse on Metaphysics, Manchester University Press.
- Levi-Strauss, Claude, Structural Anthropology, Basic Books, New York, 1963.
- Levi-Strauss, Claude, 'French Sociology', in Gurwitsch, G., and Moore, O., Twentieth Century Sociology, Philosophical Library, New York, 1945.
- Levine, D.N., ed., Georg Simmel: On Individuality and Social Forms, University of Chicago Press, Chicago, 1971.

- Lewis, David K., Convention: a philosophic study, Harvard University Press, Cambridge, Mass., 1969.
- Lewis, David K., 'Languages and Language ', in Gundersen, Keith, ed., Language, Mind and Knowledge, Minnesota Studies in the Philosophy of Science, Vol. VII, University of Minnesota Press, Minneapolis, 1975.
- Lewis, David K., 'Convention: Reply to Jamieson', Canadian Journal of Philosophy, Vol. VI, 1976.
- Louch, A.R., Explanation and Human Action, Basil Blackwell, 1966.
- Louch, A.R., 'The very idea of a social science', Inquiry, Vol. 6, 1963.
- Louch, A.R., 'On misunderstanding Mr. Winch', Inquiry, Vol. 8., 1965.
- Luce, R. Duncan, and Raiffa, Howard, Games and Decisions; Introduction and Critical Survey, John Wiley, New York, 1957.
- Lukes, Steven, Emile Durkheim: His Life and Work, Allen Lane, 1973.
- Lukes, Steven, 'Methodological Individualism Reconsidered', British Journal of Sociology, Vol. XIX, 1968.
- MacIntyre, Alasdair, 'The idea of a social science', Proceedings of the Aristotelian Society, Supp. Vol. 41, 1967.
- MacIntyre, Alasdair, 'Is a science of comparative politics possible?' in MacIntyre, A., Against the Self-Images of the Age, Duckworth, 1971.
- Mackie, J.L., 'Causes and Conditions', American Philosophical Quarterly, 1967.
- Macrae, Donald G., Weber, Fontana, 1974.
- Malcolm, Norman, 'Wittgenstein's Philosophical Investigations', Philosophical Review, Vol. LXIII, 1954.
- Mandlebaum, M., 'Societal Facts', British Journal of Sociology, Vol. VI, 1955.
- Mauss, Marcel, Oeuvres (3 Volumes), presentation de V. Karady, Les Editions de Minuit, Paris, 1968-9.
- Melden, A.I., 'Action', Philosophical Review, LXV, 1956.
- Melden, A.I., Free Action, Routledge and Kegan Paul, 1961.
- Mill, J.S., A System of Logic, Parker, 1843.
- Milligan, David, 'Philosophy and the Social Sciences' Proceedings of the Aristotelian Society, 1968.
- Monnerot, Jules, Les Faits Sociales ne sont pas les Choses, Paris, 1946.

- Nagel, E., The Structure of Science, Routledge, 1961.
- Needham, Rodney, Structure and Sentiment: A Test Case in Social Anthropology, University of Chicago Press, 1962.
- Nisbet, R., The Sociology of Emile Durkheim, Oxford University Press, 1975.
- Nozick, Robert, The Normative Theory of Individual Choice, Princeton University Ph.D. thesis, 1963 (unpublished).
- Parsons, Talcott, The Structure of Social Action, Free Press, Glencoe, Illinois, 1949.
- Parsons, Talcott, 'The role of ideas in social action', (1938) in his Essays in Sociological Theory, Free Press, Glencoe, Illinois, 1949.
- Peacocke, Christopher, 'Truth Definitions and Actual Languages' in Evans and McDowell, eds., 1976.
- Pears, D.F., Ludwig Wittgenstein, Fontana, 1971.
- Peters, R.S., The Concept of Motivation, Routledge, 1958.
- Pitcher, George, ed., Wittgenstein: The Philosophical Investigations, Doubleday, 1966.
- Pitkin, Hanna Fenichel, Wittgenstein and Justice, University of California Press, 1972.
- Poggi, Gianfranco, Images of Society: Essays on the Sociological Theories of Toqueville, Marx, and Durkheim, Stanford, 1972.
- Popper, Karl R., The Poverty of Historicism, Routledge and Kegan Paul, 1961.
- Popper, Karl R., The Open Society and Its Enemies, Routledge and Kegan Paul, 1961.
- Plath, Sylvia, 'Daddy' (poem), in her book Ariel, Faber and Faber, 1965.
- Quine, W.V.O., 'Two Dogmas of Empiricism' in From a Logical Point of View, Harvard University Press, Cambridge, Massachusetts, 1961.
- Quine, W.V.O., Word and Object, The M.I.T. Press, Cambridge, Massachusetts, 1960.
- Quine, W.V.O., 'Truth by Convention' in The Ways of Paradox, Harvard University Press, Cambridge, Massachusetts, 1966.
- Quine, W.V.O., 'Foreword' to D.K. Lewis, Convention, 1969.

- Radford, Colin, 'Knowledge -- By Examples', Analysis, Vol.27,1966.
- Rawls, John, 'Two concepts of rules', Philosophical Review, Vol.64,1955.
- Rawls, John, A Theory of Justice, Harvard University Press, Cambridge, Mass, 1971
- Raz, Joseph, The Concept of A Legal System, Oxford University Press, 1970.
- Raz, Joseph, Practical Reason and Norms, Hutchinson, 1975.
- Rhees, Rush, 'Can there be a private language?' Proceedings of the Aristotelian Society, Supplementary Volume, XXVIII, 1954.
- Richards, David, A Theory of Reasons for Action, Oxford University Press, 1971.
- Rollin, Bernard E., Natural and Conventional Meaning: An Examination of the Distinction, Mouton, The Hague, 1976.
- Rudner, Richard, The Philosophy of Social Science, Prentice-Hall, Englewood Cliffs, N.J., 1966.
- Runciman, Walter G., A Critique of Max Weber's Philosophy of Social Science, Cambridge University Press, 1972.
- Ryan, Alan, The Philosophy of the Social Sciences, MacMillan, 1970.
- Ryan, Alan, ed., The Philosophy of Social Explanation, Oxford University Press, 1973.
- Saran, A.K., 'A Wittgensteinian Sociology?' Ethics, Vol.75, 1964.
- Sartre, Jean-Paul, L'Être et Le Néant, Gallimard, Paris, 1943.
- Saussure, Ferdinand de, Course in General Linguistics, Methuen 1957.
- Schelling, Thomas C., The Strategy of Conflict, Oxford University Press, 1960.
- Schneewind, J., ed., Mill, MacMillan, 1969.
- Schiffer, Stephen R., Meaning, Oxford University Press, 1972.
- Schutz, Alfred, Collected papers: I. The Problem of Social Reality, Martinus Nijhoff, The Hague, 1967.
- Schutz, Alfred, The Phenomenology of the Social World, Heinemann, 1972.
- Schwayder, D.S., The Stratification of Behaviour, Routledge and Kegan Paul, 1965.
- Scarle, John, Speech Acts, Cambridge University Press, 1969.

- Georg Simmel, 'How is Society Possible?' (1908) reprinted in
Georg Simmel: On Individuality and Social Forms, ed. D.N. Levine, 1971.
- Skinner, Quentin, '"Social meaning" and the Explanation of Social Action'
in Laslett, ed., Vol 4, 1972.
- Specht, Konrad, The Foundations of Wittgenstein's Later Philosophy, 1967.
- Sprott, W.J.H. Science and Social Action, Watts and Co., 1954.
- Slote, Michael, 'The Theory of Important Criteria', Journal of Philosophy, Vol.63, 1966.
- Stich, S.P., 'Grammar, Psychology, Indeterminacy', Journal of Philosophy, 1972.
- Strawson, P.F., Individuals: An Essay in Descriptive Metaphysics, Methuen, 1965.
- Taylor, R., 'Knowing that one knows' Analysis, Vol. XVI, 1955.
- Ullmann-Margalit, Edna, The Emergence of Norms, Oxford University Press, 1977.
- Von Neumann, John, and Morgenstern, Oscar, The Theory of Games and Economic Behaviour, 1953.
- Weber, Max, Economy and Society (3 Volumes) Bedminster Press, eds. G.Roth and C. Wittich, 1968.
- Weber, Max, The Theory of Social and Economic Organization, ed. Talcott Parsons, tr. Parsons and Henderson, A.M., Free Press, 1964, (includes Part I of Economy and Society).
- Weber, Max, The Methodology of the Social Sciences, Free Press, 1949.
- Weber, Max, 'R.Stammlers "Ueberwindung" der materialistischen Geschichtsauffassung' (1907) in Gesammelte Aufsätze zur Wissenschaftslehre, Mohr, Tübingen, 1922.
- Wells, Rulon S., 'De Saussure's System of Linguistics' Word, 3, 1947.
- Wilson, Bryan, Rationality, Basil Blackwell, 1970.
- Winch, Peter, 'Social Science' British Journal of Sociology, Vol.7, 1956.
- Winch, Peter, The Idea of a Social Science, Routledge and Kegan Paul, 1958.

Winch, Peter, 'Nature and Convention', Proceedings of the Aristotelian Society, Vol.LX, 1959-60.

Winch, Peter, 'Understanding a Primitive Society', American Philosophical Quarterly, Vol.1.,1964.

Winch, Peter, 'Mr.Louch's Idea of a Social Science', Inquiry, Vol.7, 1964.

Winch, Peter, 'Comment' (to Jarvie,1970), in Borger and Cioffi,eds.,1970.

Wittgenstein, Ludwig, Notebooks, 1914-16, Basil Blackwell, 1961.

Wittgenstein, Ludwig, Remarks on the Foundations of Mathematics, Basil Blackwell,1956.

Wittgenstein, Ludwig, Philosophical Investigations, Basil Blackwell, 1972.

Wolff, Kurt H., ed., Emile Durkheim (1858 - 1917), Ohio State University Press, Columbus, Ohio, 1960 .
