MO.C4f

QUANTITATIVE INTERNATIONAL POLITICS AND THE FOREIGN POLICY MAKER

Patrick Laver

Department of Political Science
University of Ottawa
1986

© Patrick Laver, Ottawa, Canada, 1987.

UMI Number: EC55319

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.
In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI®

UMI Microform EC55319
Copyright 2011 by ProQuest LLC
All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

ACKNOWLEDGEMENTS

The author wishes to express his warmest thanks to Dr. John Sigler for his patient supervision of the preparation of this essay and to Dr. Leonard Lefkovitch for his comments on statistical points in section II. The author is likewise greatly indebted to Irmgard Jaknunas for her scrupulously careful transcription of the manuscript. For errors that remain, he is solely responsible.

PREFACE

The essay that follows begins with the words "This essay is just that ...". As it was being written, it was an exploration. The journey mattered more than the destination and, at the start, it was uncertain what the outcome would be. It was an exercise in what the philosopher, Gilbert Ryle, called "path-finding" as opposed to "path-making". No hypotheses were being put to the test. It was, rather, that questions were being asked.

The principal question throughout was whether those scholars working in the field of Quantitative International Politics who seek, in attempting to measure empirical phenomena in the world, to arrive at lawlike generalisations as a basis for prediction, have produced, or can produce, findings which would be of direct value to those who participate professionally in the formulation of foreign policy. The specific question addressed in the Introduction was why, in Britain at least, this whole line of enquiry should have been virtually overlooked.

The case study on "Predicting Political Instability in Tropical Africa", discussed in section II, was chosen precisely because the authors were untypical of other scholars in the field in claiming that, on the basis of their results, their methodology could usefully be applied by analysts

working for the United States government. The critique in that section was itself essentially methodological, querying, on the basis of their own criteria, the validity of some of the concepts as "operationalised" and their somewhat cavalier use of statistical techniques.

John Vasquez' "Colouring it Morgenthau" thesis, discussed in section III, was chosen, at the next stage of the enquiry, because he was untypical in offering a general explanation for the generally admitted failure of the field so far to "produce knowledge" (that is, scientific knowledge as a basis for scientific predictions). He attributed this failure to the inadequacy of Political Realism as a paradigm for guiding research - and this was another reason for the choice since, if his arguments were sound, his conclusion could be useful (if unappealing) to many traditionally-minded foreign policy makers, pointing to a need for them to re-think the way in which they conceptualise the world. The critique here was largely in terms of the philosophy of science, in relation to the work of Kuhn and Lakatos and to the standard literature on statistical significance.

At this point in the journey, when a claim to have produced a success for the field and a claim to have produced an explanation for the general lack of success had both proved ill-founded, it seemed natural to ask whether there might be

some alternative explanation for the field's inability to produce results which could convince foreign policy makers. In section IV the critique was conducted purely in philosophy of science terms and was directed, not at specific targets, but more broadly at those scholars in the field who accept, implicitly or explicitly, the feasibility of applying the so-called "covering law" model of explanation to human affairs or of discovering similarly applicable lawlike generalisations. So the essay ended not with a formal conclusion but with a further question - in effect: What now?

SUMMARY

For a number of reasons, Foreign Policy makers, in Britain at least, have shown little interest in the work of scholars in the field of Quantitative International Politics. Yet, if progress in theory-building has so far been disappointing, policy makers would be quick to take notice of theories with good predictive power.

In the United States, the State Department has been more adventurous, commissioning a study by Michael O'Leary and William Coplin on Quantitative Techniques in Foreign Policy Analysis and Forecasting (1975). A detailed analysis of the case-study on "Predicting Political Instability in Tropical Africa", of which Patrick J McGowan was the principal author, leads to the conclusion that it was not an unqualified success.

The doctrine of Political Realism, which continues to exert an influence among Foreign Policy makers, has been held to blame by John Vasquez, in his The Power of Power Politics: A Critique (1983), for some part of the failure by QIP scholars to produce much in the way of "strong findings". An analysis of his "Colouring it Morgenthau thesis" suggests that his critique may be wide of the mark.

There is no consensus among QIP scholars on a concept of theory. If the so-called Deductive-Nomological model is in-

creasingly seen, and for good reason, to be inappropriate, there is still a hankering for "lawlike generalisations", the attainment of which in International Relations Theory seems problematic.

TABLE OF CONTENTS

Acknowledgements i

Preface ii

Summary v

Table of Contents vii

Section I: Introduction 1

Section II: A Case Study of a Case Study: Patrick J McGowan et al., on "Predicting Political Instability in Tropical Africa" 14

Section III: A Critique of a Critique: "The Colouring it Morgenthau" thesis of John A Vasquez 51

Section IV: Theory and Change 80

References: 112

I Introduction

1 This essay is just that: an attempt - an attempt to probe the question of the value to the Foreign Policy maker of the research in Quantitative International Politics devoted to the discovery of laws, or lawlike generalisations, in the processes at work in the international "system". Quantitative International Politics (QIP for short) is a broad field. It might be said to encompass decision-making theory, games theory and the use of simulation techniques, all of which may be of interest to the Foreign Policy maker (or FPM for short) in that they focus directly on his or her role. But for the purposes of this essay, the prime concern is with the possibilities of describing the unfolding of events on the international scene through the use of general or partial theories with explanatory or predictive power. Can the application of scientific methods to the study of International Politics (meaning by "scientific methods" those in use in the natural sciences) produce results capable of enhancing the FPM's understanding of the world in which he must operate?[1)]

[1)] We follow the practice, fairly common in the literature, of using the masculine pronoun to refer to both sexes except where a reference to one or other is clearly intended, being obvious from the context. Anatol Rapoport, in his Mathematical Models in the Social and Behavioral Sciences (1983), uses pronouns of either gender to refer to both; this is an engaging, but confusing, effort to be even-handed.

2 Quantitative International Politics is not a totally satisfactory term for the types of research it designates. Others have been suggested: "Interpolimetrics" or the "Scientific Study of International Politics", the label preferred by Dina Zinnes (1976 p1 ff). But QIP has come to stay. It denotes a well-established field of research and the scholars engaged in it are sometimes collectively described as "the field" - like fox hunters, but riding to methods (in place of hounds) in pursuit of theories (not foxes). To name the principal riders would be one way of providing an illustrative definition of the field and avoid the danger, inherent in any formal definition, of failing to discriminate between their individual philosophies and methodologies. Yet it would be inadequate, not only because there could be no guarantee that the sample selected would be representative but also because, despite differences of approach, some broad definition is required to make enquiry feasible. Somewhat loosely, therefore, let the field be delineated as embracing those scholars who share a belief that causal patterns in international affairs can be discerned through the use of quantitative techniques. James Rosenau, a distinguished member of the group, has summarised his credo as follows: "While the paucity of knowledge in this field (sc. that of foreign policy phenomena) is indisputable, it does not mean that the behavior of nations cannot be subjected to the data-gathering procedures and criteria of evidence that are the hallmark of the scientific method. As a focus of study, the nation-state is no different from the atom

or the single cell organism." (1981 p 32). The general objective is a cumulation of knowledge in the form of theory (Hopmann, Zinnes & Singer 1980 p 1). Zinnes would add, however, that "while measurement plays a role (sc. in the work of the QIP scholar), it does not represent the central feature. What distinguishes both the ad hoc hypothesis tester and the modeler from his more traditional colleagues is the explicitness with which his research is performed" (emphasis in the original) (1976 p 15), thus permitting theories to be put to the test.

3 The term Foreign Policy maker is not totally satisfactory either. It suggests the decision-maker, the prime mover in the determination of a country's foreign policy, a Talleyrand, a Metternich or a Bismarck, for example. The net should be thrown more widely. Groups of individuals as well as individuals may take the final decisions. Professional advisers may recommend what the decisions should be, and these in turn may be advised by specialists more concerned with the analysis and interpretation of events, and of their possible development, than with policy-recommendation. The term is used here to cover all these categories. It excludes those interested in influencing policy from the outside. In particular, it demarcates, to borrow language from Milton, those in the "dust of the arena" (not always so dusty) from those in the "cloistered calm" (not always so calm) of academe. Whatever his personal predilections, the scholar is not obliged by his calling to be "policy-relevant", nor obliged not to be.

4 The Policy maker does not set the aims of scholarship although, if the piper wishes to be paid, he should not object if the payer often wishes to call the tune. What is wanted will find its market. But the policy maker's appreciation of the value of scholarship can be clouded by prejudice. This can take the form of a general distrust. "On devient forgeron en forgeant". The British example is instructive. Sir Robert Cotton, whose books formed the nucleus of what is now the British Library, failed to persuade Queen Elizabeth I to establish an academy for the study of history to train young noblemen for government service. Over a century later, in 1724, when King George I founded the Regius Professorships of Modern History at Oxford and Cambridge to train recruits for foreign service, the posts were well-paid sinecures: Modern History (that is, history from about 500 AD on) was not part of the curriculum and the professors were not obliged to teach (Kenyon 1983 pp 2 & 58-59).

5 Wicquefort, whose classic work L'Ambassadeur et ses Fonctions was highly regarded in the eighteenth century as a bible for professionals was scathing about "les gens de lettres ... qui ont contracté une trop grande habitude avec les livres, qui ont une trop forte liaison avec les prejugés des Docteurs, & qui ont plus de lecture que de bon sens" (1690 p 78). The emphasis on "good sense" is illuminating. To revert again to the British tradition, it has long been

accepted that the "negotiator" (the terms "diplomat" and "diplomatist" did not enter the language until the nineteenth century) needs not so much savoir as savoir faire. The advice proffered to future and present diplomats concentrated on the skills and qualities they should cultivate in order to be able to function effectively. This can be seen in François de Callière's De la Manière de négocier avec les Souverains (1716; English translation, 1983) in Lord Chesterfield's famous Letters to his son (1774/1893) and, more recently, in Harold Nicolson's Diplomacy (1952).

6 Some knowledge of history, the record of what has happened up to now, has of course always been regarded as indispensable for anyone participating in public affairs. And many historians, though by no means all, have agreed with the thought expressed by Bolingbroke (1754 Vol II p 345) that it was a major part of their business "to teach and inculcate ... the general rules of wisdom and good policy". In similar vein, Chesterfield stressed the importance of an acquaintance with the 'maxims of foreign courts', that is, very roughly, of the principles guiding policy (he himself being attracted to those of the Cardinal de Retz). But a widespread attitude among historians and policy-makers alike is well summed up in Jakob Burckhardt's ironical comment that "the true use of history is not to make men more clever for the next time but wiser for ever".

7 "International Relations" (the term was coined by Jeremy Bentham) first came into its own as a distinct academic discipline in the aftermath of World War I. Chairs were founded at universities in Britain and North America, often endowed by wealthy benefactors. A need was felt for research into the causes of war and the means of its prevention. Nevertheless, in Britain at least, the influence of scholars in the field on policy remained, and still remains, negligible. The few individual scholars whose views carry weight with officialdom tend to be area specialists, not theorists; and no great value is attached to theory and theorising, although they are seen by some to have marginal usefulness in broadening minds and in justifying, rather than guiding, policy (cp Cable 1981, Frankel 1981). It is symptomatic of the general sentiment in Britain that International Relations as a subject of study has little vocational relevance that, in 1977, fewer than one per cent of the members of the Diplomatic Service were graduates in the field.

8 If the value of theory in International Relations is suspect, that of attempts at scientific theorising is doubly so, British officialdom seems to have been content, indeed happy, to accept the verdict of British academics that there is little in it. Hedley Bull's celebrated paper "International Theory: the Case for a Classical Approach" did not have to set the tone; it reflected it. No British critic has attacked

Bull's conclusion that, while there was merit in a number of contributions made by theorists who adopt a scientific approach, "what is of value in them can be accommodated readily enough within the classical approach ... The distinctive methods and aspirations these theorists have brought to the subject are leading them down a false path, and to all appeals to follow them down it we should remain resolutely deaf" (ibid p 38).

9 No one would wish to accuse Bull of anti-Americanism even if it is undeniable that the targets of his attack were scholars who were American, or who were working in America, and that he regarded the very conception of a "science of international politics" as peculiarly American. The fact remains that, whereas such a conception has taken root and flourished in the United States, in Britain it has fallen almost wholly on stony ground. Steve Smith (1982) attributes this to what he calls "epistemological" bias. "The Comparative Foreign Policy school" he writes, making clear by this designation that it has no monopoly in the United States, "stresses the regularities of foreign policies; the British school concentrates on their uniqueness. In contrast to CFP's belief in the essential comparability of foreign policy, the British school emphasises the need for different approaches and methodologies for the analysis of different states' foreign policies. Rather than generate knowledge by the creation of falsifiable hypotheses and testable evidence,

the British school stresses the need for intuition and insight. Instead of the systematic collection of data, utilising, wherever possible, quantitative techniques, the British school argues for the importance of common sense and the selection of evidence."

10 That a British scholar in International Relations should refer to quantitative techniques at all in a discussion of scientific approaches to the study of the subject is unusual. Bull's prime target was Morton Kaplan (1957) although he comments briefly, and critically, on the use made of quantitative analysis by Karl Deutsch and Bruce Russett. Similarly, Charles Reynolds, in his widely read, but not very readable, Theory and Explanation in International Politics (1973) focuses on Kaplan (ibid), CA McClelland (1961), David Easton (1965) and Snyder, Bruck and Sabin (1962). The list could be multiplied indefinitely. In brief, among those scholars who adopt the scientific approach, it is the 'modellers' who are singled out for attention; the QIP community is virtually ignored. The one apparent exception is James Rosenau, from whose pen an important collection of essays was recently reissued in London, (1981) but the interest this attracted focused on his development of concepts such as "Linkage Politics", "Adaptation" and "Issue Area" rather than on his central commitment to the quest for scientific theory through quantitative methods (cp Steve Smith 1979).

11 The QIP scholar faces a double difficulty in finding a receptive audience among the FPM community. In addition to the lack of a common purpose or focus of interest between the two, there is the lack of a common language to the extent that quantitative method finds expression in the 'language' of mathematics, still regarded by many conventional FPM's as alien, if not hostile, territory. Speaking of his own "International Systems Theory", Morton Kaplan acknowledges (1969) that "it is not easily adaptable to the microstructural problems of foreign policy". The same could be said of any scientific approach; for such approaches are concerned to identify and account for the broad patterns and regularities underlying the process of international affairs and this, in the first instance, for the purpose of theory-building. The FPM, on the other hand, faced, as he often is, with immediate, and specific, practical problems, operates within a "macrostructure" which in a sense is already "given". The difference is reflected in the different sets of "variables" (though they overlap) investigated by the scientific scholar and the practitioner, with the latter leaving no detail unexamined in the search "for factors which can be manipulated to change a situation" (O'Leary & Coplin 1975 pp 18-19). But what really sets quantitative analysis apart and makes it rebarbative to the FPM is its 'methodological complexity' (Raymond Platig in his foreword to ibid p v). There is, moreover, some reluctance on the part of the FPM to be treated as himself a

variable, witness Dean Acheson's reaction when he discovered that he was so described in a study of the American decision to resist the invasion of South Korea by the North in the 1950s (ibid p 17).

12 Some QIP scholars go so far as to say that "quantitative work ... could not and should not (emphasis in the original) appeal to highlevel policy-makers" because "highly jargonised and abstract, it is alien to most policy-makers" and "inappropriately aimed" at people "who are disinterested (sic) in the results of a Markov analysis" (Wilkenfeld et al., 1980 p 249). This shows undue timidity. Oran Young displayed greater realism when he wrote (1972 pp 199-200) that "if theorists of international relations could produce even one striking success, it is hard to imagine that policymakers would refuse to pay attention to them. That is, if the theorists could produce consistently accurate predictions about any major phenomenon in the realm of international relations (such as power relationships, wars, crises, interventions or alliances), there is every reason to suppose that policymakers would take their results seriously." He makes the point that "theorists who can apply and manipulate (viable) theories ordinarily gain authority through the accuracy of their predictions rather than through any elaborate program of instructing policymakers in the intricacies of their theoretical models ... A great many policymakers dealing with economic problems have no more than a general sense of the theoretical bases of modern economics." Even at the humblest level in the

FPM community, that of the analyst who is concerned with the interpretation and forecasting of events rather than with decision taking, it is understandable that an inability or reluctance to master new techniques of potential relevance to personal expertise should induce intellectual discomfort. But at every level, what matters is that a theory should prove its predictive worth.

13 Success is the touchstone; and it has, so far, been elusive for the QIP community, as its members acknowledge. The authors of Foreign Policy Behavior (1980 p 239) confess to their "ultimate inability to generate any empirical, comparable (ie operationalisable) measures for phenomena housed within the political component (sc of scientific foreign policy inquiry)". The editors of Cumulation in International Relations Research (Hopmann et al., 1980 p 7) conclude that the papers it contains, all delivered at the annual convention of the International Studies Association held in Toronto in 1979, "indicate that the cumulative development of international conflict research still leaves much to be desired"; for what is termed "cumulation of knowledge" is regarded by the community as a test of the scientific nature and quality of a scholarly enterprise. Rosenau himself, to judge from his book The Scientific Study of Foreign Policy (1981), continues to seek new ways of laying a satisfactory foundation for the development of theory. While, however, a respectable number of books and

articles continues to flow from the pens of members of the QIP community, there is little indication that they are attracting widespread attention.[2] Interest seems to be waning in the Correlates of War project initiated by David Singer and in the Dimensions of Nations project of RJ Rummel. The Inter-University Comparative Foreign Policy Project, inspired by James Rosenau, has folded. History is inter alia, says Theodore Zeldin, the author of the monumental France 1848 to 1945, the history of the rise and fall of theories. But it would be premature to claim that the QIP movement has run out of steam. That inveterate optimist, Dina Zinnes, concludes her review of the difficulties encountered by QIP scholars in assembling "evidence on the outbreak of international conflict" (Gurr 1980 p 360) with the remark that, although "we (sc the QIP community) do not have laws, the empirical findings that have been produced, and are being produced, appear to be moving us slowly toward lawlike generalizations".

---

2) An informal survey of campus bookstores across Canada, from the Fall of 1984 to the Summer of 1985, has so far failed to detect any such works in their "required reading" sections; the following universities were covered: Victoria, British Columbia, Simon Fraser, Calgary, Alberta, Manitoba, Ottawa, Carleton, McGill, Montreal, Sherbrooke, Queen's. The hypothesis that, in every case, they were all sold out is implausible.

14 Is such optimism, even for the long term, justified? Data on international phenomena have been amassed on a vast scale, though in accordance with different taxonomies, which creates problems for comparison. An impressive number of "findings" has been published (cp Singer 1968 and Jones & Singer 1972). The testing of what Zinnes calls "ad hoc hypotheses", that is of hypotheses not deductively derived, has proceeded apace. Yet, to cite Zinnes and her colleagues again (Hopmann et al 1980 p 8), a creative, theoretical leap is needed to produce the integrated picture which is the goal of QIP research. This would seem a necessary, but not a sufficient, condition for a successful marriage between a fully fledged science of the probable and the art of the possible. A flirtation on the way need not be excluded. In the natural sciences, after all, it was feasible to make predictions on the basis of the generalisations of Kepler and Galileo before Newton formulated his laws, and on the basis of Boyle's Law and Dalton's Law before the establishment of kinetic theory. So, by way of illustration, we shall examine in section II a specific attempt to predict political instability in tropical Africa.

II A Case Study of a Case Study: Patrick J McGowan et al on "Predicting Political Instability in Tropical Africa"[1)]

1 In the mid-Seventies, an enlightened State Department commissioned, through its Office of External Research, a study (O'Leary and Coplin 1975) into whether "some of the methods of (quantitative) social scientists intent on making their work more scientifically rigorous could be applied, with advantage, by government analysts intent on preserving or increasing the more or less immediate usefulness of their work" (ibid, Foreword by Raymond Platig, as Director of the Office, p v). O'Leary and Coplin concluded (ibid pp 256 - 258) that quantitative analysis could help in several ways. It could force the analyst to be more self-aware than was the case with purely qualitative analysis, to make explicit those factors considered important, the way to measure them and the nature of relationships among them. While "the clearest evidence for the application of quantitative techniques is in the field of acquiring information, by developing coding systems to collect data from written material ... and expert-generated data systems to formalize the exchanges of views that now characterizes (sic) the information-gathering tasks of the ... analyst, there is some evidence that in areas such as political instability, international violence, and the international bargaining process, quantitative techniques can be used to generate predictions as well as identify those factors that the analysts should consider in making their predictions."

[1)] Patrick J McGowan is stated in the Introduction to Quantitative Techniques in Foreign Policy Analysis and Forecasting to be the "principal author" of this study.

2 The authors based these conclusions on the results of six case studies of which only one will be considered here: the chapter on "Predicting Political Instability in Tropical Africa". The reason for singling it out is simply that they regarded it as more successful than the others: they state baldly (ibid p 24) that it "successfully tests a predictive model of political instability in (Subsaharan) Africa (less the Republic of South Africa) based on an index of how prone each state is to coups d'état". It is interesting to note in passing that their professional pride as social scientists took precedence over their doubts, as socially concerned citizens, over the ethics of providing a tool of analysis which might be put to use in furtherance of an aspect of American foreign policy with which, at the time of writing (1974), they were evidently unhappy.

3 McGowan and his co-authors, very sensibly, take as their starting point the views and attitudes of the International Research experts in the State Department. They were given full access to the analysts concerned and were allowed to examine a wide range of official documents (though, for obvious reasons, the latter did not include material with a high security classification). They found, not surprisingly, that the prime concern of the INR analysts was with the short term prospects for political instability in individual countries. The "case-by-case approach" they write (ibid p 37) "is so ingrained in the Department's normative and intellectual

environment and bureaucratic structure, that analysts simply do not think in comparative terms." Again (ibid p 40), "On the conscious and direct level of analysis, analysts seem to agree with (those) scholars who argue that there are no systematic factors associated with instability and that these phenomena occur randomly throughout Africa. In short, no general explanations of African political instability are possible. The best one can do is to offer ad hoc and post hoc explanations, once a coup has occured in a given country. But", the authors continue, "on an implicit and indirect level (one might almost say unconscious level), the analysts appear to rely on what can be called an interest-group theory of political instability in Africa." They characterise such a theory as one which argues, very roughly, that, lacking other resources in the form, for example, of moral authority, money or efficient machinery of government, African régimes increasingly rely, when faced with demands from ethnic interest groups in the rural sector and occupational interest groups in the urban areas, on the use of force to obtain compliance from masses and élites and that this tends to lead to direct or indirect military intervention and communal violence. The authors announce their intention of putting both this implicit theory and "the explicit conception that coups and instability are random" to the test.

4 Before doing so, they make an important statement about the limits to predictability. "Insofar as

(INR analysts) deny the possibility[2] of predicting discrete events in particular countries, there can hardly be any doubt that they are correct. Even the most optimistic social scientist would also agree that such precise prediction is impossible at present. However, many social scientists - and some analysts - would agree that probabilities can be attached to the likelihood (sic) of instability in a given country and that probabilistic predictions can be made about the frequency and level of instability phenomena in Africa for a given time period. These types of predictions are the only ones (sic) that can be made from statistical generalizations and thus the only ones ever made by social scientists. We cannot predict how a particular individual will vote, but we can predict quite well what proportion of people with a given set of characteristics will vote Republican or Democrat" (ibid p 40). This passage is revealing. There is the hint that precise prediction of international phenomena may eventually prove possible, though presumably not on the basis of statistical generalisations or, at least, not on the basis of such generalisations alone. The implication, in other words, is that some day it could become feasible to progress beyond assessments of probability, to aim for a probability of 1.0, or certainty. It would be unreasonable to take the authors to task in this

---

2) "possibility" is clearly meant though "impossibility" appears in the text.

context for what is clearly an obiter dictum but the fact that the statement could be made at all is of interest as reflecting a cast of mind. So too is the denial of the possibility of predicting discrete events. In the real world, this is manifestly not the case. Predictions can be, and are frequently, made with a high degree of confidence subsequently justified. It is possible to know enough about an individual voter, in specific cases, to be virtually certain that he will vote for a particular party or abstain.

5 After this preliminary, the authors embark on their main purpose: "to demonstrate that a comparative study of political instability in tropical Africa can yield theoretical relationships sufficiently strong to build a system for statistical forecasts" (ibid p 43). They begin with an explanation of what they mean by "scientific definitions", the absolute prerequisite for quantitative analysis. Stressing that a "fundamental application of social science methods to foreign affairs analysis would be the clear-cut and consistent use of words to refer to various foreign affairs phenomena", they claim that there is no such thing "as a right or wrong scientific definition, only a definition that is more or less useful". What "X" 'really' is, is not the concern of social science. "Usefulness" is measured by the extent to which concepts, stipulatively defined, "can enter

into general sentences (sic)[3] that explain the phenomenon under investigation." There are no a priori criteria for evaluating scientific definitions other than that each must stipulate at least one empirical referent so that the presence or absence of the defined trait can be observed in the real world. At the same time, the authors add, "our (scientific) definitions can be evaluated only after empirical research has been undertaken" (ibid p 44).

6 What they mean by "evaluation" in this last phrase is not altogether clear. Presumably it is an evaluation of empirical adequacy or "fit". But how should the evaluation be carried out? On the basis of the assertion that there are no "right or wrong" scientific definitions, and of the insistence that such definitions are essentially stipulative, it looks as though, in the authors' view, empirical "fit" is established, not directly through a comparison of concept and "real world" phenomena, but indirectly through the explanatory (presumably, therefore, also predictive) success of the "general sentences" in which they feature. In the natural sciences, however, it is possible for definitions to be right

---

[3] Many philosophers would prefer, in such a context, to use the word 'statement' or, perhaps, 'proposition'; a 'sentence' is a form of words forming a grammatical whole becoming a 'statement', or other kind of expression, only when put to use in a specific situation.

or wrong. Contemporary physicists would agree, for example, that Thomson, who discovered the electron, was wrong in characterising it as well-defined in regard to spatial volume and that Bohr was wrong in characterising it as capable of having any energy level (W. H. Newton-Smith 1981 p 160). This did not prevent the definitions of these two distinguished scientists from being "useful" in the sense attached to this term by the authors, because, if partially wrong, they were partially right, and because the definitions themselves could be put to the test. The properties of an electron can be investigated empirically and established definitively. Those of a phenomenon such as political instability cannot: the concept of "political instability" is elastic. The authors are correct in suggesting that, to make it precise enough for mathematical manipulation, some at least of its defining characteristics must be stipulated, and this with some degree of arbitrariness. They are also correct in stating that, to be empirically relevant, such a concept must "contain" empirical referents ("at least one"). In such circumstances, however, to the extent that such referents are not exhaustive, in leaving no room for additional independent defining features, what is explained or predicted by general statements is not so much the conceptualised phenomenon as its stipulated empirical features. The latter become indicators of the presence of a phenomenon, and not necessarily reliable ones, because different phenomena could share one or more of the characteristics which constitute, or contribute to, their definition. Serious prac-

tical difficulties can arise in attempting to reconcile two objectives simultaneously in the formulation of concepts: the need to preserve the "validity" of a concept, its correspondence with a phenomenon in the real world, and the need to give it a measurable precision so that it can be quantitatively related to other measured concepts. Between "correspondence" with truth and mathematical tractability, there is often tension.

7 Having defined "scientific definition", the writers of the chapter go on to offer a scientific definition of "political instability". This is first, rather curiously, stated to be a characteristic of governments but the authors soon correct themselves, defining it as "a condition affecting governments in which the established patterns of authority break down, and the expected compliance to (sic) the government is replaced by political violence ... with an intent to bring about an alteration in the structure of the government." Political instability is said to take two "basic and statistically unrelated" forms: <u>élite</u> and <u>communal</u> instability respectively (Morrison & Stevenson 1971), 'basic' meaning 'conceptually distinct' and 'statistically unrelated' meaning that "instances of each type of instability do not occur together at the same time in the same countries" (sic). These are said to be the empirical referents of "political instability" defined in terms of "violence between political actors in conflict over the decisions governing the distribution of rewards in society."

8 Elite instability refers, say the authors, to events involving the threat or use of violence by members of an élite, political or military, for example, "to remove persons from their command positions in the national government" (ibid p 45). It is represented by three distinct types of behaviour: coups d'état, sudden and illegal displacements of existing political régimes by relatively small élite groups without mass participation; attempted coups, in which it is known that some combination of the following occured: the assassination, attempted assassination or arrest of some members of the governing élite; a disruption or takeover of government facilities lasting less than a week; sudden action by the armed forces explicitly aimed at the takeover of government; plots, as officially reported by the ruling authorities. To create a scale of élite instability for tropical African states, coups, attempted coups and plots are given weights of 5,3 and 1 respectively, each multiplied by the number of years for each country for which it was reported.

9 Communal instability is graded similarly (ibid p 47). It refers, say the authors, to events in which communal groups "use violence to change the distribution of authority within the general population or between the government and the group". It is represented by four distinct types of behaviour, each empirical instances of it: civil war; rebellions, either to gain increased autonomy from the national government or to

attack its supporters or agents "without aiming to secede from or monopolize the existing polity"; irredentism, where a communal group seeks to change its political allegiance from the "government of (its) territorial unit to a government, either existing or to be created, in which the decision makers share the communal identification of the ... group concerned"; ethnic violence, events of short duration in which "members of two identifiable communal groups are antagonists in violence not designed to secure independence, autonomy or political realignment". These four types of behaviour are given weights of 5,4,3 and 1 respectively, the last being attributed because ethnic violence (as defined) is less dangerous for the existence of governments.

10 What emerges immediately is that the short verbal definition of "political instability" (reproduced in paragraph 7 above) corresponds neither with common usage nor with the explications offered of élite instability and of communal instability which are said to be its types. Political instability is not a characteristic of governments but of countries or systems of government. A country can be said to be politically unstable when it undergoes frequent changes of government. There need be no breakdown in the authority of the State; a legislature may remain the arbiter; and, despite the vicissitudes of the politicians, a powerful bureaucracy can ensure orderly administration. Paradigms (to use the word in its conventional sense) of such a state

of affairs can be seen in much of post-War Italy and in France under the Fourth Republic. These two examples show also that changes of government in a politically unstable system need not, though they certainly can, be the product of a threat or use of force.

11 The authors would be entitled to counter this objection by arguing that their concern is with political instability in tropical Africa which has characteristics of its own, more resemblant of Latin America than of Western Europe. This still leaves the problem of the nature of the relationship between the concept of "political instability" and those of "élite" and "communal" instability respectively; for the instances they give of the latter are not instances of the former necessarily and in every case. A series of coups d'état in a particular country would indeed be a manifestation of political instability there. A single coup might not be, nor might relatively infrequent attempted coups or plots. Moreover, an "officially reported" plot can be an official fabrication, a convenient way of dealing with political opponents who are more of a nuisance than an immediate and serious threat to a régime (some of the reported plots against Kenyatta in Kenya and Mobutu in Zaire could fall into this category). Likewise, while it may be reasonable to regard any civil war or large-scale rebellion as itself an instance of political instability in a country (though not necessarily, let it be noted, of communal instability), not

all instances of irredentism or ethnic violence have to be so regarded. The spectrum of types of event held by the authors to be manifestations of political instability shifts from possible instances of it to possible threats to political stability and even beyond (to the extent that neither irredentism nor ethnic violence need pose a threat to a central government). In the process, "political instability" in itself ceases to be measured, ceases, indeed, to be the "dependent variable" under investigation.

12 Instead of one dependent variable, therefore, there are now two, élite and communal instability respectively, which call for "explanation". But, as it turns out, what the authors are really interested in is the concept of "coup-proneness", which has the double advantage of being both more precise and more suited to the needs of the in-house foreign policy analyst. "Which countries in tropical Africa are coup-prone?" is a more pertinent question for the FPM than "In which of these countries is there likely to be an increase in political instability?". In fact, the FPM would probably wish to see the former question divided into two separate questions: "Where (in tropical Africa) are coups likely to be attempted?" and "Where are they likely to be successful?". For this purpose, however, the scoring proposed for "élite instability" would not be appropriate. Attempted coups would include successful coups and genuine plots (it is usually not difficult to determine whether an

official report of a plot is authentic) where these last had been aimed at a take-over of power; and all "weights" would have the same value. So too would the "weights" for all successful coups. The two, overlapping but separate, categories would be necessary because, of course, the factors which make for a coup to be attempted are not co-terminous with those that make for its success.

13 That said, let it be assumed, for the sake of argument, that the authors' index for "élite instability" could serve as a rough indicator, if not as a measure, of "coup-proneness". The problem then is to establish, as they put it, whether "our measurement operations have produced scales and indexes of the phenomena in question that are 'good enough' to serve as the basis for research and possibly for policy recommendation" (ibid p 49). The adequacy of a scale depends in part on the reliability of the sources of information about the events of which account needs to be taken, and on the reliability of the "coders" in identifying such events and in attaching suitable "weights" to them. The authors are satisfied that their sources are sufficiently reliable, in that they have detected no bias in coverage (such as might occur, for example, if exclusively English-language sources were drawn upon for data on francophone countries), and that their coders are likewise sufficiently reliable in that there was agreement among them, and between them and the original investigators, about the events to select and the weights to

assign. But, while the coders no doubt conformed with the coding rules, the issue is, rather, whether the coding rules themselves were appropriate. This can have a bearing on the value of inferences drawn from a comparison of scales. Thus, in Table 1 below (ibid p 48), the authors point out, as part of an argument, that "the range in scores (the magnitude of the difference between the lowest and the highest value) is half again as great for communal instability as it is for élite instability (38 as compared with 26)" (ibid p 53) when this is, in fact, no more than an artefact of the weights.

TABLE 1

Scales of Elite and Communal Political Instability for 32 Independent Black African States, from the Date of Independence to December 1969

Elite Instability			Communal Instability		
Rank	Country	Value	Rank	Country	Value
1.	Dahomey	26	1.	Sudan	38
2.	Sudan	22	2.	Ethiopia	30
3.5	Zaire	20	3.5	Zaire	27
3.5	Togo	20	3.5	Nigeria	27
5.	Congo, B.	17	5.	Chad	17
6.	Burundi	16	6.	Rwanda	15
7.	Ghana	13	7.	Kenya	14
8.	Nigeria	12	8.	Uganda	12
9.	Sierra Leone	11	9.	Mali	6
11.	CAR	8	11.	Burundi	5
11.	Somalia	8	11.	Ghana	5
11.	Uganda	8	11.	Ivory Coast	5
13.5	Mali	7	13.	Congo, B.	4
13.5	Upper Volta	7	14.5	Somalia	3
15.5	Liberia	6	14.5	Zambia	3
15.5	Senegal	6	16.	Mauritania	2
17.	Chad	5	18.	Cameroon	1
18.5	Ethiopia	4	18.	Dahomey	1
18.5	Guinea	4	18.	Sierra Leone	1
20.	Gabon	3	26.	Botswana	0
21.	Ivory Coast	2	26.	CAR	0
24.	Cameroon	1	26.	Gabon	0
24.	Kenya	1	26.	Gambia	0
24.	Lesotho	1	26.	Guinea	0
24.	Malawi	1	26.	Lesotho	0
24.	Niger	1	26.	Liberia	0
29.5	Botswana	0	26.	Malawi	0
29.5	Gambia	0	26.	Niger	0
29.5	Mauritania	0	26.	Senegal	0
29.5	Rwanda	0	26.	Tanzania	0
29.5	Tanzania	0	26.	Togo	0
29.5	Zambia	0	26.	Upper Volta	0
Average value		7.19			6.75
Coefficient of variability		102.09			151.56
Standard deviation		7.34			10.23

Note: Date of independence for Liberia and Ethiopia set at 1956 to agree with the independence of Sudan.

Source: D. G. Morrison, R. C. Mitchell, J. N. Paden, and H. M. Stevenson, Black Africa: A Comparative Handbook (New York: Free Press, 1972), pp. 128-29.

14 The adequacy of a scale depends also on its validity, of which there are various forms. The least important is said to be face validity, the concordance of experts' opinion, with the resulting scores and rank orders. As the authors rightly remark, expert opinion can be wrong. Next comes content validity, depending on the extent to which the events and types of event taken into account are held to be representative of the concept being made operational. While this may be largely a matter of judgment (again by experts), factor analysis has, the authors say, shown that the types of event contributing to the operational definition of élite instability are associated with each other. Of course they are: there can be no coup without an attempted coup, and few attempted coups without a plot. But this does not justify the scaling proposed (cp paragraph 12 above). A test of construct validity requires, to be positive, a high correlation between the results of one attempt to measure a phenomenon and those of another, using "different indicators". Here, the authors draw a partial blank because, when they wrote, the only other social scientist, to their knowledge, who had embarked on a similar enterprise (besides the scholars furnishing the results they themselves reproduce) was Ted Gurr (1968) and, while there was a strong correlation ($r = 0.72$) between his measure of internal war and theirs of communal instability, there was none at all between his of conspiracy and theirs of élite instability. But such a test is clearly neither possible nor necessary in every case. If, for example,

the authors had chosen to measure "coup-proneness" in tropical Africa and to do so on the basis of the actual number of attempted coups in each country, no other measure could qualify as well.

15 Finally, there is criterion validity. This hangs on the proposition that, if a theory or hypothesis warrants the expectation that the scales of measurement of two sets of phenomena (such, for example, as intelligence and academic achievement) are, ceteris paribus, closely associated and if the subsequent calculation of a correlation co-efficient confirms the expectation, then it is legitimate to infer that each of the two scales provides a valid measure of what it purports to measure. "If" say the authors correctly "this sounds like circular reasoning, it is, but", they continue, "it is not illogical as long as our theoretical expectation is stated prior (emphasis in the original) to looking at the correlations (ibid p 51). On this basis, however, such inference would not be justified. It is certainly logically possible for the measures of two sets of phenomena to be positively correlated to a high degree in fulfilment of an expectation stemming from a false or unproven theory or hypothesis. This would do nothing to confirm that the scales were valid empirically. It would be necessary for the theory or hypothesis in question to be held to be true or plausible independently of the results of a calculation of "r". Even where this necessary condition was met, however, it would still not be sufficient to guarantee that the two scales

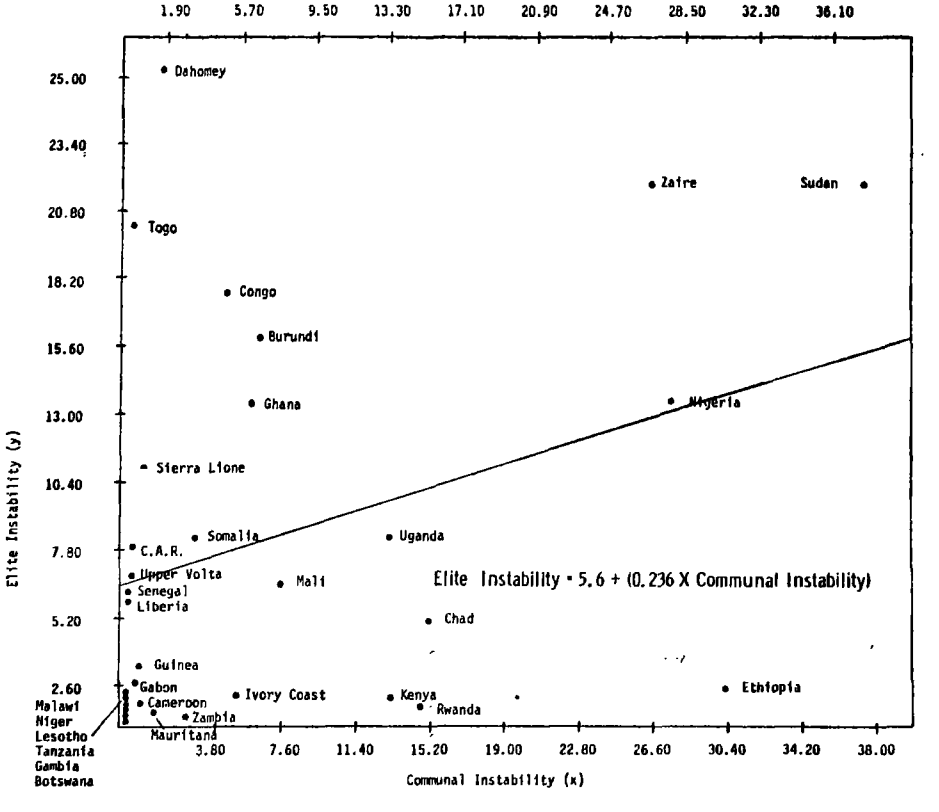
provided accurate measures of the phenomena to which they had been applied. A well-established theory or hypothesis might generate the expectation of a high positive correlation between two scales and calculation might produce one; but a range of positive values (to be on the safe side, let us say from "r" = 0.7 to just short of 1.0) could be regarded as high and considerable adjustments could be made to the scales which still allowed the correlation between them to remain within the limits of the range. A high positive correlation could thus obtain where the measures of either or both were inaccurate. If "validity" in this context did not entail "accuracy", the difficulty would disappear but, with it, the usefulness of the notion of "criterion validity" as explained by the authors. They do not, in any case, try to assess the criterion validity of their two scales although they state their intention to return to the topic before concluding their paper.

16 There is, however, a further, and very obvious, problem about the scales in Table 1. The scores for each country on each of the scales are reached by multiplying the "weights" for each event by the number of years in which it was reported for that country. These scores, however, need themselves to be weighted to take proper account of the length of time each country had been independent (if the Sudan became independent in 1956, Kenya did not follow suit until 1963) and this would entail dividing them by the number of years of independence

in each case. The result would show a less uniform distribution of "élite instability".

17 "By looking", say the authors, "at the two scales we have developed separately, we are engaged in univariate descriptive analysis". They suggest the construction of such scales every year, to produce 'fever-charts' of instability. But they then go on to draw an inference: "sharp upward shifts in the absolute and comparative levels of instability could provide <u>evidence</u> (emphasis added) for anticipatory action by the US Government" (ibid p 52). Kenya is taken as an example. "If one looked just at Kenya, one might issue unwarranted early warnings based on an absolute increase of communal instability in Kenya. But, when a comparative dimension is introduced, the 'dangerous' absolute increase of communal instability in Kenya, in the context of a continent-wide increase, might in fact lower Kenya's rank order. Alarmist speculation would then be more difficult". Here is an implicit assumption, that the countries covered by the scales are homogeneous. But no close observer of the African scene would so regard them. "Des cinq grandes régions de l'Afrique," writes Marianne Cornevin "l'Afrique septentrionale est la seule qui présente une réelle homogéneité" (1978 p 356). But, without homogeneity, no sure inferences can be drawn from statistical analysis based on the assumption of it. The notion that, on the law of averages, the FPM could some-

how downgrade an absolute increase of communal instability in Kenya, leading, or amounting, to an increase in political instability there, because of a general increase in communal instability in Africa, is misplaced; and a US Foreign Service officer in Nairobi or Washington would be misguided were he to downplay pointers to serious political instability in Kenya in the light of the situation elsewhere in tropical Africa. Coplin and O'Leary, indeed, argue, and provide evidence, against the assumption. They argue that previous incidence of coups or attempted coups in individual countries is a very good indicator of the subsequent occurrence of such events in those countries; but, if this were so, what might happen in other countries would not need to be taken into account - it would be irrelevant. The evidence is conveniently presented in the scattergram reproduced in Figure 1 (ibid p 54) of Elite and Communal Instabilities.



18 If tropical African countries really were homogeneous, random samples taken from the "population" of 32 countries should produce values converging on the slope-line (which represents the number of times by which a change in communal instability must be multiplied to find the equivalent change in élite instability). But three countries are "way out", Zaire, Sudan and Nigeria, and so are the countries with zero values on both scales. Exclude the former alone (or the latter as well) and it can be seen at a glance that the values for a significantly large number of random samples (say, with N countries = 10) could bear no relation at all to the slope: instead of a linear regression line, there is a curve running from Dahomey (as it was then called) in the top left corner of the scattergram to Ethiopia in the bottom right. From this it would follow that (at least in respect of the relationship between élite and communal instability) the 32 countries under consideration are far from being sufficiently alike to warrant conclusions being drawn from the sample averages and from the "least squares" of the deviations from them.[4)]

[4)] For lucid explanations of simple linear regression, of methods of calculation and of the use to which it can be put, see Dowdy & Wearden (1983 pp 201 - 210) and Floud (1973 pp 125 - 154).

19 The statement, moreover, that the regression line given in Figure 1 shows that communal instability "accounts for" a mere 11 percent of the variance in élite instability, even if based on a valid calculation of the average, could disguise the fact that, during the period considered by the authors, a number of coups in tropical Africa stemmed directly and discernibly from communal instability: in 1963, the coup in Togo was an act of revenge by the Kabre of the North against the Ewe in the South; in 1966, the reason the Hausa soldiery supported General Gowon's take-over of power in Nigeria was their resentment against their Ibo officers; in 1968, in the Congo, the rural Mbochi displaced the urbanised Lari (Cornevin 1978 pp 310 - 11). Yet, in the light of the crossed-lag correlations given in Figure 2 below, the authors feel entitled to say "it would appear that communal instability does not lead to subsequent élite instability, which is not what one might have expected" (1976 pp 55 - 56). Their conclusion "that policies designed to prevent élite instability need have no impact on communal instability, and vice versa" (ibid p 57) is stricto sensu true. There is no necessary connection between the two, but there can be, and has been, a contingent connection which it would be unwise for the Foreign Policy analyst to overlook.

FIGURE 2

Relationships Between Types of Political Instability in Africa: Cross-Lagged Correlations

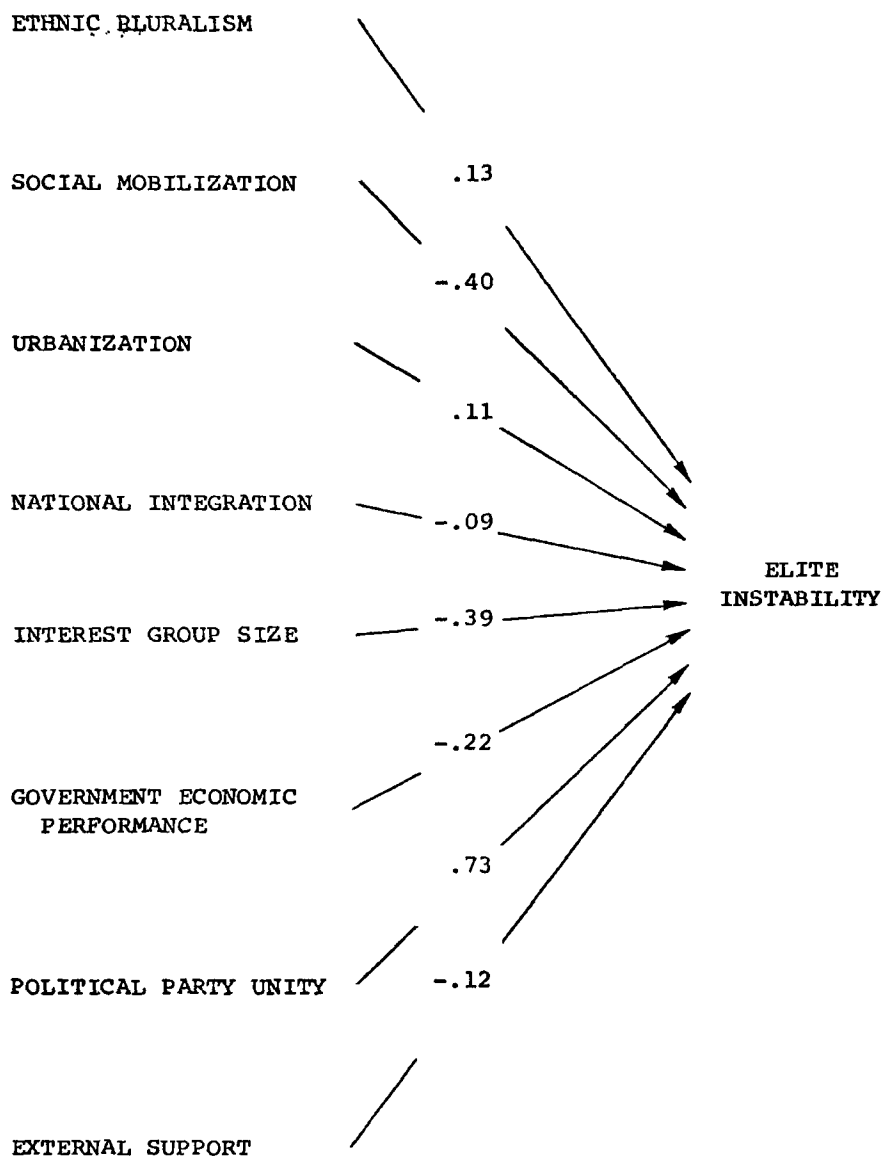
	1965 to 1969	
	Elite Instability	Communal Instability
1960 to 1964		
Elite instability	.97	.40
Communal instability	.16	.47

Source: D. G. Morrison and H. M. Stevenson, "Political Instability in Independent Black Africa: More Dimensions of Conflict Behavior Within Nations," Journal of Conflict Resolution 15 (September 1971): 365.

20 The authors "are now ready to deal with the critical question of whether explanatory theory, that is, strong theoretical relationships, can be identified ... Let us return ... to the contention ... that it is impossible to explain coups and civil wars in Africa except after the fact and with reference to each particular case. Are instances of élite and communal instability distributed randomly in Africa? Do they not have measurable and theoretically comprehensible correlates? ... Quantitative methods ... are appropriate for investigating these questions (and) an appropriate statistical method for assessing the truth content of alternative causal explanations of political and economic phenomena is regression analysis" (ibid p 57). The results of their multiple regression analysis of the factors "affecting" élite instability in tropical Africa are reproduced in Figure 3 (ibid p 61).

Figure 3

Causal Model of Elite Instability in Black Africa



The "indicators" of the factors are given in Table 2 (ibid p 72).

Table 2

Bivariate Correlations of Indicators of Factors affecting African Elite Instability

(N = 32)

Factor and Indicator		Elite Instability r
I.	Size Population, 1969	.20
II.	Ethnic pluralism Number of spoken languages, 1967	.35*
III.	Social mobilization Percent of workers in agriculture, 1968	-.49**
IV.	Urbanization Percent increase in percent of population in cities, 1955-65	-.38*
V.	National integration Number of commercial vehicles, 1966	-.20
VI.	Interest-group size Total armed forces men, 1967	.31
VII.	Government economic performance Cumulative balance of trade, 1963-68, as percent of 1967 GNP	-.34*
VIII.	Political party unity Number of illegal parties, 1957-69	.79**
IX.	External support Aid from ex-metropole per capita, 1969[b]	-.27

*Correlation significant at the .05 level.
**Correlation significant at the .01 level.

21 If the values given in Table 2 for correlation differ from those given in Figure 3 for regression, this is because the two operations differ in purpose and in computation: very roughly, the purpose of correlation is to determine the degree of linear association between variables, whereas the purpose of regression analysis is prediction (Dowdy and Wearden 1983 p 231). "In doing regression analysis" say the authors, "one wants to achieve two things, first to account for as much variance as possible in the phenomenon to be explained (this represents a crude index of the explanatory power of one's theory) and second, to examine the relative impact of each independent variable on the dependent variable (this permits assessments of which factors are most powerfully related to the dependent variable)"(ibid p 59). But a number of conditions has to obtain for the use of a regression analysis to be appropriate: for a start, the independent (better called the "predictor") variables must be measurable and measured without error.

22 It is not shown by the authors, and it is not immediately apparent, that this condition is met, that the "indicators" can be held accurately to measure the predictor variables. Examine Table 2, the factors and their indicators, <u>seriatim.</u> Leaving <u>Size</u> apart, accepting that it may have little bearing on élite instability, the first question to ask is whether the number of languages spoken in a country is an adequate measure of <u>Ethnic Pluralism.</u> There is an obvious

problem of classification involving the need to distinguish between a language and a dialect. In Cameroon, there are said to be several hundred distinct languages in the South - people who speak one cannot understand people who speak others - but the languages can still be grouped into families (as in Western Europe) (Baumann & Westermann 1957 pp 438 - 439); ethnic rivalries can exist between people who speak even distinct dialects as well as between people who speak different languages and different families of language (contemporary Zimbabwe provided an excellent example).

23 Similar criticisms can be levelled at the choice of indicators for the other chosen factors. Social mobilisation is vaguely characterised as "the emergence, because of social change and modernization, of masses of people who have rising expectations and who are making demands on the government" and "the percentage of workers in agriculture is an indication of its absence" (O'Leary and Coplin 1975 pp 51 - 60). This needs greater elaboration. Neither the absolute percentage of agricultural workers in the economically active population nor changes in that percentage need be in themselves pointers to discontent. In many African countries governments have adopted stick-and-carrot policies to induce people to stay on, or go back to, the land. In both Kenya and Tanzania, for example (to take one "capitalist" and one "socialist" country), such policies have been successful, largely, in preventing a mass rural exodus to the towns, but they have aroused much muttering

although, in the former, and until the country was hit by drought, the land reform programme, the shifting of subsistence farmers into the cash economy, was going well. Discontent, and hence instability, is linked not to the question of how many people are on the land but to the question of how many people are happy there.

24 A detailed analysis of the "gaps" between the remaining factors and indicators listed in Table 2 would be tedious but they should be briefly outlined. What makes urbanisation a factor for instability is not, in itself, the rate of increase in the percentage of a country's population living in the towns (which may stem partly from natural increase, given the very high birth-rates prevalent in most parts of tropical Africa, and partly from immigration) but the degree of poverty, resulting from un- or under-employment or inadequate wage-levels. As an indicator of national integration, the number of commercial vehicles in a country is clearly unsatisfactory. As an indicator of interest group size, it is not enough to look to the size of the armed forces; the latter are not invariably the only interest group in a country and may contain different interest groups within their own ranks. As an indicator of government economic performance, the ratio of balance of trade (payments?) to GNP matters less than the extent to which the government concerned is felt to be responsible for the state of an economy's performance which is frequently influenced by events recognised to be outside

national control. As an indicator of (the lack of?) political party unity, the number of illegal parties will not do: there may be serious rifts within the ruling party in a one-party state in which there are no illegal parties, in part because there are none. Finally, as an indicator of external support, the value of per capita aid from the ex-metropole is extremely crude, taking no account of the type of aid or of the effects of the ways in which it may be used.

25 The "gap" between the "indicators" and the "factors" indicated above could suggest to the FPM that some further refinement in measurement is required, perhaps a variety of indicators, in each case calling for each of the factors to be treated as a dependent variable itself, and itself subject to a possible multiple regression analysis. Such a procedure would, however, widen the "gap" to the extent that the pre-indicators of the indicators could not account for more than a certain percentage of their variance.

26 A further problem for the FPM over the authors' selection of "factors" will be evident. Admittedly, the selection was made on the basis of the views of the INR community and of interested scholars at the time. Yet it is permissible to ask whether discretion or ignorance was at play. A prime factor making for élite instability, or élite stability, over the period considered by the authors, for any particular country in tropical Africa, was the interest of

certain outside powers in preserving or upsetting régimes in power. Need more be said about the rôle, sometimes crucial, of external intelligence services, other than that the extent and effectiveness of their operations are difficult to quantify? What of the effects of climatic change or of (unacceptable degrees of) corruption? The FPM seems entitled to be somewhat sceptical over the authors' very precise claim that the factors they have selected are a 'powerful explanation' of élite instability in tropical Africa in that they "account for" 72 percent of the variance in that dependent variable (ibid p 67), and "that significant theoretical relationships among variables do exist".

27 The next step in the authors' argument is to attempt to show that the employment of an additional statistical technique, discriminant function analysis, combined with the results of their multiple regression analysis of élite instability, "could be used to construct a system for forecasting coups d'état and related events in tropical Africa" (ibid p 68). Discriminant function analysis, they point out, is widely used in the business world: to discriminate, for example, between good and bad credit risks in the light of statistical profiles based on past experience. They state the "technique to be reliable enough to be a <u>useful aid</u> to decision making (emphasis added) <u>and this is all a quantitative technique can claim to be</u> (emphasis in the original)" (ibid p 70). In the absence of an explicit theory (sc for forecasting

the incidence of coups in particular countries in tropical Africa), such analysis is based on the assumption "that the future will be like the past" (ibid p 69), a generalisation often used by INR analysts. Taking five of the indicators/ factors in Table 3 which they regard as being the most significant in "accounting for" élite instability (but note that the indicator for interest group size is no longer the number of men in the armed forces but the percentage of workers (among the economically active in the cash economy?) in the public sector), they produce a Table (Table 3 below (ibid p 71)) illustrating how states with at least one successful coup from the date of independence (or 1956, whichever was the later) differ in factor scores from states in which there were no successful coups through to the end of 1969. The first stage of the discriminant analysis is to show up such quantitative differences.

Table 3

Group and Total Average Scores on Discriminant Function Analysis Variables

Factor and Indicator	States with One or More Coups to 1969 (N = 14)	States with No Coups to 1969 (N = 18)	Average Value of All Cases (N = 32)
Social mobilization, percent workers in agriculture, 1968	80.4	87.0	84.1
Interest group size, percent workers in public sector, 1965	36.9	26.1	30.8
Government economic performance, cumulative balance of trade, 1963-68 as percent of 1967 GNP	-7.9	23.7	30.8
Political party unity, number of illegal parties, 1957-69	5.4	0.8	2.8
External support, aid from ex-metropole per capita, 1969, $ U.S.	3.5	7.5	5.8

28 Allowing, which is a large allowance, that the indicator/factor scores reflect reality, two questions arise for the FPM about the picture which the table reveals. The first arises from the comparatively insignificant deviation from the average, in respect of social mobilisation and interest group size, of states in the two categories, suggesting prima facie, on the basis of the analysis, that they can be virtually discounted for the purpose of discrimination. The second arises from what appears to be an arbitrary distinction between states that have had no successful coups up to the end of 1969 and those that have had them. For reasons given earlier, "coup-proneness" might legitimately be held to cover both attempted and successful coups. What should matter, statistically, is the frequency of attempted coups between the date of independence (or 1956) and the end of 1969, weighted for the length of independence. The occurrence of one coup, or attempted coup, in any particular country could indicate very little on the score of future "coup-proneness". A successful coup mounted with popular support, or acquiescence, against an unpopular régime could lead to élite (and political) stability for a long time to come.

29 The second stage of a discriminant function analysis, according to the authors, is "to establish a discriminant function and, on the basis of the weightings assigned by

this equation and the score of each African state on each variable (ie the indicators/factors enumerated above), to assign probabilities of group membership to each state". Table 4 below is the result (ibid p. 72).

Table 4

Probabilities of Group Membership of Coup-Prone and Non-Coup-Prone States

Classification	Coup-Prone Group p	Non-Coup-Prone Group p
I. States with one or more coups through 1969		
Nigeria	1.00	.00
Zaire	.99	.01
Sudan[a]	.99	.01
Dahomey[b]	.99	.01
Ghana[b]	.97	.03
Congo, Brazzaville[a]	.95	.05
Togo	.94	.06
Sierra Leone[a]	.93	.07
Upper Volta	.74	.26
Somalia	.68	.32
Uganda[b]	.56	.44
CAR	.43	.57
Burundi	.41	.59
Mali	.28	.72
II. States with no coups through 1969		
Cameroon	.62	.38
Ethiopia	.34	.66
Chad	.22	.78
Mauritania	.22	.78
Senegal	.21	.79
Gambia	.21	.79
Guinea	.17	.83
Tanzania	.11	.89
Kenya	.10	.90
Niger	.10	.90
Malawi	.09	.91
Lesotho[b]	.07	.93
Rwanda	.04	.96
Ivory Coast	.02	.98
Botswana	.01	.99
Liberia	.01	.99
Zambia	.00	1.00
Gabon	.00	1.00

[a]Countries that experienced a serious coup attempt between January 1, 1970, and December 31, 1972.

[b]Countries that experienced a successful coup between January 1, 1970, and December 31, 1972

30 It is a pity, in this respect as in others, that the authors do not detail their calculations (seeing that replicability, the checking of results, is held to be a hallmark of QIP method). "We will not discuss the discriminant functions created (sic) in our analysis. We will focus on the ... probabilities (resulting) and on what they have to say about the relevance of quantitative techniques to the problem of forecasting coups" (ibid p 71). But what they have to say is minimal: "statistical methods treat classes (emphasis in the original) of cases like coup-prone and non-coup-prone African states ... What will actually happen (emphasis in the original) in a given country cannot be predicted by any methodology available or likely to be available in the foreseeable future". Nevertheless, "individual cases can be fitted into general statistical patterns, and this could be of great help to the intelligent analyst" (ibid p 74). To a large extent, the first statement nullifies the second, even when the statistical analysis undertaken is unexceptionable.

31 Discriminant function analysis is intended to discriminate: either sheep or goats, no hybrids, no "grey areas". But in Table 4 above, there is a grey area, indicated by the boxed scores. There is thus no clear demarcation between the sets of states held to be coup-prone or otherwise. Moreover, it is important to note that the degree of probability of a state's membership of the set of coup-prone states is not the same as the degree of probability that a state classified as a member

of that set will experience a coup, at least within a specified period of time (that chosen by the authors being between January 1970 and March 1974). Nigeria is assigned a probability of 1.0 for membership of the set of coup-prone states, entailing the absolute certainty of such membership, but suffered no coup during that period. It was not until 1975 that General Gowon was displaced by General Murtallah Mohammed's coup and until 1976 that the latter was himself displaced by assassination. Cameroon, with a coup-prone probability, according to Table 4, of 0.62, remained stable for as long as President Ahidjo was in power and only recently has there been a reported coup attempt against his successor. At the other end of the scale, Zambia is assigned a 0 probability of belonging to the set of coup-prone states, meaning, strictly, that it was impossible for that country to be coup-prone; but there is no incompatibility in the logic of the Table between the impossibility of being coup-prone and the possibility of actually experiencing a coup. This must reduce its usefulness to the FPM.

32 In the final section of their chapter, the authors discuss the utility of the techniques they have employed to the Foreign Affairs analyst. In fact what they do is to change tack. Table 4 is forgotten. They look, instead, at the factors listed in Table 3 above as contributing to, or inhibiting, élite instability, and ask, in effect, whether changes in the measures of those factors can be of help in discerning a trend

towards a greater or lesser probability of a coup in any given country (though, this time, they are careful to stress the impossibility of predicting the date or the period of time at, or within, which the coup will take place). Assuming that the factors used are sufficiently relevant (a topic discussed earlier), this could be no more, as the authors emphasise, than a heuristic device (ibid pp 74 - 76).

33 The results of this case-study of a case study must be judged against the claims (quoted in paragraph 1 above) that "quantitative techniques can be used to generate predictions as well as identify those factors that the analysts should consider in making their predictions". What has emerged is that the predictions generated need to be checked by the individual analyst in the light of his knowledge of the individual countries of concern to him. In this study of tropical Africa, however, no prediction has been made which could better common sense, or intuition; and some quasi-predictions could have positively misled in the absence of informed judgment. The result is disappointing. But the authors are not alone in this. Across the field, there have been many "findings", but few of predictive value. Is there an explanation for this? One QIP scholar claims to have found one. This claim will be examined in section III.

III A Critique of a Critique: the "Colouring it Morgenthau" thesis of John A Vasquez

1 Before attempting to consider this question directly (and the word 'consider', rather than 'answer' is used advisedly), we must therefore examine John Vasquez' The Power of Power Politics: A Critique (1983), the main argument of which had been very fully summarised in an article which appeared in the British Journal of International Studies (1979) under the title "Colouring it Morgenthau: new evidence for an old thesis on quantitative international politics". In what follows the quotations will be taken from the article except where the book provides important supplementary material. The thesis falls into two parts: "The first is that international relations inquiry, as practised in the US, has developed as a science along the lines suggested by Kuhn (1970) and is currently in a normal science stage with all the major activities of the field (theory construction, data collection and research) being directed by the Realist paradigm. The second is that the failure of the behavioural research of the sixties to produce many strong findings constitutes an anomaly that challenges the belief that the Realist paradigm can adequately explain behaviour" (p 211). The focus will be on the second part of the thesis, although the first will be touched on. The reason is quite simply that, if justified, the claim would be of practical as well as theoretical importance. As Robert Rothstein has remarked (1972 pp 388 - 392), the doctrine of Political Realism seems to

exercise a powerful hold on the thinking of statesmen and diplomatists, at least of the traditional major powers: it represents them as custodians of an ancient wisdom; it appears to offer a picture of the international world as it really is and a recipe for success through the pursuit of a "national interest" defined in terms of power and security.

2 Vasquez takes Hans Morgenthau's Politics among Nations (1960) as the exemplar, or paradigm in the conventional sense, of Realist scholarship and there can be no doubt, historically, that that work has had considerable influence not only on academia but also on FPMs, particularly in the State Department in the decades immediately following the Second World War. On the basis of a textual and content analysis of the book, Vasquez 'operationalises' the fundamental assumptions of Realism as follows:

> "(a) Nation-states or their decision-makers are the most important actors for understanding international relations.
>
> (b) There is a sharp distinction between domestic politics and international politics.
>
> (c) International relations is the struggle for power and peace. Understanding how and why that struggle occurs and suggesting ways for regulating it is the purpose of the discipline. All research that is not at least indirectly related to this purpose is trivial".

It is worth comparing this list with the six principles of Political Realism as formulated by Morgenthau himself (ibid pp 4 - 15):

"1 Political realism believes that politics, like society in general, is governed by objective laws that have their root in human nature. In order to improve society it is first necessary to understand the laws by which society lives ...

2 The main signpost that helps political realism to find its way through the landscape of international politics is the concept of interest defined in terms of power ... It sets politics as an autonomous sphere of action and understanding apart from other spheres such as economics (understood as interest defined as wealth), ethics etc. ... Statesmen think and act in terms of interest defined as power.

3 The key concept of interest is not endowed with a meaning that is fixed once and for all ... The goals that might be pursued by nations in their foreign policy can run the whole gamut of objectives any nation has ever pursued or might possibly pursue ... The same applies to the concept of power.

4 Political realism is aware of the moral significance of political action ... It considers prudence - the weighing of the consequences of alternative political actions - to be the supreme virtue in politics. Ethics in the abstract judges action by its conformity with the moral law; political ethics judges action by its political consequences.

5 Political realism refuses to identify the moral aspirations of a particular nation with the moral laws that govern the universe.

6 The difference between political realism and other schools of thought is real and profound. The political realist is not unaware of the existence and relevance of standards of thought other than political ones ... but cannot but subordinate those other standards to political ones."

3 What emerges immediately is that Morgenthau's brand of Political Realism is a normative doctrine, not a descriptive theory. He advises (ibid p 5) that we should "put ourselves in the position of a statesman who must meet a certain problem of foreign policy under certain circumstances, and we (should) ask ourselves what the rational alternatives are from which a statesman may choose who must meet this problem under these

circumstances, presuming always that he acts in a rational manner, and which of these rational alternatives, this particular statesman, acting under these circumstances is likely to choose. It is the testing of this rational hypothesis against the actual facts and their consequences that gives meaning to the facts of international politics and makes a theory of politics possible" (emphasis added). This last sentence is obscure. Uppermost in Morgenthau's mind seems to be the thought that statesmen can be led astray by Idealism; for, a few pages later (pp 13 - 15), he gives three examples of actions in the field of foreign policy plainly dictated by "legalistic-moralistic" considerations rather than by the national interest as he conceived it. But he does not offer a general theory to account for actual adherence to, or deviation from, rationality. Indeed he does not offer a theory of behaviour at all. If, therefore, behaviouralist scholars said to be working within the Realist paradigm have failed to explain the actual behaviour of states, that failure cannot of itself be held to undermine the Morgenthau doctrine, unsatisfactory or inadequate though it may be for other reasons.

4 It must be said, however, that Vasquez has every right to try to extract from Morgenthau those assumptions which are of scientific value in guiding research, leaving aside the latter's normative principles. Moreover, it might be said that it does not matter very much if he fails to do full justice to Morgen-

thau in the process provided that he succeeds in capturing what most scholars would accept as the essential components of a Realist position. On the first point, it is worth noting that Morgenthau nowhere talks of a struggle for power and peace; for him there is a problem of peace because there is a struggle for power. On the second, a Realist, Robert Gilpin, has offered (1984 pp 290 - 1) a formulation which seems more satisfactory, his fundamental assumptions being that: international affairs are basically conflictual; that the group is the essence of social reality; and that the quest for power and security is the prime motive in all political life. This list does some justice to the jaundiced view of humankind which lies at the root of much Realist thinking; and its emphasis on the social group, rather than on the nation-state, accords with Morgenthau's view that the latter is not a sempiternal phenomenon and with the view, not basically at odds with a Realist philosophy, that domestic politics are not irrelevant to a state's external conduct. In other words, a critic could argue that, if Vasquez succeeds in undermining his version of a Realist paradigm, it does not follow that any alternative Realist paradigm would necessarily be discredited.

5 Accepting for the sake of argument that the Vasquez paradigm of Realism is 'scientifically' accurate, in the sense that it reflects the thrust of research undertaken by behaviouralist scholars in the field, we must next ask whether it is 'scientifically' adequate. For Vasquez, this means that it

must satisfy the following three criteria (1979 p 216):

> "- The Realist paradigm should tend to produce hypotheses that fail to be falsified.
> - The central propositions (i.e. those embodying the "core") of the Realist paradigm should tend to produce hypotheses that fail to be falsified.
> - Realist hypotheses that fail to be falsified should be of scientific importance."

On the same page he writes: "Whether a paradigm has produced knowledge can be determined by the empirical content of its theories, i.e. by the extent to which the theories give rise to hypotheses that fail to be falsified", explaining in a footnote that this criterion was suggested by Imre Lakatos (1970 p 116). This needs to be read in conjunction with Vasquez' own definition of "paradigm" (1983 p 5): "The concept of paradigm ... could be stipulatively defined as the fundamental assumptions scholars make about the world they are studying. These assumptions provide answers to the questions that must be addressed before theorizing even begins. For Kuhn ... such questions are: what are the fundamental units of which the world is composed? How do these units interact with each other? What interesting questions may be asked about these units? By responding to these questions, the fundamental assumptions form a picture of the world the scholar is studying and tell the scholar what is known about the world, what is unknown about it, how one should view the world if one wants to know the unknown, and finally what is worth knowing". He goes on to say that such assumptions focus attention on certain phenomena and govern the application of

certain concepts to them; propositions are developed by specifying relationships between concepts; and theories are developed by specifying relationships between propositions.

6 Although he invokes the names of two philosophers of science (each offering a rival normative theory about how scientists should go about their business), Vasquez is faithful to neither. In the case of Kuhn he concedes this; for in the Postscript (1970 p 175) to the second edition of The Structure of Scientific Revolutions Kuhn wrote: "... in much of the book the term 'paradigm' is used in two different ways. On the one hand, it stands for the entire constellation of beliefs, values, techniques, and so on shared by the members of a given community. On the other, it denotes one sort of element in that constellation, the concrete puzzle-solutions which, employed as models or examples, can replace explicit rules as a basis for the solution of the remaining puzzles of normal Science" (i.e. science conforming to the norms shared by the scientific community). Neither of these definitions suits Vasquez because he is anxious to differentiate the methodology of behaviouralist scholars in the field from their Realist assumptions and because, so far at least, their efforts have produced little in the way of exemplary puzzle-solutions (requiring, as they do, the backing of tried theory).

7 The all-important difference between Kuhn's concept of a paradigm and that of Vasquez is that the former provides a guide for a programme of research. It tells scientists not only what to study but how to go about it. Vasquez, on the

other hand, gives us a Realist paradigm which amounts to nothing more than an injunction to regard nation-states as the fundamental units in the world and (whether or not domestic politics are to be taken into account - a question left open by proposition (b) of the paradigm) to explain why they struggle for power and peace. It is only in this sense that his paradigm can be said to "guide" hypotheses (1979 p 213) or to "direct" data-making or hypothesis-testing (ibid p 214). That it fails to meet the specifications he himself lays down for a paradigm is of minor importance compared to the fact that it cannot be said to be capable of producing knowledge. Vasquez talks at one point (ibid p 217) of "the extent to which Realist hypotheses ... adequately test the theoretical premises of the paradigm" and of the latter's ability "to accurately predict behaviour", but none of the hypotheses tested is deducible from it. It predicts nothing. It is not a theory; rather, it is a set of assumptions and, while it may be possible to disagree with what the three propositions could be taken to imply (if, for example, the first was taken to mean that nation-states and nation-states alone were worth study), who would be prepared to deny that any body of theory about international relations worth the name would have to try to take account of the behaviour of nation-states and of the problem of preserving peace in the face of their unending quest for power and/or security? If Vasquez' argument were correct, the scholars whose findings were summarised by Singer and Jones (1972

Section II B pp 155 - 209 and Section III B pp 325 - 395) had been addressing the wrong problems. It is difficult to take such a claim seriously.

8 Let us assume, however, and, again, for argument's sake, that Vasquez' Realist paradigm could be interpreted in a slightly different way, as equivalent to the proposition: that if behaviouralist scholars study the behaviour of nation-states and the manner in which they interact in the struggle for power and peace (even ignoring domestic politics in the process), it is reasonable to predict, given the correctness of their methodology (or what Vasquez calls the 'adequacy' of quantitative analysis (1979 p 226)), that their labours will produce a fair number of explanatory hypotheses which do not "fail to be falsified". But since, the argument would continue, over 90 % of over 7.000 Realist hypotheses examined have been falsified (ibid p 223), that proposition itself must be regarded as falsified and the conclusion must be that the focus of their enquiries has been misdirected. This raises the question as to what Vasquez understands by "fail to be falsified".

9 As noted earlier, he says that the phrase was 'suggested' as a criterion by Lakatos but, while it may be possible to read such a suggestion into what Lakatos wrote on the page in question (1970 p 116), it is worth looking at his <u>ipsissima verba:</u> "For the naive falsificationist a theory is falsified

by a(n) ... 'observational' statement which conflicts with it (or which he decides to interpret as conflicting with it). For the sophisticated falsificationist a scientific theory T is falsified if and only if another theory T 1 has been proposed with the following characteristics: (1) T 1 has excess empirical content over T: that is, it predicts novel facts, that is, facts improbable in the light of, or even forbidden by, T; (2) T 1 explains the previous success of T, that is, all the unrefuted content of T is included (within the limits of observational error) in the content of T 1; and (3) some of the excess content of T 1 is corroborated". Lakatos says later (ibid p 119): "Contrary to naive falsificationism, no experiment, experimental report, observation statement or well-corroborated low-level falsifying hypothesis alone can lead to falsification. There is no falsification before the emergence of a better theory". Whatever the merits of Lakatos' position (not surprisingly a matter of some dispute among philosophers of science), it is clear that on his criterion of falsifiability, unduly generous though it may be, the theories that may underlie the Realist hypotheses surveyed by Vasquez "fail to be falsified"; for there is as yet no better theory or set of theories, that is, no theory answering to the description of Lakatos' T 1, in the field. Certainly, none of the theories from which Vasquez' 'non-Realist' hypotheses may be derived would appear to qualify. These hypotheses number about 550, or about 7 % of the sample, and, according to him, show a 'failure' rate of a mere 83 % (1979 p 222).

It is permissible to wonder whether, if Vasquez had focused on these and these alone, he would have concluded that there was some 'non-Realist' paradigm worth pursuing.

10 In fact, Vasquez offers his own criterion (1979 pp 217 - 8): "A hypothesis can be said to have failed to be falsified to the extent to which, when tested by inductive and descriptive statistics, it is found to be statistically significant and to have a strong measure of association". He specifies that a hypothesis fails to be falsified if it is significant at the .05 level and if the measure of association is at least .72, with both requirements having to be met. Such a hypothesis is said to have 'predictive power'. But it is important to note that falsifiability is not here the issue; nor does Vasquez claim that it is. A significance test is a procedure for deciding between two statistical hypotheses: one, typically but not necessarily, positing no relationship between two variables (the so-called "null hypothesis"); the other positing some relationship. The question asked is: what, given the truth of the null hypothesis, would be the probability of obtaining by chance (say, sample error) the measured result actually obtained? If the researcher decides in advance that he can afford to take the risk of a 1 in 20 chance of rejecting a null hypothesis which is in fact true, he will set the significance level at .05. There is nothing sacred, however, about an .05 level; and whatever level is set necessarily preserves some possibility of the

truth of the rejected hypothesis (a consideration advanced by Gillies (1973 pp 187 - 194) against his own proposal for a 'falsifying rule', incorporating significance tests, for statistical hypotheses). Equally, a significance test cannot guarantee the truth of a hypothesis accepted as a result of it. Moreover, we need to bear in mind with Morrison and Henkel (1970 p 185) that "the level of significance of a given sample statistic under the typical null hypothesis is a monotonic increasing function of the sample size. A correlation coefficient of .11 can be the basis for rejecting a typical null hypothesis at the .05 level of significance in one study, but in another study of the same population a correlation coefficient of .25 will not reject the null hypothesis at the same level of significance, if the sample sizes are respectively large enough (300) and small enough (50)".

11 What matters for Vasquez, however, is 'predictive power': "A hypothesis is as useful as the percentage of variance it explains" (1983 p 178); and "Measures of association (which have passed a particular significance test) provide an estimate of how many successful guesses a scholar can expect to make by using a particular hypothesis as a guide to prediction" (ibid p 179). The implications of the first of these statements are discussed below but it must be said at once that the second should be treated with caution. Vasquez correctly claims (ibid p 177) to be following Blalock (1960) in stressing the need to take account, in assessing the usefulness of a statistical hypothesis, not only of the result of a significance

test but also of the strength of the correlation coefficient. But the two taken together are not by themselves sufficient to justify prediction. A correlation coefficient indicates the strength of a relationship, not its nature. By itself, on the assumption it has passed a significance test, it has no explanatory power and need not have predictive power. According to Blalock, that possible privilege must be reserved for the "regression equation", such equations being "the 'laws' of a science" (ibid p 275); his point is, rather that a regression of one variable upon another is unlikely to be of predictive value unless their correlation coefficient is "reasonably high (say, .7 or above)" (ibid p 299).

12 Nevertheless, and contrary to the first of the two statements of Vasquez quoted in the preceding paragraph, this does not mean that low coefficient values necessarily lack theoretical importance. In their critique of David Singer's Correlates of War project (Hoole & Zinnes 1976 pp 55 - 57), Job and Ostrom point out (and it is not a point which Singer controverts in his reply (ibid pp 127 - 134)) that "by focusing on highly significant (i.e. both statistically significant and high-value) relationships, the COW researchers have no guarantee of pin-pointing relevance and in fact may tend to miss much of the information presented in their correlation matrices. ... a very low relationship between international violence and a certain variate may actually be very 'important' from a theoretical perspective and worthy of further investi-

gation". Singer and his associates had been engaged in a "brush-clearing" operation, selecting high-value bivariate relationships for inclusion in a multivariate analysis, but such analyses will frequently have to take account of a fairly wide range of contributory factors, many of which individually may have a low-value correlation with the dependent variable, if they are to "explain" a satisfactory percentage of variance in the latter. Mono-causal explanation enjoys no monopoly. It follows that Vasquez was not entitled to dismiss correlation coefficients of less than .72 as not being useful (let alone as "falsified") and hence that we can have no confidence in his assigned failure rate for Realist hypotheses.

13 If it is important to distinguish between hypotheses which are simply being 'explored' and those which are, or may have been, specifically designed to put a theory to the test, it is just as important not to write off a scientific research programme which gives rise in its early stages to hypothesised relationships which fail to "fit the facts", particularly when their failure leads to the formulation of better theory. A pertinent example is afforded by Vasquez' treatment of Rudolph Rummel's article "The Relationship between National Attributes and Foreign Conflict Behavior" (1968 pp 187 - 214). The hypotheses there 'tested' numbered over 2.500, a very substantial proportion of all the hypotheses included in the survey, and Rummel conceded (ibid p 213) that the findings had been "almost entirely negative for the nations and time periods covered". Accordingly, to dispose of the

ad hoc explanation that this one article had contributed, unduly and artificially, to the high failure rate of Realist hypotheses over all, Vasquez "repeated (his analysis) without the article, and the results were substantially the same" (1979 p 226 footnote). But the picture looks a bit different in the light of subsequent developments. In the article itself, Rummel explained why his findings had been so poor: in effect stemming from his failure to connect his 'attribute system' to his 'behavioral system' "by the distances between nations on the attribute dimensions" (1968 p 214). He later modified this position, too, and came up with a 'model' "with an average multivariate correlation of .70 or 49 % of the variance" (1976 p 237), that is, within a whisker of satisfying Vasquez' .72 criterion.

14 It would be wrong to suggest that Vasquez overlooked these later findings of Rummel. Indeed, he refers to them specifically in the chapter of his book reviewing theory and research in the 1970s (1983 pp 204 - 227) which was clearly written after his article was published and not foreshadowed in it. "Rummel", he claimed (ibid p 222), "has produced two pieces of evidence (1972) that undercut the third assumption of the Realist paradigm (cp § 2 above). The first is that status-field theory can produce very high correlations, indeed some of the strongest published in the 1970s, when used to predict foreign policy behavior. The second is that as long as foreign policy behavior is treated as a unidimensional

struggle characterised by conflict-cooperation, efforts to predict behavior will not be very successful." In other words, Rummel's research programme had left the Realist track. But the claim is curious for two reasons. In the first place, 'status', however defined, has always been regarded, by classical Realists at least, as an essential component of power. In the second, while it certainly seems mistaken to treat conflict-cooperation as a single continuum, no commitment to such a view is implied by the assumption that nation states "struggle for power and peace", let alone entailed by a Realist outlook. It follows that since, as Vasquez declares in the article (1979 p 222), "inter-nation conflict-cooperation is the most frequently employed dependent variable (sc in the hypotheses which figure in the Jones and Singer compilation (1972)), constituting about sixty percent of all the employed dependent variables", part of the reason for the failure of those hypotheses that failed could be ascribed to the 'single continuum' presumption rather than to Realism.

15 Vasquez found also (ibid) that "national power is the most frequently employed independent variable, (likewise) constituting almost sixty percent of all the employed independent variables ... (and that) the most frequently (sic) hypothesis in the field is the one that employs national power to explain inter-nation conflict-cooperation (41.7 % of the hypotheses)". Now, although "national power" is undeniably

a Realist concept, there is, again, a reason why the failure of hypotheses structured in this way to be accepted as satisfactory ought not to be laid at Realism's door, the consideration advanced in the preceding paragraph apart. It is a truism that, generally, nation states will seek, in the light of their "power", to enhance or maintain that power relative to that of other nation states, of a group of them or of one rival in particular, but this does not mean that "total power" as quantified by measures of, say, GNP, defence expenditure, land-mass etc., is necessarily proportionate to the "amount" of power a particular state can or will bring to bear in a particular situation. Power is not fungible; and a distinction must be made between "total" power, available power and effectively usable power (this last being extremely difficult to measure for particular circumstances, even approximately). Not all the elements of power are imponderable, but some are; and of those that are ponderable (in some sense), not all are ponderable in the same scales (what is the relative weight of good generalship to physical fire-power?). "Can" does not imply "will", moreover, and not simply because misjudgements are possible on the part of decision makers about the relative capabilities of their own and other nation states. History is replete with examples of lesser 'powers' thwarting greater, over major as well as minor issues (the Vietnam War is a case in point). It is scarcely surprising that "power" should be an unreliable predictor variable, in any calculus Realist or otherwise.

16 Vasquez' final comment on the relative value of the Realist and non-Realist hypotheses he had tested concerns their relative "scientific importance". Of the former, a mere 157 had "failed to be falsified" and, of these, very nearly 70 % had been judged "trivial"; the corresponding figures for the latter were 24 and nearly 55 % respectively. Hypotheses were deemed trivial "if they were (a) similar to common sense generalisations ... (e.g. nations tend to enter alliances before they fight wars); (b) correlated measures of the same concept (e.g. economic capability with quality of consumer goods; or (c) were highly idiographic and therefore of little importance for building general theory (e.g. population size and number of airline flights" (ibid p 225). Vasquez grants that the results of this test should be interpreted with caution in view of the small number of cases under consideration and of the subjective nature of the concept of 'triviality' but, in the matter of criterion (a), he seems over-severe. Of course the scientist is interested in "novel facts"; and a body of theory is often held to be strikingly confirmed when it predicts one (a classic example being the discovery of the Planet Neptune which accounted for the otherwise anomalous behaviour of Uranus). Yet a scientific research programme must, so to speak, learn to walk before it can run. Its theories must be able to account for known facts and, in the field of international relations (as indeed elsewhere), they can take the form of very obvious generalisations. Karl Popper has provided an example (1952 p 264): "If we explain ...

the first division of Poland in 1772 by pointing out that it could not possibly resist the combined power of Russia, Prussia and Austria, then we are tacitly using some trivial universal law such as 'If of two armies which are about equally well armed and led, one has a tremendous superiority in men, then the other never wins'". Regardless of whether such a proposition can properly be characterised as a "universal law", the point is made: just as statistical significance should not be equated with theoretical or scientific significance, so the latter should not be equated with practical significance. A scientific corroboration of a seemingly trite generalisation may be uninteresting for the scholar but of value to the policy maker.

17 The comparison between Realist and non-Realist hypotheses is a crucial component of Vasquez' argument because "the decision-rule that proportionally more Realist than non-Realist hypotheses should fail to be falsified or be of scientific importance ... permits an empirical determination of the adequacy of the Realist paradigm" (1983 p 176), albeit a tentative one (given the possibility of ad hoc explanations for the disappointing results in the field). Consequently, it is of some moment that successful non-Realist hypotheses should be properly categorised as such. Yet of the 11 hypotheses so categorised by Vasquez (ibid p 198), 4 would seem to be candidates for the Realist group. Three

taken from OR Holsti's "Cognitive Dynamics and images of the Enemy: Dulles and Russia" (abstracted in Jones & Singer (1972 pp 174 - 5)), are "somewhat successful in predicting the conditions under which Dulles would perceive Soviet policy as hostile"; but was not Dulles the key US decision maker in foreign policy at the time? And the fourth, taken from HR Alker's "Dimensions of Conflict in the General Assembly" (abstracted in Jones & Singer (ibid pp 278 - 80)) predicting UN votes on self-determination from membership of the "Old European" group, would seem to fall into the same Realist slot as his finding in the same article that various indicators of East-West alignment were good predictors of voting on East-West issues. The result, however, of transferring just these four 'non-trivial' hypotheses from the non-Realist to the Realist camp is to reduce the former's success rate from 45.8 % to 35 % and to raise the latter's from 30.5 % to 32.3 %, hardly a reasonable margin for the application of the decision-rule. Of the seven remaining successful non-Realist hypotheses, moreover, six, taken from J Laulicht's "An Analysis of Canadian Foreign Policy Attitudes" (abstracted in Jones & Singer (1972 pp 180 - 181)), afforded good predictions of attitudes towards coexistence/disarmament and internationalism on the part of business, labour, political élites and the voting public; while the seventh, taken from Bruce Russett's "International Communication and Legislative Behavior: the Senate and the House of Commons" (abstracted ibid pp 194 - 5), is a .86 correlation between the degree

of involvement of a US state in Anglo-American trade and the degree of responsiveness of its congressmen to the United Kingdom on foreign policy questions. The question arises as to whether it makes sense to see in the success of these hypotheses any threat at all to the adequacy of Realism, whether, indeed, they can fairly be regarded as incompatible with it. But, if the answer to the second question is 'no', the success rate of non-Realism is zero.

18 "It is already clear", writes Michael Banks, "that Vasquez' The Power of Power Politics is the most important single work to have emerged from the behavioural movement in international relations ... The findings strike at Realism like an Exocet missile" (1985 pp 220 - 25). Vasquez' own claims are more restrained. He allows that "the youthfulness of the field" and "measurement error" may account, or may still be defended as ad hoc explanations, for the "absence of much produced knowledge" and accepts that it would be a mistake to abandon large data projects "producing new Realist indicators" already under weigh. At the same time, he insists that "until a paradigm is found that shows promise of adequately explaining behavior, there will be no major progress in research. This implies that the Realist paradigm must be rejected as the dominant paradigm in the field ... no new projects guided by (it) should be permitted to occupy a large amount of the (available) intellectual energy and

financial resources" (1983 pp 226 - 27).[1] But, if the considerations advanced above have merit, the conclusion that follows is that Vasquez has made Realism the scapegoat for the alleged failings of behaviouralist research in international relations generally. If his missiles have been directed against Realism, they have struck the whole QIP field.

19 How much damage has been inflicted? Those prejudiced against the quantitative approach to international relations may be inclined to hail Vasquez' results with glee. Nevertheless, as has been remarked by Blalock, an acknowledged high priest of multivariate analysis, "there are (sic) ... a number of complex, technical and sticky methodological problems confronting all the social sciences ... and one important ... consequence is that we seldom, if ever, experience any dramatic breakthroughs in (them)" (1984 p 38 and p 43 respectively). The book from which these quotations are taken, Basic Dilemmas in the Social Sciences gives an

---

[1] Consonant with these remarks, Vasquez has himself produced, in collaboration with Richard Mansbach, an alternative paradigm to Realism: In Search of Theory: A New Paradigm for Global Politics (1981), on which a critical note is appended. It is difficult to envisage how the authors could "operationalise" their paradigm, let alone do so without the use of at least some 'Realist indicators'.

admirable non-technical account of the daunting difficulties facing any quantitative analyst attempting to tackle "the complications produced by multiple causation" and, in a real world full of ambiguities, to "move from many facts to fewer propositions" (to reproduce language from some of the chapter headings). It is important to note both that these problems of methodology, particularly in the field of 'measurement', may well help to explain the slow growth of "the Scientific Study of International Politics" and that those engaged in such study are not unmindful of them. They are extensively discussed for example in Hoole & Zinnes' Quantitative International Politics: An Appraisal (1976), a work which does not even figure in Vasquez' bibliography. Yet the latter has posed an important question (even if he was not the first to do so). There is now an abundance, perhaps over-abundance, of data, of "operationalised" concepts and statistical correlations, available for study, but not much 'explanatory' or 'causal' knowledge. Where are the theories? What kind or kinds of theory are to be looked for?

A P P E N D I X

In Search of Theory: A New Paradigm for Global Politics (1981) by Richard Mansbach and John Vasquez (referred to hereafter as MV) is specifically presented by the authors as an alternative to "Realism". The paradigm itself embodies three fundamental propositions, each, as MV see it, correcting a basic Realist tenet: (1) There are actors on the world stage other than nation-states; (2) There is no essential difference in principle between domestic and international politics; and (3) the values which guide international actors are not confined to power and/or security maximisation.

MV's system has been described by Peter Willetts (1984 p 96) as an "Issue-", as distinct from a "Power-", "paradigm". The Issue, what is at issue, is the "allocation" or "disposal" of stakes, by and between international actors. A stake can be concrete, valued for itself (e.g. a chunk of territory); symbolic, i.e. valued not only for itself but also because it represents other values; or transcendental, i.e. valued above other values. Issues become "salient" when they dominate the attention of major international actors and those with the highest salience are the "critical" issues on the world agenda. Critical issues involving (not necessarily always) transcendent stakes, tend to "absorb" lesser issues involving concrete and symbolic stakes. Thus US intervention in Vietnam was not just to protect the South (concrete), but to forestall a "domino-effect", the fall of neighbouring coun-

tries to communism (symbolic), and, beyond that, to maintain American credibility as defender of the "Free World" (transcendental).

An important corollary of the paradigm proper is that relations between international actors cannot be scaled on a single continuum ranging from cooperation to conflict. The actors can be friends or foes (depending on the previous pattern of interaction); but foes can reach agreements with each other and friends can take negative actions against each other, without necessarily modifying friendship or hostility. This creates both problems and opportunities. Actions that are perceived as "out of character" give rise to "cognitive dissonance", leading to possible misinterpretation of motives, but the very awareness of this danger can be useful in formulating policy.

If there is an implicit prescriptive element in this approach, it is that international actors require "sensitivity" rather than the rationality or prudence of the Realist. This means, however, that the paradigm lacks a dynamic. There are no "forces", psychological or otherwise, built into it; no criteria for deciding which values, or mix of values, should prevail; no calculus for distinguishing the short, medium and longer term. In short, the paradigm does not help us to "explain" international behaviour although it helps us to interpret it.

MV do not seem to realise this. Somewhat bravely for theorists, they actually provide an account of the origins of the original cold war and of subsequent developments in the relationship between the two super-powers up to and including the Soviet intervention in Afghanistan at the end of 1979 (1981 pp 388 ff). Their purpose, they say, is to enable the reader to apprehend why the cold war occurred and to grasp the causal dynamics that might rekindle it (ibid p 388), but they appear unaware that, even before their book went to press, clouds were already gathering. They have a tendency both to misread the historical record and to "reify" issues, giving the impression at times that the latter can jostle for position on the world agenda.

The failure of MV to see that Détente was waning in the late 70s stems in part from their failure to understand what it was all about. They date its origin (not unreasonably, though other dates might have been chosen) to the aftermath of the Cuban missile crisis of 1962. That crisis had prompted the realisation on the part of the two super-powers that the critical issue was no longer the struggle between communism and capitalism (socialism and imperialism), on which they could not be expected to agree, but the avoidance of nuclear catastrophe, on which they could. The rules of the game (what MV call the "allocation mechanism" for the disposal of disputed stakes) therefore needed to be changed. Coercion, one such mechanism, had to make way for agreement to bargain and to operate within the framework of agreed

principles. While communism/capitalism remained a salient issue, the fact that neither Nixon nor Brezhnev could be regarded domestically as being 'soft' on it helped to make it possible for both sides to explore ways and means of putting the relationship on a new footing. This resulted in the Agreement on Basic Principles of 1972 and SALT I of the same year.

So far, fair enough. But consider the following passages (ibid from pp 443 - 5): "The most salient issues ... were control of the arms race ..., the maintenance of nuclear parity at bearable cost, and the old division-of-the-spoils issues in Europe"; "Détente coincided with a reduction in the potency of the actor dimension (i.e. the degree to which each side was influenced in its approach to issues by its attitude (friendly, trusting, or otherwise) towards the other) and movement towards a stake dimension, as a result of which both actors began to view the key issues as separately negotiable"; "The Helsinki agreement was effectively a peace treaty ending World War II ... the transformation of a de facto into a de jure situation". But it may be doubted whether the maintenance of nuclear parity was in fact an objective of either side, whether the actor dimension had lost its potency and whether the division-of-spoils issue in Europe was still salient. The Helsinki Final Act was, besides, strictly de facto in the sense of being morally and not legally binding. Most important of all, perhaps, while certain key issues might form

the subject of separate negotiations, for Kissinger all issues were linked: the Russians were expected to play ball across all the boards.

Yet the point that needs to be made is not so much that MV's history of recent relations between the Soviet Union and the United States is open to question, still less that their system has encouraged them to misread the "facts"; it is rather that their marriage of the "issue-area" thinking of Rosenau and O'Leary & Coplin with "cognitive dissonance" thinking to produce a "theory" of "cognitive interdependence" gives any theorist, regardless of paradigm, another useful conceptual tool. It is an essential part of Chomsky's critique, for example, in Towards a New Cold War (1982) that the anti-communism issue has deliberately been exploited in US domestic politics, kept "salient", for reasons unrelated directly to the nature and extent of the "Soviet menace".

As a description of events, the MV theory is far from comprehensive; and this is partly by intention. MV do not purport to pass judgment on what, for many, would be the crucial question of responsibility: who was primarily to blame for the ups and downs in US/Soviet relations? "This important question ... is, on the whole, impossible to answer objectively, given limited access to classified documents" in the countries concerned (ibid p 451). Nor is the theory meant to be prescriptive: the authors strive to keep it

"value-free" while expressing a personal preference for hastening the demise of a declining cold war through mutual abandonment by Washington and Moscow of any desire to take advantage of each other's difficulties or to overthrow the other's system (ibid p 461). MV have useful insights to offer on the mechanics, in the abstract, of the growth and waning of trust and mistrust. It is left to the policy maker or to the concerned citizen to devise inputs for their conceptual machine.

IV Theory and Change

1 The works examined in sections II and III of this essay were selected because both dealt with topics of obvious interest to a FPM. If there were a statistical method of predicting coups demonstrably more reliable, or less fallible, than the judgment of the man in the corridor or "in the field", decision makers and their advisers would be quick to take notice. So too with any well-grounded claim that Political Realism had been scientifically discredited as a paradigm for research. Both pieces, however, have also served to illustrate the tendency, not uncommon in the QIP community, to be admirably explicit about method while not being as explicit about the underlying rationale. "What distinguishes the QIP field", writes Dina Zinnes, "is its approach (of which) the major characteristic is the attempt to be as explicit as possible about all aspects of the presentation of an argument. It could be said that the central goal of the QIP researcher is to put his entire argument in the public domain so that others may judge, accept or reject, and build upon the research" (Hoole & Zinnes 1976 p 5) (emphasis added). The book from which this quotation is taken was the fruit of a praiseworthy effort to focus attention not only on the "methodology", but also on the "philosophy of science", employed in QIP research. Yet, in the latter context, there is a wide divergence of opinion among the contributors about what constitutes a proper theory.

2 Caporaso (1976 pp 355 ff) defines "theory" in a manner closely resembling the orthodox positivist views of Carl Hempel (the so-called Deductive Nomological model outlined in paragraph 7 below). David Singer (1976 a pp 24 - 26), believing that "the notion of 'causality' may be inappropriate in the social sciences", declares that the name of the game is prediction: the model that best integrates the (statistical) predictors is said to "explain" the outcome; but what makes an explanation satisfactory is "a highly subjective matter". Later (1976 b p 133), he writes: "A theory ... is a set of logically interdependent hypotheses (and assumptions), most of which have been empirically confirmed, brought together for the explicit purpose of illuminating causal patterns". North, Holsti and Choucri leave the question open, while conceding it is fundamental: "the use of a particular statistical technique is <u>in itself</u> (emphasis added) a theory about the structure of data, and there are some implicit assumptions about the impact of the findings of alternative data structures. Indeed, the data themselves reflect an implicit or underlying theory that has guided their initial compilation. The theory of data is an area that yet remains insufficiently investigated" (1976 p 452). The second sentence is clear enough; the third obscure; and the first comes perilously near to substituting method for theory (a form of Hilary Putnam's "method fetishism" (1981 pp 188 - 200).

3 Quot homines, tot sententiae (or almost). A few examples, taken from the writings of scholars who did not contribute to the Hoole & Zinnes volume, will suffice. McGowan, following Van Dyke, states that a "theory is an explicit formulation of determinate relations between a set of variables in terms of which a fairly extensive class of empirically ascertainable regularities (or laws) can be explained" (1975 p 65) ... "If we have generated a theory of the causes of X, we can predict its future occurrence; X would take place when the necessary and sufficient conditions reoccur. This is prediction based on explanatory theory". Earlier on the same page, he speaks of the need "to engage in deductive theorizing in an effort to relate empirical statistical laws to general causal forces, such as the need for national security ...". Wilkenfeld et al., (1980 p 241) regard a scientific theory as a "system of statements. Some of the statements are a priori (the assumptions, premises or axioms); others are logically implied from these and represent theorems which take the form of falsifiable statements about the real world. Theory, then, is a deductively connected system of axioms and theorems". But they go on to draw a distinction between the basic underlying preconditions of foreign behaviour and those factors or events which actually precipitate behaviour, the former being amenable to systematic enquiry, the latter not (ibid p 244). This implies

that theorising can cover necessary, but not (necessarily) sufficient, conditions.

4 Bueno de Mesquita can provide the last example. He has not only given an account of his concept of theory but also produced a major work The War Trap (1981) expounding one on the initiation of conflict. For him, a theory is an explanation of relations among variables. Referring to "the well-known statement that theories can never be proved true", he adds that "they must, nevertheless, either be logically true or logically false ... The truth of a theory resides in whether or not its conclusions (that is, theorems) can be arrived at without faulty logic. If a deduction follows logically from a set of assumptions then that deduction is necessarily true under the precise conditions assumed in the theory. The truthfulness of the deduced relationship among variables, given the assumptions of the theory, is not an empirical question. Consequently, the deduction need not be subject to empirical evaluation to determine its truthfulness. But when one studies data that approximate, but do not comply precisely, with the assumptions of one's theories, then empiricism becomes crucial. For it is only by observation that one can determine whether a theory's generalizations are useful explanations of reality ... It is as inappropriate to speak of assumptions being true or false as it is to speak about a theory being empirically (as opposed to logically) true or false" (1980 p 363).

5 The purpose of the theory he elaborates in the book is to "provide lawlike general statements about the initiation ... of international conflicts" (1981 p 181) and Bueno de Mesquita stresses that his generalisation must be regarded as universal (i.e. not subject to spatio-temporal restrictions): "If the logic is temporally bounded, then the theory cannot provide a vehicle for dealing with future events" (ibid p 133). The generalisation that emerges from the theory is "the expectation (sic) that wars ... will be initiated only when the initiator believes the war will yield positive expected utility" (ibid p 27) (where, following von Neumann & Morgenstern (1944), 'expected utility', as a numerical measure, is the product of a numerical value assigned to a desired outcome and of the probability, again expressed numerically, of obtaining it). The assumption is made that the 'initiator' is a rational actor. Yet, on Bueno de Mesquita's own showing (ibid pp 129 - 134), some of the wars he considers (between 1816 and 1974) were started in the absence of such utility for the initiator: the supposedly necessary condition was not there. He is not bothered by this because his concept of theory precludes the empirical testing of the consequences of axioms and assumptions and merely requires the latter to 'approximate' reality; but the result is to rule out his generalisation as 'lawlike'. In reply to his critics (Majeski & Sylvan 1984 and Yuen Foong Khong 1984), he takes refuge in the Lakatos formula on falsification

(cp section III paragraph 9), claiming, in effect, that his theory should stand, since there is no better alternative in the field (1984 pp 342 - 344). He does not pause to consider whether his theory is a theory in Lakatos' sense, let alone whether Lakatos' position can be accepted without question (cp Newton-Smith 1981 chapter IV and Keat & Urry 1982 pp 49 - 50).

6 In their concluding comments, Hoole and Zinnes discuss the criteria for evaluating the quality of theories (e.g., elegance, parsimony, 'goodness of (empirical) fit', etc.) (1976 pp 463 - 67), but are silent on the question of what makes a 'theory' a theory. This is somewhat paradoxical. Given the different conceptions expressed, it might have been expected to merit examination: the QIP community is, after all, in business to produce theory and scientific theory at that; the 'house epistemology' of the behaviouralist social scientist (to borrow a phrase from Michael Shapiro (1981)) entails a commitment to "methodological monism", the doctrine that the natural and human sciences form a unified corpus of knowledge brought, or to be brought, to light through the application of scientific method(s); and, for 'positivist' philosophers (such as Braithwaite (1953), Nagel (1961) and Hempel (1965)), the so-called Deductive-Nomological model (or DN model for short) represents the ideal to which scientific theories should, if possible, conform (failing which, the scientist must fall back on what Hempel calls "Inductive-Statistical" explanations).

It should, therefore, be of concern to the QIP scholar to be able to form a judgment as to whether, for him, this ideal is attainable. The intending theory-builder should have some fairly precise idea of the kind of theory he is hoping to build, rather than hope for inspiration to emerge from the chatter of the computer.

7 The DN model is beautifully simple. Hempel sets it out schematically as follows (ibid p 336):

> "(A full) explanation may be conceived as a deductive argument of the form
>
> (If) $C, C_2, \ldots, C_k$)
> (and given) $L, L_2, \ldots, L_r$) Explanans-sentences
>
> (then) E Explanandum-sentence".

Here, the 'C' sentences represent the antecedent conditions and the 'L' sentences the general laws which, taken together, entail the event (or another general law) represented by the E sentence. To qualify as a "general law", a sentence must be an "exceptionless generalisation" or, as it is also called, a "universal conditional": "If C, C_2, ..., and given L, L_2, ..., then <u>always</u> E". The model is also known as the "covering-law" model because what is to be explained (the Explanandum) is so by virtue of being "covered" by a (more) general law. A valid explanation invariably justifies a prediction (though the converse does not hold) and, if the pre-

diction fails, the purported explanation fails as an explanation. It is important to be aware of the "Fallacy of the Consequent": the truth of an "Explanandum-sentence" does not entail the truth of a "Law-sentence" from which it might have been deduced (a true sentence may be a valid conclusion to a syllogism in which the major or minor premise is false). And, while an accurate prediction may corroborate a law, no finite series of such corroborating instances can establish its truth absolutely. For a "full explanation", the antecedent conditions must, taken together, be both necessary and sufficient: a necessary condition is one without which an event cannot take place; a sufficient condition, one with which it must take place.

8 Questions have been raised about the explanatory adequacy of the model, even for the natural sciences (Keat & Urry 1982 and Armstrong 1983); about whether we can say of any law at present accepted as such in the natural sciences that it is other than approximately true (Newton-Smith 1981) or probably false (Hesse 1974). Doubt has been cast on the falsifiability of any theory (Quine 1961). But these debates can be left to the philosophers. What matters for our present purposes is that there are some laws of which we can be sufficiently confident to treat as "exceptionless" within their range: Newton's celestial mechanics made it possible

(not alone, of course) to set a man on the moon. Such laws are essentially deterministic; and the DN model is likewise. If it is applicable in the "social" world as it is in the "physical", it should be possible in principle accurately to predict individual events in the former by deduction; it could be methodologically sound, in practice, to be doubtful about a (non-probabilistic) exceptionless generalisation in cases where a proposition validly deduced from it failed to "fit the facts". Few QIP scholars seem to be determinists (in any of the several senses that can be given to "determinism" as a philosopical position) but, whatever the philosophical persuasion of any particular scholar in the field, the question of the relevance to it of the DN model needs to be addressed. Should political scientists seek to construct theories that conform to it?

9 In a celebrated article published nearly ten years ago, Almond and Genco answered: "No" (1977). Basing their case on Karl Popper (1972), they affirm roundly that:

> "if we search only for generalizations and regularities in political processes, if we couch our explanations only in terms of the covering-law model, ... we are committing ourselves - whether we recognize it or not - to a disciplinary research program designed to strip away the cloud-like and purposive aspects of political reality in order to expose its 'true' clock-like structure. If politics is <u>not</u> clocklike in its fundamental structure, then the whole program is inappropriate. We believe this to

> be the case: the current quandary in political science can to a large extent be explained by the fact that, by themselves, 'clock-model' assumptions are inappropriate for dealing with the substance of political phenomena" (1977 pp 504 - 5).

Popper might well agree with the thought underlying this passage; but it is important to note that he had simply been posing a problem: "What we need for understanding rational human behaviour ... is something intermediate in character, between perfect chance and perfect determinism - something intermediate between perfect clouds and perfect clocks ... what we want to understand is how such non-physical things as purposes, deliberations, plans, decisions, theories, intentions, and values, can play a part in bringing about physical changes in the physical world" (1972 pp 228 - 9 fn). This is, however, not so much an argument against attempts to apply a deterministic model of explanation to human affairs, as an expression of unease about the propriety of doing so.

10 Alasdair Macintyre develops the argument:

> "The key part that beliefs play in defining political situations, and the fact that beliefs are always liable to be altered by reflection upon the situation, including reflection about the beliefs of other agents, has a crucial consequence: that we cannot ever identify a determinate set of factors which constitute the initial conditions for the production of some outcome in conformity with a law-like regularity" (1971/1973 p 184).

But what is the force of "cannot" in this sentence? Macintyre himself is unsure: "Am I saying what the limits of enquiry are <u>as of now,</u> or what the limits <u>as such</u> are? I am strongly inclined to say that at the moment we have no grounds for answering this question as it stands either way" (ibid p 185). If so, he

should have limited himself to "cannot now" or "cannot for the foreseeable future". At the same time, "cannot ever" clearly reflects his convictions.(Such a contention might be supported in a slightly different way: if the determinate set of factors required could always be identified, human choice would be revealed as illusory; for it would imply perfect powers of prediction and, with it, perfect powerlessness on the part of human agents to influence the course of events. This may not be logically inconceivable (Costa de Beauregard 1968) but we should then be living in a world in which QIP scholars and FPMs alike would be redundant).

11 Macintyre recalls (ibid p 180) a debate between Lord Macaulay and James Mill, the father of JS, in which, interestingly enough, the expected roles of the philosopher and of the historian are (by normal positivist standards) reversed. For Mill was arguing, against Macaulay, that empirical facts cannot yield lawlike generalisations justifying confident prediction, that there is no specifically political science; and modern behaviouralists in the field of International Relations have repeatedly chided their "traditionalist" confrères for not following Macaulay's line. Yet, although Hempel at one time championed the view that the DN model applied to historical explanation (1942) and although a political scientist like Martin Nicholson can remark that "the historian relates to the social scientist as a consumer of his generalisations, and a provider of the raw material, but not as a maker of the

generalisations" (1983 p 186), surprisingly little attention has been paid by International Relations theorists to discussions in the "philosophy of history" which have a direct bearing on our present problem. The views of Ernest Nagel are a case in point.

12 Nagel devotes a whole chapter of his monumental work The Structure of Science (1961) to "Problems in the Logic of Historical Inquiry"; and his first concern is to establish whether or not it is feasible to provide "covering-law" explanations for particular historical events. The underlying rationale is simple. If every event has a cause, it cannot be properly "explained" (in accordance with positivist doctrine) unless it can be shown to be an instance of a general law expressible as a universal conditional; it must be deducible from an exceptionless generalisation taken together with adequate statements of antecedent conditions. It is notorious that the "causal generalisations" proffered by historians fail to meet this specification. But can it be met? Nagel, in tackling this question, takes a particular event, the proclamation of the royal title of Queen Elizabeth I of England when she first ascended the throne: "Elizabeth, by the Grace of God, Queene of England Fraunce and Ireland, defendour of the fayth &c". The problem is to "explain", in accordance with the DN model, the Queen's decision to 'etceterate' her title. For many people, the explanation given by the historian FW Maitland

(from whom Nagel took this example), and given in terms of the Queen's judgments and decisions, would seem satisfactory: her father, Henry VIII, following his quarrel with the Pope, had added to his title the words "and Only Supreme Head on Earth of the Church of England called Ecclesia Anglicana"; when he was succeeded by his elder daughter, Mary, who was (politically) a devout Catholic, the Pope's authority was re-acknowledged and Henry's addition to the title was dropped; when Mary was succeded by her sister, Elizabeth, however, the latter judged it expedient to strike a balance between offending Catholics, on the one hand, and Protestants on the other; hence the "&c". A fully deductive formulation of Maitland's argument, without recourse to generalisations, has been offered by the philosopher Alan Donagan (1970 pp 240 - 1). For Nagel, though, the argument lacks a major premise. He tries one out:

> "Whenever an individual is compelled to announce publicly which one of several alternative policies he is ostensibly adopting, the circumstances under which the announcement is made being such that he believes that the proclamation of a definitive commitment to any one of these policies at that time is fraught with grave perils for himself, the individual will formulate his announcement in ambiguous language" (ibid p 556).

Quite obviously, says Nagel, this statement is false. It cannot serve as a universal conditional. Yet it seems close enough to the tacit generalisation which, in Nagel's eyes, would be needed to make Maitland's argument an "explanation".

13 In his "Function of General Laws in History" (1942/1965), Hempel had argued that the generalisations actually given by historians, when explaining events in history, should be regarded as no more than "explanation sketches", the implication being that, if historians took the trouble to give full explanations, they ought to be able to indicate the universal conditionals forming the major premise(s) of the argument. Nagel disagrees. "The possibility of establishing strictly universal generalizations (sc in the field of human history) cannot be excluded on principle. However, no such generalizations seem to be currently available.[1)] Moreover, ... if well-founded universal laws about human conduct are ever established, they are likely to be formulated in terms of highly refined distinctions which fall outside the customary range of interest of historians" (1961 p 557). A passage in the previous chapter of his book, devoted to the social sciences, makes the reference to "highly refined distinctions" clear: "There are some grounds for doubting that the social sciences are likely to refine their current distinctions beyond a certain point ... unless indeed these disciplines as presently constituted are transformed almost beyond recognition. For suppose that in order to obtain universal social laws, it would be necessary to classify social phenomena in part by reference to minutely

---

1) What Nagel says for history, goes for the social sciences as well. Popper and Hempel, among others, have tried to provide examples to the contrary, but unconvincingly (see Donagan 1964/1966 pp 142 - 5 with references).

differentiated physical and physiological traits of human participants ... (e.g.) differences in chemical composition of their blood, and variations in the spatial distribution of their nerve filaments ..., it may not be advantageous to abandon (a statistical generalization about them) in favor of (a) strictly universal one based on (such variables) ... Few if any social scientists would be competent to analyze social phenomena in terms of those variables" (ibid p 507). The same would go for historians if "Elizabeth's proclamation of her ambiguously worded title could be strictly deduced from premises containing, among others, assumptions formulated in quantum theoretical terms (sic) about the state of her glands, the condition of her neural synapses, the organization of her brain cells, and the intensities of various physical stimuli to which she was exposed" (ibid p 557). In Nagel's view, therefore, we must be content, at least for the foreseeable future, in both history and the social sciences, with what he calls "statistical generalizations". These are "explanations" rather than "explanation sketches" but, of course, an argument based on one "is not a formally valid deductive one, and its premises entail (sic) the conclusion not with necessity but only with some 'degree of probability'" (ibid).

14 "Most if not all historical explanations, like explanations of human conduct in general ... mention only some of the indispensable conditions for those occurrences" (ibid p 559).

Thus, for practical purposes, while being a professional determinist (ibid p 597) and an adherent of the "covering-law" doctrine (ibid p 33), Nagel reaches the same conclusion, though for different reasons, as Macintyre. So does the mathematician, Anatol Rapoport. He states flatly: "There are no 'laws' governing human behavior analogous to physical laws. On the metaphysical level one can, of course, insist that all human actions are, in the last analysis, physical processes ... however, this assumption has no practical significance ... The appeal to the reductionist thesis, whatever force it may have in a philosophical discussion about 'free will', is in no way a contribution to (scientific) method" (1976 p 21). Rapoport had in mind the parallel sometimes drawn (as by Lewis Richardson, the pioneer of arms-race modelling) between political processes and the weather, in terms of complexity and predictability: he considered the analogy misleading because, for the meteorologist, unlike for the political scientist, there was no question "about the relevance of the variables constituting (his) model or the underlying laws of their interaction" (ibid).

15 There are, of course, other possible analogies, drawn from the natural sciences, to which appeal might be made by those who continue to hope that the DN model may be applicable, and not just in principle, to social phenomena. Although, for example, most philosophers and physicists today would dis-

agree not only with Nagel, but also with Einstein and Planck, in holding that the physical processes at work at the sub-atomic level (the concern of quantum physics) are irreducibly indeterminable, it is accepted that they give rise, at a "macro-level", to processes which can be described and predicted in terms of deterministic laws (in spite of continuing debate about the nature of those laws). Applying this analogy to political processes, one might suggest that, notwithstanding the (apparently) random quality of individual human decisions and choices, these might sum, so to speak, into aggregates of behaviour which could be similarly described and predicted. Statistical analysis alone would not suffice to disclose the laws involved, but it could perhaps point the way. To quote Nagel again (1961 p 508), but for a purpose different from his: "Had Galileo sought to establish the laws for freely falling bodies simply by correlating observed data, he would certainly have found that the velocity of falling bodies varies with their weight and shape ... and that there is only a high correlation rather than an invariable proportionality between the distances bodies fall and the squares of the elapsed times of their fall ...". Is this the sort of situation Dina Zinnes had in view when she spoke of the need for a "creative leap" into theory? Another, but quite different, analogy might be drawn from molecular biology. We now know about the double helix and the basically very simple structure of the genetic code; and, if the potential combinations and

recombinations of the elements of the code seem endless, a (horrifyingly) promising start has been made to bringing the details, as it were, of genetic reproduction under human control. Blalock has suggested (perhaps it would be more accurate to say that the suggestion could be read into his remarks) that political processes which seem very complex on the surface might be regarded as combinations of several simpler processes (1984 p 114). Could Factor Analysis be used to attempt to disentangle them? The question can at least be posed despite the cautionary notes struck by Scott Armstrong (1967) and by Gould, who gives a hilarious account of the alleged discovery of "mental energy" by Charles Spearman, the inventor of the rank-correlation coefficient that bears his name (1981 pp 256 - 72).

16 There remains room for doubt. Whether "free-will" can be explained away (by showing that apparently free choices and decisions are in fact governed by universal laws), or whether it can be explained (by showing, however paradoxical this may seem, that human freedom is genuine but determined), is a matter for philosophers and theologians. In the world in which we live, however, it is an empirical fact that even professional determinists behave, at least when they leave their studies, <u>as if</u> they could choose and decide freely: it would seem perverse to say that Caesar <u>had</u> to cross the Rubicon. And this fact would seem to entail, as Macintyre argues, that, where someone's decision is required to trigger an event, there is always some possibility that that event will not happen, though

the probability of its occurrence may seem overwhelming and though the presence of all the factors deemed to constitute the necessary conditions for its occurrence be established. If this argument is not accepted, however, let us return to Nagel's thesis that, if there are universal laws governing human conduct, some of these may be accessible only to neurophysicists. Theoretically, this could be a defensible position; comfortable for a determinist philosopher, if awkward for a similarly committed political scientist. But, if the thesis were unfounded - and Nagel hints that it was put forward at least half in jest - are there other grounds for thinking that any such universal laws might, for practical purposes, be out of reach?

17 In an intriguing footnote to the passage in which they discuss the points of substance (empirical, rather than methodological or philosophical) to have emerged from the review of QIP contained in their book, Hoole and Zinnes write:

> "One of the more interesting issues raised was the degree to which it is reasonable to expect a theoretical statement in international politics to contain the same parameter values for all cases for all years. This issue is of potential interest for all scholars in the field of QIP, and we have not previously seen so much importance attached to it in regard to the development of theory in the international politics field. See chapters 7, 8, 10, 11, 17, 18, 20 and 21." (1976 p 474) (Emphasis added).

The issue is of more than "potential" interest, however. It is crucial for those who hope to build theory in accordance with the DN model. The reason is this (to adapt slightly a

remark of Richard Ashley): theories and models that allow for structural change need to explain it, if the "covering-law" doctrine is accepted, by reference to "a higher-order causal structure" (Ashley 1980 pp 304 - 5; cp also App B). In other words, if QIP scholars come up with "findings" which are clearly not spatio-temporally invariant, these findings cannot directly reflect exceptionless generalisations or universal conditionals: variations in the strength of a correlation coefficient, or in the "relative potencies" of the independent variables in a multiple regression equation, from one period of time, or from one region, to another, themselves require to be "explained" by a higher-order generalisation. Yet we must face the fact that such findings are not unknown (Hoȍle & Zinnes 1976 esp. chapters 20 and 21). This presents problems. Leaving aside the, relatively minor, technical consideration that it can be "difficult to distinguish true change from random measurement error", except where changes in scores are substantial (Blalock 1984 p 72), there is the philosophical difficulty that one could never be sure that any generalisation which purported to "explain", or to help to "explain", what Ashley (following Choucri & North 1975) calls "phase-shifts" (1980 pp 332 - 4) in the statistical regularities, would itself be invariant. For practical purposes, it is reasonable to treat the "laws" of physical nature as unchanging: Planck's Constant can be regarded as just that - a constant. And it <u>may</u> be reasonable, with David Hume (of whose "regularity" theory of causation the DN model is a development),

to treat human nature "as always and everywhere the same", even if the putative neuro-physical laws governing human conduct cannot (yet) be fathomed. But we live in a world which is partly of man's creation and the man-made institutions and technologies that influence the pattern of political relationships are subject to steady or abrupt change. At the very least, it must remain an open question whether any laws governing such change, if they exist, could be grasped and formulated by human intelligence.

18 It follows from all this that the QIP scholar might be unwise to try to build a theory on the DN model, a context-free theory, that is, in which the theorems were logically related in a deductive fashion (given that such theorems would have to take the form of a universal conditional or of a proposition derived from one). This may seem obvious to the great majority of such scholars, and to political scientists generally, who may therefore rightly feel that they escape the strictures of Almond and Genco (paragraph 9 above). But the argument had to be developed for two reasons. First, a few scholars still do believe, or appear to believe, in the feasibility of formulating authentic covering-laws for the social sciences (in addition to the examples cited above, there is Martin Nicholson whose book <u>The Scientific Analysis of Social Behaviour</u> (1983) is in the mainstream of "Sunday-School Positivism"). Second, and more important, it forms an indispensable background to any discussion of the nature of a

"lawlike generalisation". According to North and Willard, "proponents of 'hard' theory (now) ... assume that human affairs are indeterminate - in the sense that iron-clad laws of behaviour do not apply - but unlike their 'soft' theory colleagues, they accept the idea that outcomes can be explained and predicted in terms of statistical probabilities" (1984 p 23) (emphasis added). What is the justification for regarding statistical generalisations as being able to explain and predict outcomes, as being (potentially) lawlike?

19 Before embarking directly on a discussion of this question, we may recall Nicholson's remark (paragraph 11 above) to the effect that the historian should be seen as a 'consumer' of the generalisations of the social scientist and Nagel's conclusion (paragraph 13 above) that the generalisations of both should alike be seen as 'statistical' in character (each author is, of course, referring to generalisations that are "context-free"). Historians may not relish Nicholson's proposed division of intellectual labour; and few would see it as their business to produce specifically statistical generalisations. But the underlying assumption is the same: namely that, however they may be arrived at and defended against subsequent challenge, the context-free generalisations used by social scientists and historians as (part of) their frameworks of "explanation" share the same logic. For Nagel, both sets of generalisations are probabilistic, explicitly or otherwise. The argument elaborated so far supports this view, though it is unfortunate that Nagel

should describe them as 'statistical', since this might be taken to prejudge the sense in which statistical generalisations can properly be regarded as probabilistic.

20 So it is not inappropriate to turn once again to a discussion in the "Philosophy of History", for a critical look at what Hempel has called the "Inductive-Statistical" model of scientific explanation (or "IS model" for short) (1965 pp 380 - 403). In an article which might have been overlooked by social scientists if Keat and Urry had not drawn attention to its importance (1982 pp 12 - 13), Alan Donagan (1966) analyses the model. It goes:

> "Given that an event of the kind A has occured (A_1), and that the <u>statistical</u> probability of an event of the kind B, given A_1, is equal to 1 - e (where e is small), high <u>inductive</u> probability is conferred on the statement that an event of the kind B will occur" (ibid p 131).

Thus, to take a Hempel illustration: if someone, blindfold, were to draw a marble from an urn which he knew to contain 1.000 marbles, one black and the rest white, all thoroughly stirred beforehand, he could, if he drew a white marble, explain this by the high <u>inductive</u> probability of doing so (given that its <u>statistical</u> probability was 0.999). <u>Prima facie,</u> this looks like an "explanation"; but is it? Turning Hempel's argument round, Donagan asks: "Is it the case that when you bought a lottery ticket, which had only one chance in a hundred thousand of winning, this fact <u>explains</u> why you lost? If it does, was it not irrational to buy the ticket at all? And what is to be said when, against the odds, you draw a black ball or

a winning ticket?" (ibid p 133). Donagan continues:

> "In cases of this sort the obvious thing to say is that there is no explanation of any individual outcome. You will be deceived into imagining that there is only if you confound what it was reasonable to expect with what has been explained (emphasis added). Reasonable expectations and explanations differ fundamentally ... With respect to explanation, chance situations where the odds are equal do not differ from those where the odds are fifty to one or a thousand to one" (ibid).

Richard Jeffrey puts much the same point slightly differently:

> "(Where) the probability of the phenomenon to be explained is so high ... as to make no odds in any gamble or deliberation ... (Hempel's) inferential model of explanation seems unexceptionable. But where the strength of the inference is more modest, I think it simply wrong to view the inference as an explanation" (1969/1971 p 27).

The IS model has also been criticised by other philosophers [2] but what matters is to note that a statement of the form "Given A, then, with probability p, B", or a statement reporting a high correlation between A and B, may lead us to expect, but does not "explain", "B, given A". If Mrs. T goes to the Bingo Parlour practically every night of the week, that may be said to "explain" the fact that she is rarely at home in the evening; it does not explain itself - she may be

---

[2] Notably Wesley Salmon, who shows that the model satisfies neither necessary nor sufficient conditions for "explanation" and that there may, in certain circumstances, be fairly high inductive probability for the occurence of an event which is mathematically highly improbable (1970/1975). These arguments are endorsed by Jonathan Cohen, but the latter's attempt to rescue the notion that statistical generalisations can "explain" individual events would seem to depend too heavily on the validity of a "propensity" theory of probability (1977 pp 295 - 309), although he hints at its limitations (ibid pp 22 - 24).

an inveterate gambler, in need of familiar company, or in love with Mr. W, another assiduous attender. This is not to say, however, that statistical generalisations cannot form part of (the evidence for) an explanation.

21 A somewhat similar analysis can be applied to the complex statistical generalisations which take the form of multivariate equations. Blalock (1984 p 91) sees these as the "causal laws" of social science and gives as an example the following linear equation:

$$Y = a + b_1X_1 + b_2X_2 + b_3X_3 + u$$

where Y is the dependent variable, X_1, X_2 etc. the independent variables, a and b are constants and u is the disturbance term. The changes in the value of Y are commonly said to be "explained" by changes in the bX's, with u accounting for the "unexplained" variance. It is possible (and statisticians have to guard against doing so artefactually) to construct such an equation without a disturbance term and thus to "explain" the total variance in Y. This begins to look like "total explanation"; but if the equation "tells us that, if X_1 were to change by one unit, and if X_2 and X_3 were held constant, then the <u>expected</u> change in Y would be b_1 units in the Y variable, (which) means that <u>over the long run,</u> changes in Y would <u>average out</u> to this amount" (ibid), such a "causal law" does not purport to "explain" why a particular sack of potatoes (say) sold at a particular price, let alone

why that or any particular sale took place. Assuming that the parameters hold constant, we should have grounds for expecting the price to fall within a certain range over a certain period. By contrast, a very full, and some would say adequate, explanation could be given ex post facto of the purchase of a sack of potatoes by, say, Mrs. T.

22 "Explanation" and "Prediction" are closely related. According to Nicholson, they become (within the DN model at least) "very close to the same thing, differing only in the temporal detail that with explanation the explanandum has already happened while with prediction it has yet to come" (1983 p 151). For Hempel, however, while "every adequate explanation is also a potential prediction", the thesis that "every adequate predictive argument also affords a potential explanation" remains an open question (1965 pp 374 & 376). Thus, to use his example, an outbreak of Koplik spots predicts measles, but does not explain it. Similarly, a high correlation between A and B may make A a good 'predictor' of B, without "explaining" it; the explanation may lie in a "common cause", C. Perhaps, therefore, it would not be unreasonable, for our present purposes, to proceed on the basis that, although (unless, of course, we were contemplating any "ultimate regularities" that might govern the behaviour of the cosmos) we would expect there to be an explanation in principle of why A predicts B, where A was not (part of) the explanation of B, we need not know that explanation before accepting that

A was a good predictor. Other things being equal, we can reasonably expect Mrs. T to continue to be a regular visitor to the Bingo Parlour. Consequently, if it is accepted that an "If ..." clause should not be construed as an "explanation" of a "then, with a probability p ..." clause, it would not follow that statistical generalisations could not serve as a justifiable basis for prediction.

23 It would be an understatement to say that there is, as yet, no agreed theory of probability and no agreement over all on the logic of statistical inference. Nevertheless, most philosophers would agree with Birnbaum when he writes:

> "In order to use statistical inference, it must be the case that every unit of analysis in the population (sc in the statistical sense of that word) has some chance (no matter how small) of being included in the sample. Hence, attempting to use statistical inference to predict the future (treating present data as a sample from data now and in the future) ... (is) simply invalid ... Theoretical reasons must be adduced to validate such generalisations (sc beyond the data) ... Statistics cannot do our scientific thinking for us." (1981 p 9)

Such "theoretical reasons" - perhaps it would be better to call them assumptions - necessarily include the assumption that other things will remain more or less equal: the notorious ceteris paribus principle, without which, indeed, prediction would not be feasible. They also include, for multivariate equations, the assumption that the parameters really are constants and that supposedly independent variables are not just correlated with one another (as is usual in non-experimental settings) but will

not change as a result of changes in the others (cp Blalock 1984 p 92). But (as was noted in paragraph 17 above), there is evidence of "parameter instability" in some of the "findings" of QIP scholars.

24 The problem does not end there. It is crucial for statistical inference that what Birnbaum calls the "units of analysis" should be homogeneous in the sense in which one "death", let us say, can be regarded as "identical" to any other "death". Taking "wars" as his example, Rapoport writes:

> "It is not obvious that these large complexes of events are instances of the same event in the same way as, say, eclipses of the sun or tornadoes are obviously instances of the same event. The question arises as to how many of these classes of complex events can be subsumed under the same category. Are the wars waged by the Comanchees, by the princes of eighteenth century Europe, by the Crusaders, and by the United States since 1950 simply variants of a well-defined event and so deserving a generic name? The answer to this question is far from obvious ... Nevertheless, it is important to keep (it) in mind ... The quantitative empirical approach to social phenomena leans heavily on statistical methods that, in turn, rest on probabilistic assumptions that in turn, are bases on the notion of repeated 'identical' events" (1976 p 15).

The same question can be asked, if less obviously, of "States". Treating states as homogeneous, when they are not, and "averaging out" their behaviour, leads to fallacious inference (cp section II). It can also lead to the formulation of statistical generalisations which are liable to be misinterpreted. Zinnes and Wilkenfeld give a classic example (1971 pp 168 - 9 with references): using factor analysis, Rummel had concluded - a con-

clusion supported independently by Tanter - that there was no relationship between the domestic and the foreign conflict behaviour of states; but by dividing states according to form of government and redoing the Rummel and Tanter analysis within each of the resulting groups, "Wilkenfeld found significant correlations between certain of the domestic conflict factors and some of the foreign conflict factors". Aggregate data analysis depends for its usefulness on the data not being over-aggregated.

25 The purpose of these observations is not to cast doubt on the feasibility of statistical inference in the field of International Relations. That would be quite unwarranted. It is simply to indicate some of the assumptions that must be made, and validly made, before a statistical generalisation can be used as the basis for a valid predictive inference. These particular assumptions might be seen collectively as assumptions of homogeneity: homogeneity of the units of analysis in the present data, homogeneity of the present statistical "population" with any future "population" regarding which an inference is to be made and, with reference especially to multivariate equations, homogeneity of present and future parameters. Indeed, the "assumptions" are, rather, conditions that have to be satisfied; and for certain, particular, contexts in the social sciences we can be reasonably sure that they <u>are</u> satisfied. But can we be so sure of the

possibility of formulating statistical generalisations governing human conduct which meet these conditions across the board? Can we hope to arrive at social science equivalents of the context-free probabilistic generalisations to be found in the natural sciences, e.g. certain laws of genetics or the laws of radio-active decay, which assert invariant probabilities?

26 To ask these questions is to ask whether Dina Zinnes was being realistic in claiming that the QIP community appears to be moving slowly "toward lawlike generalisations" (1980 p 360). For, although philosophers of science may disagree about the logical characteristics of a scientific law (cp Nagel 1961 pp 47 - 78), invariance must be one of them (if it makes sense to talk of "scientific laws" at all). Yet to say that it was improbable that such laws exist in the domain of enquiry covered by the social sciences might appear paradoxical (given that we make generalisations all the time about how normal people can be expected to behave under normal circumstances) and even contradictory (since such a statement might itself be regarded as a statement of invariant probability). A better way of approaching the problem would be to ask, instead, how an invariant <u>statistical</u> generalisation, if stumbled across, could be recognised as such and put to the test. By contrast, a probabilistic, yet lawlike, statement of the half-life of $plutonium_{239}$ can be backed by reference to a <u>theory</u> of radio-active decay amply supported by observation. Like many other

scientific laws, moreover, its lawlike character is supported definitionally: if a quantity of supposed plutonium decayed at a very different rate from the range expected, it would not be "plutonium".

27 The use in QIP literature of expressions such as "lawlike generalisation" conveys the impression that some seekers after theories of International Relations (scientific theories, that is) still hanker for theories on the DN model or its equivalent. Kenneth Waltz, for example, whose book Theory of International Politics (1979) has not been examined in this essay because the theory he offers is not quantitative ("statistical operations", he writes, "cannot bridge the gap that lies between description and explanation" (ibid p 3)), seems to evince such a yearning. As an "instrumentalist", he holds that assumptions contained in a theory are not factual, yet he castigates as "false" the assumption that balance of power systems require a minimum of three players (ibid pp 117 - 8). At another point, he says of his own balance of power theory that its "predictions are indeterminate" and then proceeds to cite the formalisation of the Franco-Russian alliance in 1894 as a "hard case" confirmation of it (ibid p 125). Such nostalgia is understandable. But that it exists at all points to a lacuna. There seems to be little clear conception among QIP scholars as a whole, and certainly no consensus, about what kind or kinds of theory-model to follow. There is no clarity about what "lawlike

generalisation" might be taken to mean. If there is an epistemologist in the house, his services are urgently required. Meanwhile, the FPM will have to wait.

REFERENCES

Section I: Introduction

Bolingbroke, HStJ, Viscount (1754) Works, David Mallet, London

Bull, H (1966) "International Theory: the Case for a Classical Approach", World Politics, 18, 3, pp 361 - 377

Cable, Sir J (1981) "The Useful Art of International Relations", International Affairs, 57, 2, pp 301 - 314

Callières, F de (1716/1983) The Art of Diplomacy, trans. HMA Keen-Soper & KW Schweizer, Leicester University Press, Leicester

Chesterfield, PDH, Earl of (1774/1893) The Letters, ed. J Bradshaw, Swan Sonnenschein, London

Easton, D (1965) A Systems Analysis of Political Life, Prentice-Hall, Englewood Cliffs, NJ

Frankel, J (1981) "Conventional and Theorising Diplomats", International Affairs, 37, 4, pp 537 - 548

Hopmann, PT, Zinnes, DA &Singer, JD (eds) (1980) Cumulation in International Relations Research, Monograph Series in World Affairs, Vol 18, Bk 3, University of Denver, Denver, Colorado

Jones, SD & Singer, JD (1972) Beyond Conjecture in International Politics, FE Peacock, Itasca, Ill.

Kaplan, M (1957) System and Process in International Politics, Wiley, New York

(1969) "The New Great Debate: Traditionalism vs Science in International Relations", in K Knorr & JN Rosenau (eds) Contending Approaches to International Politics, Princeton University Press, Princeton, NJ

Kenyon, J (1983) The History Men, Weidenfeld & Nicolson, London

McClelland, CA (1961) "The Acute International Crisis", in K Knorr & S Verba (eds) The International System, Princeton University Press, Princeton, NJ, pp 182 - 204

Nicolson, H (1952) Diplomacy (2nd edition), Oxford University Press, London

O'Leary, MK & Coplin, WD (1975) Quantitative Techniques in Foreign Policy Analysis and Forecasting, Praeger, New York

Rapoport, A (1983) Mathematical Models in the Social and Behavioral Sciences, Wiley, New York

Reynolds, C (1973) Theory and Explanation in International Politics, Martin Robertson, London

Rosenau, JN (1981) The Scientific Study of Foreign Policy (revised and enlarged edition), Frances Pinter, London, & Nichols, New York

Singer, JD (ed) (1968) Quantitative International Politics: Insights and Evidence, The Free Press, New York

Smith, S (1979) "Brother, can you Paradigm?: A Reply to Prof Rosenau", Millenium, 8, 2, pp 235 - 245

(1982) "Foreign Policy Analysis: British and American Orientations and Methodologies", paper read to the annual meeting of the International Studies Association, Cincinnati, March 24

Snyder, RC, Bruck, HW & Sapin, B (1962) Foreign Policy Decision making, Free Press of Glencoe, New York

Wicquefort, M de (1690) L'Ambassadeur et ses fonctions, (dernière édition), Pierre Marteau, Cologne

Wilkenfeld, J, Hopple, GW, Rossa, PJ & Andriole, SJ (1980) Foreign Policy Behavior: the Interstate Behavior Model, Sage, Beverly Hills & London

Young, O (1972) "The Perils of Odysseus: On Constructing Theories of International Relations", in R Tanter & RH Ullman (eds) Theories and Policy in International Relations, Princeton University Press, Princeton, NJ

Zinnes, DA (1976) Contemporary Research in International Relations: A Perspective and a Critical Appraisal, The Free Press, New York

(1980) "Why War? Evidence on the Outbreak of International Violence", in TR Gurr (ed) Handbook of Political Conflict, The Free Press, New York, pp 331 - 360

Section II: A Case Study of a Case Study: Patrick J McGowan et al on "Predicting Political Instability in Tropical Africa"

Baumann, H & Westermann, D (1957) Les peuples et les civilisations de l'Afrique, French trans. L Homburger, Payot, Paris

Cornevin, M (1978) Histoire de l'Afrique contemporaine, (2nd edition), Payot, Paris

Dowdy, S & Wearden, S (1983) Statistics for Research, Wiley, New York

Floud, R (1973) An Introduction to Quantitative Methods for Historians, Methuen, London

Gurr, TR (1973) "A Causal Model of Civil Strife using New Indices", American Political Science Review, 62, pp 1104 - 24

Morrison, DG & Stevenson, HM (1971) "Political Instability in Independent Black Africa", Journal of Conflict Resolution, 15, pp 347 - 68

Newton-Smith, WH (1981) The Rationality of Science, Routledge & Kegan Paul, Boston, Mass., and London

O'Leary, MK & Coplin, WD (1976 Quantitative Techniques in Foreign Policy Analysis and Forecasting, Praeger, New York

Section III: A Critique of a Critique: the "Colouring it Morgenthau" thesis of John A Vasquez

Banks, M (1985) "Where we are now", Review of International Studies, 11, 3, pp 215 - 233

Blalock, HM (1960) Social Statistics, McGraw Hill, New York

(1984) Basic Dilemmas in the Social Sciences, Sage Publications, Beverly Hills & London

Gillies, DA (1973) An Objective Theory of Probability, Methuen, London

Gilpin, RG (1984) "The Richness of the Tradition of Political Realism", International Organisation, 38, 2 pp 287 - 304

Hoole, FW & Zinnes, DA (eds) (1976) Quantitative International Politics: An Appraisal, Praeger, New York HZ

Job, BL & Ostrom, CW (1976) "An Appraisal of the Research Design and Philosophy of Science of the Correlates of War Project" in HZ pp 43 - 66

Jones, SD & Singer, JD (1972) Beyond Conjecture in International Politics, FE Peacock, Itasca, Ill.

Kuhn, TS (1970) The Structure of Scientific Revolutions (2nd edition), International Encyclopaedia of Unified Science, 2, 2, University of Chicago Press, Chicago

Lakatos, I (1970) "Falsification and the Methodology of Scientific Research Programmes", in Lakatos & Musgrave (eds) Criticism and the Growth of Knowledge, Cambridge University Press, London, pp 91 - 196

Mansbach, RW & Vasquez JA (1981) In Search of Theory: A New Paradigm for Global Politics, Columbia University Press, New York

Morgenthau, HJ (1960) Politics among Nations (3rd edition), AA Knopf, New York

Morrison, DE & Henkel, RE (1970) "Significance Tests Reconsidered" in Morrison & Henkel (eds) The Significance Test Controversy, Aldine Publishing Co., Chicago

Popper, K (1952) The Open Society and its Enemies (2nd edition), Routledge & Kegan Paul, London

Rothstein, RL (1972) "On the Costs of Realism", Political Science Quarterly, 87, 3

Singer, JD (1976) "Rejoinder to the Critiques" in HZ pp 128 - 145

Rummel, RJ (1968) "The Relationship between National Attributes and Foreign Conflict Behavior" in JD Singer (ed) Quantitative International Politics: Insights and Evidence, The Free Press, New York, pp 187 - 214

(1976) "Comments on the Reviews of the Dimensionality of Nations Project" in HZ pp 219 - 243

Vasquez, JA (1979) "'Colouring it Morgenthau': new evidence for an old thesis on international politics", British Journal of International Studies, 5, 3, pp 210 - 228

(1983) The Power of Power Politics: A Critique, Rutgers University Press, New Brunswick, NJ

Willetts, P (1984) "The Politics of Global Issues: Cognitive Actor Dependence and Issue Linkage" in RJ Barry Jones and P Willets (eds) Interdependence on Trial, Pinter, London, pp 83 - 110

Section IV: Theory and Change

Almond, GA & Genco, SJ (1977) "Clouds, Clocks and the study of Politics", World Politics, 29, 4, pp 489 - 522

Armstrong, DM (1983) What is a Law of Nature?, Cambridge University Press, London

Ashley, RK (1980) The Political Economy of War and Peace, Pinter, London, & Nichols, New York

Birnbaum, I (1981) Introduction to Causal Analysis in Sociology, Macmillan Press, London

Blalock, HM (1984) Basic Dilemmas in the Social Sciences, Sage Publications, Beverly Hills & London

Braithwaite, RB (1953) Scientific Explanation: A Study of the Function of Theory, Probability and Law in Science, Cambridge University Press, Cambridge

Bueno de Mesquita, B (1980) "Theories of International Conflict: An Analysis and an Appraisal" in TR Gurr (ed) Handbook of Political Conflict: Theory and Research, The Free Press, New York, pp 361 - 98

(1981) The War Trap, Yale University Press, New Haven & London

(1984) "A Critique of 'A Critique of The War Trap'", Journal of Conflict Resolution, 28, 2, pp 341 - 60

Caporaso, JA (1976) "A Philosophy of Science Assessment of the Stanford Studies in Conflict and Integration" in FW Hoole & DA Zinnes (1976) pp 354 - 82

Choucri, N & North, RC (1975) Nations in Conflict: National Growth and International Violence, WH Freeman, San Francisco

Cohen, LJ (1977) The Probable and the Provable, Clarendon Press, Oxford

Costa de Beauregard, O (1968) "Time in Relativity Theory: Arguments for a Philosophy of Being" in JT Fraser (ed) The Voices of Time, Allen Lane Penguin Press, London

Donagan, A (1964) "The Popper-Hempel Theory Reconsidered" in WH Dray (ed) (1966) Philosophical Analysis and History, Harper & Row, New York & London pp 127 - 59

(1970) "Can Philosophers Learn from Historians?" in HE Kiefer & MK Munitz (eds) Mind, Science and History, State University of New York Press, Albany, pp 234 - 50

Gould, SJ (1981) The Mismeasure of Man, Norton, New York & London

Hempel, CG (1942) "The Function of General Laws in History" reprinted in Hempel (1965)

(1965) 'Aspects of Scientific Explanation' and Other Essays in the Philosophy of Science, The Free Press, New York

Hesse, M (1974) The Structure of Scientific Inference, Macmillan, London

Hoole, FW & Zinnes, DA (1976) Quantitative International Politics: An Appraisal, Praeger, New York

Jeffrey, RC (1969) "Statistical Explanation vs Statistical Inference" in Salmon et al (1971) pp 19 - 28

Keat, R & Urry, J (1982) Social Science as Theory (2nd edition), Routledge & Kegan Paul, London & Boston

Mcintyre, A (1971) "Is a Science of Comparative Politics Possible?", reprinted in A Ryan (ed) (1973) The Philosophy of Social Explanation, Oxford University Press, London, pp 171 - 88

Majeski, S & Sylvan, S (1984) "Simple Choices and Complex Calculations: A Critique of The War Trap, Journal of Conflict Resolution, 28, 2, pp 316 - 40

McGowan, PJ (1975) "Meaningful Comparisons in the Study of Foreign Policy" in CW Kegley et al (eds) International Events and the Comparative Analysis of Foreign Policy, University of South Carolina Press, Columbia, SC, pp 52 - 87

Nagel, E (1961) The Structure of Science: Problems in the Logic of Scientific Explanation, Harcourt, Brace & World, New York

Newton-Smith, WH (1981) The Rationality of Science, Routledge & Kegan Paul, Boston, Mass., & London

Nicholson, M (1983) The Scientific Analysis of Social Behaviour, Pinter, London

North, RC, Holsti, OR, Choucri, N (1976) "A Reevaluation of the Research Program of the Stanford Studies in International Conflict and Integration" in Hoole & Zinnes (1976) pp 435 - 59

North, RC & Willard, MR (1984) "The Post-Behavioural Debate: Indeterminism, Probabilism and the Interaction of Data and Theory" in M Banks (ed) Conflict in World Society: A New Perspective on International Relations, Wheatsheaf, Brighton, & St Martin's Press, New York, pp 22 - 38

Popper, K (1972) "Of Clouds and Clocks: An Approach to the Problem of Rationality and the Freedom of Man" in Objective Knowledge: An Evolutionary Approach, Clarendon Press, Oxford

Putnam, H (1981) Reason, Truth and History, Cambridge University Press, Cambridge

Quine, WVO (1961) "Two Dogmas of Empiricism" in From a Logical Point of View, Harvard University Press, Cambridge, Mass.

Rapoport, A (1976) "Mathematical Methods in Theories of International Relations: Expectations, Caveats and Opportunities" in DA Zinnes & JV Gillespie (eds) Mathematical Models in International Relations, Praeger, New York, pp 10 - 36

Salmon, WC (1970) "Statistical Explanation" reprinted in Salmon et al (1971)

Salmon, WC, Jeffrey, RC & Greeno, JG (1971) Statistical Explanation & Statistical Relevance, University of Pittsburgh Press, Pittsburgh

Salmon, WC (1975) "Theoretical Explanation" in S Körner (ed) Explanation, Basil Blackwell, Oxford, pp 118 - 145

Scott Armstrong, J (1967) "Derivations of Theory by Means of Factor Analysis, or Tom Swift and his Electric Factor Analysis Machine", The American Statistician, 21, 5, pp 17 - 31

Shapiro, MJ (1981) Language and Political Understanding: The Politics of Discursive Practices, Yale University Press, New Haven & London

Singer, JD (1976a) "The Correlates of War Project: Continuity, Diversity and Convergence" in Hoole & Zinnes (1976) pp 21 - 42

(1976b) "A Rejoinder to the Critique" in Hoole & Zinnes (1976) pp 128 - 45

von Neumann, J & Morgenstern, O (1944) Theory of Games and Economic Behavior, Princeton University Press, Princeton, NJ

Waltz, KN (1979) Theory of International Politics, Random House, New York

Wilkenfeld, J, Hopple, GW, Rossa, PJ & Andriole, SJ (1980) Foreign Policy Behavior: the Interstate Behavior Analysis Model, Sage, Beverly Hills & London

Yuen Foong Khong (1984) "War and International Theory: A Commentary on the State of the Art", Review of International Studies, 10, 1, pp 41 - 63

Zinnes, DA & Wilkenfeld, J (1971) "An Analysis of Conflict Behavior of Nations" in WF Hanrieder (ed) Comparative Foreign Policy: Theoretical Essays, McKay, New York