1

MicroD

National Library
of Canada

Bibliothèque nationale
du Canada

Canadian Theses Service

Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, tests publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30.

THE SOCIAL PSYCHOLOGY OF DETERRENCE THEORY

by

Kevin P. Leppmann

A thesis submitted to the
Norman Paterson School of International Affairs
in partial fulfilment of the requirements for the degree of
Master of Arts in International Affairs

Carleton University
Ottawa, Ontario, Canada

January 1988

© copyright

1988, Kevin P. Leppmann

Permission has been granted to the National Library of Canada to microfilm this thesis and to lend or sell copies of the film.

The author (copyright owner) has reserved other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without his/her written permission.

L'autorisation a été accordée à la Bibliothèque nationale du Canada de microfilmer cette thèse et de prêter ou de vendre des exemplaires du film.

L'auteur (titulaire du droit d'auteur) se réserve les autres droits de publication; ni la thèse ni de longs extraits de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation écrite.

ISBN 0-315-46257-4

The undersigned hereby recommend to the Faculty of
Graduate Studies and Research acceptance of this thesis,
submitted by Kevin P. Leppmann, in partial fulfilment of
the requirements for the degree of Master of Arts.

Brian W. Tomlin, Director
The Norman Paterson School of International Affairs

Professor L. Strickland, Supervisor

Carleton University

Abstract

The intellectual adequacy of deterrence theory is challenged on three levels. First, the orthodox epistemology and method of behavioural science which informs deterrence theory is re-appraised in light of recent post-behavioural criticism. The behavioural research on the psychology of deterrence is then surveyed within the context of this critique, and is found to be a self-reproducing intellectual exercise serving to legitimize the international security status quo. Finally, the epistemological status of deterrence theory's chief proponent, the Strategist, is assessed within the framework of Kruglanski's social psychological Theory of Lay Epistemology, highlighting the Strategist's normative commitment to Realism and a complementary epistemological commitment to rational analysis.

Acknowledgements

Many theses reflect the shadow agendas of candidates' supervisors or department. I was very fortunate to have two supervisors, Dr. Lloyd Strickland and Dr. John Sigler, who, for good or ill, supported me in formulating and realizing my own particular intellectual journey. I am deeply grateful for their remarkable encouragement, trust, and patience throughout the preparation of this thesis.

Four other friends also deserve special mention. Dr. Robert Malcolmson, of Queen's University, must take responsibility for pointing me in my present direction. Dr. Brian Mandell provided very helpful advice and guidance, particularly in earlier drafts of the second chapter.

Peter Lindfield has been invaluable both as a friend and as a dialectician; his constant nagging was instrumental to the development of my own thinking (though he would deny any responsibility for what appears in the following pages).

Finally, the completion of this thesis would have been impossible without the love, support and apartment of my companion, Andrea Fessler.

Table of Contents

Abstract	ii
Acknowledgements	iii
Table of Contents	iv
INTRODUCTION	1

CHAPTER ONE:

The Political Psychology of Behavioural Science	6
Behavioural Science Orthodoxy	7
Mandate	8
Epistemology	10
Method	12
Recent Challenges to the Orthodoxy	14
A Proliferation of Paradigms	15
The Fact-Value Fallacy	18
Theory as Rhetoric	19
Bias Toward the Status Quo	21
Toward Emancipatory Social Theory	22

CHAPTER TWO: Psychology and Deterrence

Two Ambiguities of This Century	25
The Emergence of Deterrence Theory	28
Defining Deterrence	28
Defining Deterrence Theory	32
The Rational Actor	36
The Second Wave	36
Game Theory as Rational Analysis	37
Game Theory and Strategy	41
Critique of the Rational Actor	48
The Political Irrelevance of Game Theory	49
The Psychological Irrelevance of Game Theory	51
The Normative Inadequacy of Game Theory	57
The Reification of Strategic Analysis	61

Table of Contents (cont'd)

Modifying the Rational Actor	62
Abandoning Deterrence Theory:	
The Social Psychology of Conflict	63
Saving Deterrence Theory: The Third Wave	67
Dilemmas of Political Psychology:	
Radicals and Traditionalists	75
Rejecting the Rational Actor	77
The Poverty of a Behavioural Deterrence Theory ...	77
Toward an Emancipatory Theory of Deterrence	81
 CHAPTER THREE: Deterrence Theory as Lay Epistemology ..	83
A Recapitulation	83
Arie Kruglanski's Theory of Lay Epistemology	87
Kruglanski's Critique of Prevalent Theories of Human Inference	87
Kruglanski's Theoretical Alternative	93
Implications for an Analysis of Deterrence Theory	97
Deterrence Theory as a Lay Epistemology	99
The Hard Core: Realism	100
The Protective Belt: Rational Analysis	103
Alternative Epistemologies	106
 CONCLUSION: Social Science as Political Activism	110
 Notes	112
References	116

INTRODUCTION

Few historical discontinuities have been more recognized and less understood than the dawn of the nuclear age. The 1945 explosions over Japan symbolized not only the death of the old European-centered modern world, but also the birth of a new post-modern world, a "global village" rife with global conflict and insecurity. Every new era, however, has its roots in the old, and two of the most stubborn and enduring inheritances that plague the contemporary era are the unrestrained technology which first emerged with the second industrial revolution of the late nineteenth century, and the hopelessly anachronistic Westphalian system of sovereignty, designed almost four centuries earlier.

Albert Einstein's well-known diagnosis of this hybrid is worth yet another recital: "The unleashed power of the atom has changed everything save our modes of thinking, and we thus drift toward unparalleled catastrophes." Most observers agree that the "unleashed power" is here to stay. But what about our "modes of thinking"? Those who wield the unleashed power are usually rather sanguine, professing that they have changed their thinking: While war is still a central reality for much of the Third World, the argument usually goes, those First and Second World leaders, who hold the globe hostage, have abandoned the historical

symbiosis between sovereignty and war. "Nuclear weapons," they claim "have kept the peace for forty years." The dominant rationale for maintaining huge arsenals, we are told, is not for war-fighting, but for deterrence. What had forever been only a by-product of sovereignty would now move to centre-stage.

But this truncated version of the Westphalian global order is rather schizophrenic, and the ongoing debates over defence policy serve to highlight these state leaders' futile attempt to "live simultaneously in the two worlds -- the nuclear, scientific world and the pre-nuclear military and political one." (Schell, 1982) Practicing deterrence writ large, it turns out, has been more politically controversial, more socially and economically costly, and more militarily unstable precisely because of the tensions concomitant with abandoning large scale war while maintaining sovereignty.

Various observers, using various discourses, have diagnosed the problems of a deterred world, and the problems of constructing an alternative normative order which better manages and resolves social conflict. But within social science, the contemporary era's intellectual vanguard for human emancipation, the primary intellectual contribution to the debate has been the unprecedented behavioural-scientific program of strategic studies. This program, supported by a variety of disciplines, concerns

itself, implicitly or explicitly, with the psychological nature of a posited behavioural relationship of coercive influence between two or more nations (or their prominent decision makers), in other words, deterrence. The products of this inquiry usually have been directed toward the state, in an attempt to enlighten diplomacy and foreign policy. But by any reasonable account, this program has failed miserably in the way of a new international security agenda, producing and reproducing an ever more dangerous national security culture. Beneath the amoral, objective social scientific clothes lies an immoral emperor of global militarization. The irony is that an intellectually weak but disproportionately influential branch of social science has become part of the problem rather than part of its solution.

The following inquiry does not attempt to advance our understanding or explanation of deterrence. My purpose is to launch a persuasive attack against the legitimacy of deterrence theory and those strategists who produce it. My attack has three components, corresponding to the three subsequent chapters.

The first chapter will discuss the main disciplines which inform deterrence theory. Despite normative and methodological diversity, I will focus on the trans-disciplinary consensus that endures -- a consensus that views social scientists as potentially privileged

observers -- and the implications of recent challenges to this orthodox view. Some of these challenges derive from critiques of positivist philosophy; but instead of appropriating the critique of the natural sciences *in toto* (as the behavioural sciences have attempted to appropriate method), some social scientists have themselves provided a valuable critical perspective specific to the methods and mandate of social science, with a call to move away from instrumental knowledge, toward emancipatory theory.

In the context of this critique, the second chapter will provide a chronological and selective account of the last forty years of deterrence theory. The emphasis will be on continuities and discontinuities in the often sterile analysis of deterrence as a *behavioural relationship between states*. This theoretical work, while sometimes delivering instrumental knowledge, has delivered very little in the way of emancipatory theory, and therefore cannot really get us what we need. Such emancipatory theory, rather than positing more or less immutable laws of state behaviour, would have to address the historically situated adherence to a set of normative ideas and intellectual practices that vitiates theorists' attempts to transcend the deterrence agenda.

One possible emancipatory theory is elaborated in the third chapter. I will outline the main features of Arie Kruglanski's "Theory of ~~Day~~ Epistemology," which

synthesizes a large portion of earlier affective and cognitive research in attitudes and social judgment. I will suggest how such a theoretical framework answers the challenges discussed in the previous chapters, and I will present an alternative deterrence theory as a "lay epistemology," a social psychological construction that, like any other construction, can claim no ultimate foundation in transcendental Truth. My object of study, then, is not the phenomena of deterrence, but rather, the social psychological context which conditions and sustains the phenomenon of deterrence theory.

My purpose throughout is to reclaim Professor Herbert Kelman's mandate that social science should lend hands and feet to morality. I conclude with some of the implications such a new conception has for peace research, peace education, and the urgent task of "changing our modes of thinking."

CHAPTER 1: The Political Psychology of Behavioural Science

Over 2000 years ago, long before the current division between political and psychological inquiry, Plato addressed the contrast between episteme and doxa, between the knower's search for stable and immutable truths, and the human experience of change, contingency, and disorder. This dichotomy emerged again as a signifying characteristic of the European Enlightenment, with the commitment to discovering immutable rules of natural governance, to separate the everyday contingent noise from the foundational language of eternity. With only a few qualifications, the behavioural sciences took the rational method and epistemology as their model, separating themselves from the more methodologically diverse world of other social sciences.

The behavioural scientific search for order, predictability, and ultimately, control, has both psychological antecedents and political implications. Whether humans are innate seekers of stability and control or have merely been socialized in a Western Platonic civilization to value these ideals, they are nevertheless psychologically reinforced. Learning and maturation, at least in our culture, are almost always equated with the detection of stable natural and social patterns, in order to better cope with the world around us. The hegemonic

means for this end has, of course, been the edifice of science, both natural and social: enduring truths are to be achieved with the acquisition and cumulation of scientific knowledge.

But this devotion to stability as a definition of knowledge is profoundly political. Non-analytic political philosophers, from Nietzsche to the Frankfurt school and beyond, have warned of various political inter-relationships between power and knowledge. What is to be accepted as knowledge, what is to be valued as useful knowledge, and what political privileges are to be afforded those with the privileged status as "knowers"? Whatever the epistemology ascribed, it is clear that the acquisition and wielding of knowledge are both psychological and political activities, and nowhere have these activities been more acute than in the behavioural sciences.

BEHAVIOURAL SCIENCE ORTHODOXY

Certainly, much of deterrence theory, at least in its original form, has emerged from outside the behavioural sciences proper, in think-tanks and other state-directed institutions, under the name of strategic studies (with its pre-nuclear Clausewitzian ancestors). As well, many behavioural scientists concerned with the issues that deterrence theory ostensibly speaks to -- international security, coercive influence, etcetera -- avoid reference

to deterrence in their work, for a variety of normative and theoretical reasons. But over the years, two major developments have implicated behavioural science: First, strategic studies, due in part to specific government funding, has become a major branch of political science. Second, behavioural scientists outside strategic studies have also attempted to advise defense policy.

Mandate

In this context, the behavioural sciences refer primarily to American social psychology; the behavioural influence in international relations theory, and the interdisciplines of conflict analysis, political psychology and peace research. Of course, none of these discourses is concerned only with deterrence, nor even solely with war and peace, and there is significant but often confusing overlap between them. But being summarily defined as behavioural science suggests complementary, if not similar, mandates. For while there may be sharp dissension regarding normative disposition within and between these discourses, a general theme is the commitment to "inform social policy," to contribute to an enlightened rational society by providing "expert" knowledge to those policy-makers most able to use it. Some participants would take exception to this mandate, but they are exceptions that prove the rule.

For example, the most influential behavioural text on the *Theory of International Politics* (Waltz, 1979) concludes with a chapter on how best to use this theory toward the "management of international affairs," relying almost exclusively on national policy-makers to do the managing. Political psychologist Morton Deutsch (1983) claims that the major social value of his discipline "...resides not in specific recommendations but rather in its providing organizing frameworks, clarifying ideas and systematic concepts for helping those who are engaged in practical political work..." (p. 224); he later elaborates on the specific "...social need for better ways of managing conflict," and refers to the "... relatively small number of social scientists ...[who]... have been working to provide the knowledge that will lead to more constructive conflict resolution." (Deutsch & Shichman, 1986, p. 219). As well, there have been recent calls for behavioural scientists to directly advise foreign policy makers (Kahan et al., 1983; Tetlock, 1986; Swingle, 1986).

Not all researchers see the state as the most effective or appropriate target for their expertise. In particular, some peace researchers have consistently noted that the "knowledge" they provide threatens the *raison d'être* of the same state apparatuses that almost all other academic communities serve (see Kelman, 1981). Thus, in Anatol Rapoport's words,

It is precisely the conditions that make a 'rational' prevention of war impossible that also make the creation of a supranational rational agency impossible. If such an agency could be created, there would be no need for it. (quoted in Eckhardt, 1983, p. 236)

To pursue this argument would anticipate some of the issues addressed below, but it appears that virtually all behavioural scientific research that contributes to deterrence theory follows the orthodox view that expert knowledge is to be acquired and disseminated to those most able to use its power, and this is usually the state.

Epistemology

Given the ends of instrumental knowledge, the behavioural scientific consensus on means is a method premised on an epistemology ostensibly drawn from the natural sciences, namely the empiricism emulated by Galileo and Newton, and defended *post hoc* by positivist philosophy.

Gergen (1986) provides an excellent summary of the empiricist's vision of knowledge, worth quoting at length:

At the outset one commences with the assumption of a physical world independent of the knower. The acquisition of knowledge first requires the observer's careful sensitivity to this world (recording the contours of the world as sense data), and the preliminary formulation of hypotheses. These hypotheses then become the basis for logical deductions, which are tested more systematically against new observations. The new observations, in turn, inform the observer of the validity of his/her hypotheses, and enable alterations to be made in such a way that the theory becomes an increasingly precise map of the physical world. Notice here that we have a finely honed narrative in which there is a causal arrow released at the beginning (with the material world), and which, through a series of repercussions (including induction,

deduction, test revision, and so on), ultimately finds its target in the form of an enhanced state of knowledge. (p. 4)

While many social psychologists and all nonjustificationist philosophers would see this version as a straw man, it is the only position that can legitimate the orthodox method of behavioural science.

This method gives pre-eminence to vulnerable theories (linguistic or mathematical systems) as the day-to-day repository of cumulative knowledge about the "real world." Theories consist of hypothetical propositions concerning operationalized concepts, with rules of correspondence between signifiers and signified, so that counter-factual evidence (however defined) can refute the theory. The "knower" ultimately seeks universal and incorrigible covering laws governing the relationships among empirical phenomena, so that particular events in the world can be explained and perhaps predicted via deductive-nomological inferences of causation.

This behavioural science appropriation from the natural sciences seems ironic for two reasons. First, "rigorous" social inquiry clearly predates success in discovering the "laws" that govern the physical universe; only the apparent success of the natural sciences (via technology) can explain its credibility as a new method for a much older project. Second, Rorty (1982) has observed that the main reason for philosophy's obsession with the vocabulary of

the scientific method was for the very fact that it was devoid of moral significance and human interest: "So, they said, the more metaphysically comfortless and morally insignificant our vocabulary, the likelier we are to be 'in touch with reality'..." (p. 193). Nevertheless, the empiricist epistemology and its consequent method are clearly hegemonic in the behavioural sciences, as expressed, for example, in M. Deutsch's (1983) loyalty to Nagel and Waltz's (1979) chapter on theory.

Method

The implications of this epistemology frame the never ending intra- and inter-disciplinary debates over methodology. In psychology, the issues focus on "the experiment" -- how to properly use statistics and experimental design to ensure internal and external validity and reliability -- as the critical crucible of reality. In the behavioural study of international relations, an assortment of tools -- content analysis, social indicators, mathematical models, computer simulations -- are continuously defended and criticized. And the interdiscipline of political psychology is said to practice an even more eclectic "methodological opportunism" (M. Deutsch, 1983, p. 224).

But perhaps the best illustration of the behavioural science adherence to the scientific method is the

discussion of the level-of-analysis problem in the study of international relations. Waltz (1959), in discussing his "three images" of the causes of war (denoting three levels of explanation) implies that each image must be evaluated on its predictive and prescriptive capacity. Jervis (1976) follows the Waltz argument with a more explicit criterion: the level of analysis chosen "is not arbitrary and is not a matter of taste -- it is the product of beliefs (or often hunches) about the nature of the variables that influence the phenomena that concern one." (1976, p. 15; see also Jervis, 1986; and Larson, 1985) The issue, then, is over which "dependent variable" (person, group, state) accounts for the most "variance" in observation, or else which "independent variable" can be more readily manipulated in practice. What is most curious is that these authors all recognize the normative antecedents and implications of the issue. Not only is the choice of analytic level the "product of beliefs" but it also has "...important moral implications. When all people would respond the same way to a given situation, it is hard to praise or blame the decision maker." (Jervis, 1976, p. 15)

As I will show in the second chapter, deterrence theory, while premised from its inception on a very narrow interpretation of mandate and epistemology, has taken much longer to address these discussions over methodology. Yet to the degree that it has incorporated advanced

methodological expertise, the apparent methodological successes of the behavioural sciences have been partially eclipsed by emerging doubts regarding its long-term viability.

RECENT CHALLENGES TO THE ORTHODOXY

Since the early 1970s, reflection on the part of a few insiders has led to a picture of the behavioural sciences as incompetent, disunited, or both. For the social psychologist Shaver (quoted in Kiesler & Lucke, 1976, p. 141), the crisis in his discipline has been

...that despite the development of an impressive methodology, many American social psychologists felt they had neither contributed substantially to the improvement of social life nor succeeded in developing a coherent scientific discipline. (see Strickland, et al., 1976)

In the whole field of international relations, an "inter-paradigm debate" has been diagnosed, with no consensus on an "underlying view of the world" (Banks, 1985), and an unhealthy parochialism on the part of the "behavioural paradigm" (Alker and Biersteker, 1984). The interdisciplines discussed above have tended to ignore these events, but in particular, the product of strategic studies -- deterrence theory -- has appeared to prosper as "probably the most influential school of thought in the American study of international relations." (Jervis, 1979) This prosperity, I will argue, has been due in large part to its paradigmatic isolation, and its consequent success

in avoiding the challenges facing the whole of the behavioural sciences.

While these recent challenges have sometimes been dismissed as mere methodological bickering, a consensus has emerged among some revisionists that the problem may lie with the commitment to an impoverished epistemology and its consequently impoverished method. A singular event in this realm was the behavioural scientists' reception, both hostile and sympathetic, to a conceptual study on the history of science.

A Proliferation of Paradigms¹

Since Popper, of course, the orthodox epistemology of positivism has been challenged by an ongoing debate between positivist and nonjustificationist philosophers over whether knowledge is passively gathered from an objective "real world," or actively constructed within an inter-subjective social world. But Kuhn's (1970 [1962]) attack on the "textbook" positivist portrait of evolutionary cumulative science highlights the social-psychologically necessary but historically contingent adherence to shared beliefs. Science progresses, Kuhn argues, not via small data-driven discoveries, but rather, with Gestalt-like shifts in the researchers' cognitive perspective and consequent language. Research is "...a strenuous and devoted attempt to force

nature into the conceptual boxes supplied by professional education." (p. 5) Hence, as many others have suggested, all observation is theory-impregnated; psychologically speaking, there can be no intelligible sensations which are not conditioned by cognitive expectations. Thus the puzzle-solving of "normal science" can only occur within an accepted paradigm, which

provides a map whose details are elucidated by mature scientific research. And since nature is too complex and varied to be explored at random, that map is as essential as observation and experiment to science's continuing development." (p. 109)

These paradigms are usually specific exemplary theories yielding a concrete historical discovery, from which methodological rules are intuitively defined. In what Kuhn called a "crisis" stage, a paradigm is destabilized by the emergence of persistent anomalies in observation, and several competing theories may come to exist simultaneously. These theories are incommensurable, and with a revolution from one paradigm to another, the choice of a new alternative exemplary theory cannot rest solely on rational, logical criteria.

While Kuhn's account has been widely criticized and adjusted, the signal interpretation -- that science does not and cannot rely on a privileged "objective" knower using a wholly rational, "objective" method -- seemed to strike a sensitive nerve throughout all of social science. Was political science, for example, in a Kuhnian crisis

where, as some suggested, the behavioural "program" would win-out over other methodological commitments? (Ball, 1976) Was the lack of consensus in the study of international relations due to competing incommensurable paradigms, with paradigmatic partisans constantly talking past each other? (Banks, 1984, 1985) Was the "crisis" in social psychology justifiable as a legitimate stage in any "mature" scientific discipline? (Kiesler & Lucke, 1976)

One of the more reasoned responses to these questions is Ball's (1976) attempt to shift the focus away from the search for social scientific paradigms, to an investigation of how paradigms endure. He draws on Lakatos' account of how methodological rules themselves form "research programs," which endure in science, despite the supposed instability of refutable theories: each program is built on a "hard core" of assumptions that are not directly refutable, and a "protective belt" of auxiliary models and assumptions are accumulated around this core. Successful programs are "good swimmers" because the protective belt insulates the crucial grounding assumptions which, like Kuhn's paradigmatic theories, are necessary for normal science. But Ball suggests that in any social science, this hard core derives from some "model of man" upon which research questions can be formulated and investigated. He posits the "rational actor" as one such model:

The rational choice program's model of man -- which may be traced back to Hobbes -- views men as rational

self-interested calculators. Men are assumed to be rational, and rationality is in turn defined in instrumental terms: to be rational is to choose that course of action...which will be most efficient in satisfying one's own ordered preferences. The objection that these basic assumptions are themselves 'unrealistic' or are 'refuted' by the 'facts' of human behavior in the 'real world,' carry little weight with contemporary 'positivist' theorists...(167)

The Fact-Value Fallacy

Ball implies that in the social sciences, the hard core of any research program is normatively² derived -- that a cognitive perspective on the social world is informed by beliefs, values, interests, ethics -- and this seems to me to be the fundamental social-psychological moral of Kuhn's and Lakatos's accounts. Thus we should not be surprised to find that much of the behavioural international relations theory had been "coloured" by the traditional assumptions of Morgenthau (Vasquez, 1975); in fact, the traditionalist/behaviouralist fight over methodology (cf. Knorr & Rosenau, 1969) may mask the more important consensus on world view. Similarly, we should not be surprised that the priorities of cognitive psychology are informed by a specific set of socio-political values and interests (Sampson, 1981). Nor should we be surprised that peace research is criticized for being both "too normative" and "too scientific" (Kelman, 1981). We should not be surprised because facts do not give birth to theories any more than counter-facts slay them.

Thus, in Banks (1985) words,

It is naive and superficial to try to discuss IR solely on the basis of 'the facts.' This is because whatever facts are selected -- any at all -- are literally abstract. They are chosen from a much bigger menu of available facts, because they are important. The question is: why are they important? And the answer to that is: because they fit a concept, the concept fits a theory and the theory fits an underlying view of the world. (p. 7)

And in Gergen's (1982) words,

Socioeconomic classes, ideologies, profit, loss, deviance, and so on do not exist "out there," waiting for discovery like some sequestered planet. Rather, they are made distinguishable by scientists motivated by specific tastes, values, needs, or motives...[And] ...in the designation of "what there is," the terms of description, the explanation for activity, and the process of gaining knowledge itself, scientists generate statements of the good, the proper, or the desirable in social life... In effect, one may properly replace the traditional belief in the separation of fact and value with the apothegm: no description without prescription. (pp. 28, 33-34)

Theory as Rhetoric

This relativism of theory and theoretical language was implied in Kuhn's often misunderstood notion of incommensurability, with respect to which he has more recently elaborated:

In applying the term 'incommensurability' to theories, I had intended only to insist that there was no common language within which both could be fully expressed and which could therefore be used in a point-by-point comparison between them." (quoted in Brown, 1983, p. 3)

As Rorty (1982) points out, the success of Galileo et al. did not

point to any epistemological moral... Kuhn and Dewey

suggest we give up the notion of science traveling towards some end called "correspondence with reality" and instead say merely that a given vocabulary works better than another for a given purpose... Galileo's terminology was the *only* "secret" he had -- he didn't pick that terminology because it was "clear" or "natural" or "simple," or in line with the categories of the pure understanding. He just lucked out. (pp. 191, 192-93)

Gergen has extensively discussed the implications of this revisionist view of theoretical language for social inquiry. Such an inquiry, he argues, must confront impermanence. Rather than trying to "factor noise out" of the apparent flux of the social world, to uncover universal and trans-historical laws, social inquiry should welcome its fundamentally historical-contingent nature, and recognize that all forms of representation, no matter how seemingly abstract, are nevertheless steeped in the theorist's social and political milieu. To have an absolute consensus on denotative definitions for terms, there must exist, at the least, "...*continuing or recurring properties* available to public scrutiny." (p. 81) Such stability does not actually exist in the social world. In the case of psychology, for example, the virtually infinite and particularistic varieties of human action force the theorist (as with anyone else) to employ "motive" language, describing human actions with respect to inferred ends. Such terminology, defined in terms of the end points of action, "...appear[s] not only to describe, but simultaneously to explain." (p. 90) It is for this reason,

that psychological theory does not

rest on unambiguous evidential grounds; rather, when empirical support has garnered a theory, it can remain "support" only as long as there is a willing cartel to support a line of interpretation... The theory does not in this case furnish new knowledge; it primarily furnishes intelligibility for what is already known. [Thus]...research must be viewed primarily as a rhetorical implement. Its chief function is to lend power (persuasive impact, appeal, felicity) to the theoretical language. (p. 77)

Bias Toward the Status Quo

In the case of informing social policy, this receptivity will require, further, that the theoretical knowledge be communicated *in the evaluative terms set by the policy-maker*:

Suppose we picture the "value-free" social scientist walking up to the divide between "fact" and "value" and handing his predictions to the policy-makers who live on the other side. They will not be of much use unless they contain some of the terms which the policy-makers use among themselves. (Rorty, 1982, p. 196)

Furthermore, these terms often prove to be self-reproducing. Once the privileged subject communicates the newly manufactured knowledge, the object of inquiry has been potentially "contaminated" with whatever enlightenment this knowledge affords (Gergen, 1982, p. 22-26). Thus, just as the distinction between fact and value is dissolved, so too is the distinction between the privileged subject and an inert object of inquiry, implying that theoretical logics may become reified in actual social practice.

But herein lies the final challenge to behavioural science. For if successful theory requires a receptive audience, there will always be a conservative bias in favour of those theories which are consistent with orthodoxy or the received view. *The rhetorical power afforded the behavioural scientist appears to be the power to reproduce and sustain the social and political status quo rather than to offer emancipation from it.* Thus, Morawski & Goldstein (1985) find that psychological research on issues of nuclear war consistently reflected the political climate of the times. Similarly, Banks (1984) finds that the evolution of international relations theory has had a symbiotic relationship with trends in foreign policy. And finally, Apfelbaum and Lubek (1976) can make a persuasive, if over-stated, case that the priorities of most conflict analysis (and certainly deterrence theory) reflect a Cold War pre-occupation with symmetrical, manageable conflict resolution over the legitimacy of conflict as a force for social change (see also, Kelman, 1981).

TOWARD EMANCIPATORY SOCIAL THEORY

M. Brewster Smith (1983), in a personal account of postwar American social psychology, has

...yearn[ed] for a social psychology that bears usefully on human problems — not so much in the technological spirit of human engineering as in the emancipatory spirit (Habermas, 1971) of empowering

people to cope more competently.... the more I take advantage of my senior status to try to gain a philosophic perspective on our activities as psychologists, the more I am convinced that the positivist tradition in which I was raised and to which the mainstream of the field still adheres is essentially misguided. (p. 177)

To Smith, it is not the original mandate of the behavioural sciences that is at issue, but rather, the instrumental knowledge that has been projected as the best means available. As one of many alternatives, Gergen (1982) proposes criteria for a new "generative" theory, whose proponents would be less dependent on method, more sensitive to the historicity of their research, more concerned with the social negotiation required for communal knowledge, and more self-conscious of (and content with) their non-neutral political role in transforming society -- in brief, to answer the challenges discussed above not by interpreting merely what is, but by suggesting what *might* be. (See also Sampson, 1981.) Such an emancipatory theory would downplay the utility of instrumental knowledge, provide support for ethical debate, and affirm what Rorty (1982) has claimed is the real aim of all social science:

This is the attempt to free mankind from Nietzsche's "longest lie," the notion that outside the haphazard and perilous experiments we perform there lies something (God, Science, Knowledge, Rationality, or Truth) which will, if only we perform the correct rituals, step in to save us. (p. 208)

If not the "longest lie," deterrence theory has certainly been a lie of sorts. As a research program, strategic studies has continued to flourish in paradigmatic

isolation, declaring Olympian value-neutrality, and providing formidable rhetorical support for an international security status quo. Any inquiry into deterrence, in order to serve the true mandate of social science, must at the very least complement the strategists' search for instrumental knowledge with some form of emancipatory theory. This is my purpose in the third chapter. But first, within the context outlined in this chapter, I will suggest all the reasons why an alternative deterrence theory cannot come too soon.

CHAPTER 2: Psychology and Deterrence

Two Ambiguities of This Century

Since the seventeenth-century Peace of Westphalia, the hegemonic normative global order has been a European invention -- with its zenith in the nineteenth century -- of independent, (potentially) equal, impermeable, sovereign nation-states, who interact from time to time in an anarchical environment. States within the system are afforded the exclusive right to govern their own citizens, and to resort to force in settling conflicts with other states. Whether this conception of global order has ever been an accurate account of reality, its predominant influence on the thoughts and actions of those in service to states continues to the present. But as an appropriate ideal for managing the interests and security of these states and their citizenry, Westphalia's viability has been challenged by new circumstances. Examples would include the end of the European multipolar consensus and leadership over the globe (signalled by World War I), the emergence of a bipolar hegemony of two geopolitically and ideologically opposed superpowers (signalled by the end of World War II), the spread of nation-state status to many former European colonies with Southern agendas very often in conflict with the apparent interests of the North (signalled by the growth of the United Nations), and the emergence of viable

non-state actors (e.g., multinational corporations, the Palestine Liberation Organization) who affect both the economic and security interests of nation-states.

Of relevance to this discussion, however, has been the role of 20th century technology in transforming the nature, and hence the *raison d'être*, of military forces. We live in a world of unprecedented accumulation and frequent use of sophisticated military force to resolve human conflict. Nevertheless, the utility of this accumulation and use has become most ambiguous. Westphalia's requisite link, between national security and military power, between sovereignty and war, no longer obtains, particularly for the North. A disarmed loser of the last world war is now among the world's most prosperous states, and two heavily armed winners have become muscle-bound superpowers unable to defend against each other's nuclear arsenals or successfully resort to conventional force in the service of other foreign policies (e.g., Vietnam and Afghanistan).

Coincident with a new ambiguity of military power has been the varied and often ambiguously defined roles of Western intellectuals in responding to problems of international security. No longer are questions of global politics the exclusive domain of diplomatic and military historians (the Westphalian vanguard), nor of normative political theory. Journalists, social scientists, academic and military strategists, political philosophers,

statesmen, and lay-persons all vie for participation in discourse about global politics. So the dissolution of the political consensus that sustained the Westphalian order until this century has been matched by a dissolution of the intellectual consensus on what global politics is and ought to be about.

The Western discourse on international security is no exception. National security policy is almost exclusively informed by the discipline of strategic studies, the domain of government-funded think-tanks as well as many political science curricula. Deterrence theory is the single most influential intellectual product of this discipline. But, other social sciences have addressed questions of international security, and, as discussed in the first chapter, these disciplines draw on different epistemological and normative traditions. Some of this work has influenced the policy debates within governments, and the priorities of strategic studies; and some theorists have at one time or another been involved in policy formulation. But it is often unclear what particular interests are being served by the production of theory (cf. Chomsky, 1982). The discussion below will focus primarily on the theory of deterrence as it has evolved from the theoretical work both within and outside direct state sponsorship.

Something less than a full historical account, the

following discussion might be seen as a selective genealogy of strategic studies. It will attempt to highlight various discourses that have contributed to the production of deterrence theory, within the context of the previous chapter's critique. Specifically, what have been the various methodologies appealed to in providing a protective belt for the research program of deterrence theory, and what are the requisite normative commitments which inform its *hard core*? How have critics challenged deterrence theory, and how have they failed in challenging its hegemony as a "good swimmer"? Finally, how has the obsession with a behavioural deterrence theory unwittingly vitiated attempts to challenge the logic and consequent dangers of a deterrence status quo?

THE EMERGENCE OF DETERRENCE THEORY

Defining Deterrence

As discussed earlier, social signifiers cannot be defined unequivocally, because their intelligibility derives from the contingent social context in which they are used. Deterrence, therefore, is a meaningless proposition until people use it, along with a variety of other concepts, to explain the past, prescribe the present, or predict the future. Even if we agree that deterrence refers to behaviour modification via threat, it is unclear whether such an outcome is due to fear, rational

calculation, misperception and so on, nor is it clear whether the behaviour of the proverbial misbehaving child is isomorphic with that of potential criminals, state bureaucrats and militia, or even collective state actors. The assumption of isomorphism, however, has been assumed by many theorists (as with other terms like "aggression" and "cooperation"), and this helps to explain the continuity between "deterrence" and "deterrence theory."

The widespread currency of the deterrence concept is illustrated by the conspicuous absence of the term in Bernard Brodie's frequently cited mandate for an American nuclear-armed military:

The first and most vital step in any American security program for the age of atomic bombs is to take measures to guarantee to ourselves in case of attack the possibility of retaliation in kind... Thus far the chief purpose of our military establishment has been to win wars. From now on its chief purpose must be to avert them. It can have almost no other useful purpose. (Brodie, 1946, p. 76)

Brodie had been one of the first to account for the revolutionary implications of nuclear weapons for national security: that with the inevitable end to an American nuclear monopoly, the final success of offence over defence would bring an unprecedented challenge to sovereignty, and "that while the idea of deterrence per se was nothing new..., what was distinctively new was the degree to which it was intolerable that it should fail." (Brodie, 1978, p. 66)

This was a Westphalian response to the nuclear age, and

it ushered in what Jervis (1979) has called the "first wave" of deterrence theory.³ The suggestion that deterrence was "nothing new" reflects the traditional *a priori* assumptions of a Hobbesian international state of war-prone anarchy and potential aggression, and a primary role for the military in securing the state against such aggression. But the nature of this security was certainly new: no longer could there be a healthy Clausewitzian role for the actual use of military force. Brodie's formulation was really not a theory at all, but rather a proto-theory, implying a much more amorphous Realist⁴ theory of international relations. (In fact, Brodie's writing remained ambivalent toward the Clausewitzian view; compare Brodie, 1946, 1959, 1978.) His rather benign proto-theory, unhampered by subsequent elaboration and obfuscation (Bundy, 1983, calls it "existential deterrence"; Morgan, 1977, calls it "general deterrence") suffered three mortal weaknesses: ambiguity, internal contradiction, and limited scope of reference.

First, its ambiguity is illustrated by a dilemma that has become a major problem for all deterrence theories -- namely, whether deterrence rests on rational or non-rational elements of human behaviour: Does the threatened retaliatory strike represent an attempt to rationally manipulate the adversaries rational calculation of costs and risks (via denial or punishment) or is it an

"existential" fear that retaliation might occur for *non-rational* reasons such as revenge or reflex action? Is the mechanism of manipulation one of *rational* self-interest maximization or the sponsorship of *non-rational* terror? Does the aggression to be deterred derive from an evil (but *rational*) intention of a revolutionary adversary or from systemically conditioned misperception as a consequence of states in anarchy? Leaving aside the problems in defining rationality (see below), this ambiguity clearly causes great difficulty for the formulation.

Second, the theory does not address the other side of the equation. Jonathan Schell's (1982) articulation deserves quotation at length:

In pre-nuclear military strategy, the deterrent effect of force was a useful by-product of the ability and willingness to wage and win wars. Deterrence was the shadow cast by force, or, in Clausewitz's metaphor, the credit that flowed from the ability to make the cash payment of the favorable decision by arms. The logic of pre-nuclear deterrence escaped circularity by each side's being frankly ready to wage war and try for victory if deterrence failed. Nuclear deterrence, however, supposedly aims solely at forestalling any use of force by either side, and has given up at the outset on a favorable decision by arms. The question, then, is: Of what object is nuclear deterrence the shadow? Of what cash payment is it the credit? The theoretical answer, of course, is: The retaliatory strike. Yet since in nuclear-deterrence theory the whole purpose of having a retaliatory capacity is to deter a first strike, one must ask what reason would remain to launch the retaliation once the first strike had actually arrived... In sum, the doctrine is based on a monumental logical mistake: one cannot credibly deter a first strike with a second strike whose *raison d'être* dissolves the moment the first strike arrives. (p. 202)

Finally, if a strategic nuclear first strike was not the

only threat to national security (as all strategists maintained) then how would a nuclear-armed militia deter or otherwise respond to lesser contingencies -- such as an attack on other intrinsic interests like Western Europe, or attacks on strategic interests which threaten the geopolitical balance of power? When the adversary is also armed with nuclear weapons, one's own use of them in response to these other aggressions might bring on the very national suicide that is to be avoided at all costs. This theory, then, provided neither an intellectually credible rationale for a strategic retaliatory capability, nor a pragmatically credible mandate for deterring lesser threats. This absence of credibility manifested itself politically in the controversial 1950s U.S. policy of "massive retaliation." The signal intellectual response was Henry Kissinger's (1957) book on *Nuclear Weapons and Foreign Policy*.

Defining Deterrence Theory

As a student of nineteenth century diplomatic history, Kissinger was understandably pre-occupied with the successful pursuit of American foreign policy; the existence of nuclear weapons, in his view, did not (or at least could not) vitiate a continued pursuit of national interests, such as extended deterrence to protect European allies. Because the revolutionary Sino-Soviet bloc did not

recognize the Western consensus on international relations, diplomacy could not avoid war, but merely mitigate its consequences. The only agreement between East and West was the desire to avoid all-out nuclear war, but this made lesser aggression more likely. Thus, given his Realist theory of international relations, Kissinger attempted to reclaim a Clausewitzian role for military forces, to deter piecemeal aggression. The retaliatory capability was still necessary to deter all-out war, but since this was the least credible threat, a further capability for deterring and perhaps fighting "limited war," or even "limited nuclear war," was also necessary.

In legitimizing this policy (which indeed foreshadowed actual American and NATO strategy), Kissinger elaborated a complex psychological theory of deterrence. "Massive retaliation" had "...given the Soviet leadership an opportunity to strive to neutralize us psychologically" because such a threat was not credible. "The United States must study the psychology of its opponents as carefully as they have studied ours... in the relation among states, will may play as great a role as power." He dwelled on the Soviet fear and insecurity deriving from their revolutionary status, and the "fearful psychological handicap" of a U.S. reliance on a strategic deterrent. He painted a clear picture of a relentlessly aggressive Soviet adversary using protracted psycho-political means to take

advantage of a defensive Western alliance. And he praised the Soviets for their accurate perception of Western vulnerability to intimidation, while he condemned Western misperception in ascribing legitimate motives to revolutionary powers. The only appropriate response would be to define a set of limited diplomatic-military strategies, keeping in mind that the purpose "... is to affect the will of the enemy, not to destroy him, and that war can be limited only by presenting the enemy with an unfavorable calculus of risks."

Kissinger's theory of deterrence attempted to address the credibility problem by emphasizing the psychological nature of the Cold War. All of his propositions about this nature were based on vague hypotheses about human (and specifically Soviet) behaviour and intentions, without external validity nor even internal consistency. What is important, however, is that his psychological theory of deterrence could not be distinguished from his normative theory of international relations. His observations of perception and intention, aggression and insecurity, will and resolve, appeared to simultaneously describe and prescribe state behaviour, and his theory of deterrence was only relevant within a Cold War environment of inevitable armed conflict. Thus, in the language used in the first chapter, the *hard core* of his deterrence theory derived from a normative, Realist theory of international

relations.

Kissinger's last chapter chided the American security apparatus, and all of American culture, for its "empiricist" isolation from the cunning, dangerous and sometimes tragic world of great power conflict. It is ironic, then, that his book preceded an ostensibly empiricist theory of deterrence concerned with analyzing and solving the same psychological problems of influence, credibility and resolve which he had diagnosed. In Brodie's reflection on the "Development of Nuclear Strategy," he claims that

Virtually all the basic ideas and philosophies about nuclear weapons and their use have been generated by civilians working quite independently of the military... Whatever the reasons, they must include prominently the fact that to the military man deterrence comes as the by-product, not the central theme, of his strategic structure. Any philosophy which puts it at the heart of the matter must be uncongenial to him. (1978, p. 67)

So, deterrence would no longer be the shadow of military use, but rather, the shadow of a new "second wave" deterrence theory, to be manifest in the new academic discipline of strategic studies, with methodological support from the behavioural sciences. Such a theory would be sensitive to the intricate psychological aspects of diplomatic influence necessary to pursue national interests in a world of unmitigated and often latent violence. Much of this discipline's product was Kuhnian puzzle solving -- systems analysis and operational research to direct weapons

procurement policy and civil defence (e.g., Kahn, 1960, 1962; Wohlstetter, 1959) -- but the paradigm most often exploited was the behavioural theory of rational decision.

THE RATIONAL ACTOR

The Second Wave

The rational paradigm⁵ actually had its roots in much of the classical Realist theory. Such accounts often provided a rational historical account, explaining state behaviour by appealing to an often tautological definition and description of power politics. But the second wave, appealed to a new positive, axiomatic definition of rationality in hopes of escaping this tautology.

George and Smoke (1974), in one of many revisionist accounts of deterrence theory (see below), suggest that because of American politics, culture, and geography, the main focus of deterrence theory has been on the more likely, but also more problematical, threats below the strategic level:

The problem for American foreign policy has been that of "projecting" deterrence beyond her own borders to cover other parts of the world, seeking to protect some third party -- an ally, or a neutral we wish to safeguard. Accordingly, theory about the deterrence of limited threats to the U.S. has focused almost exclusively upon a "three-nation problem"... A relatively simple set of ideas about this problem has emerged which without much oversimplification could be presented in the form of a simple model embracing the following propositions:

(1) The principal communist powers perceive an interest in attacking or encroaching upon various nations or regions which the U.S. has an interest in

preserving self-determination.

(2) Such attacks can be deterred if these powers calculate that the expected costs and risks (negative utilities) of attacking outweigh the expected benefits (positive utilities):

$$C + R > B$$

(3) Such expectations will be held if these powers believe that the United States will defend the attacked nation. More precisely, the expectations will result if the estimated probabilities of U.S. defense, multiplied by the costs and risks that defense would impose, exceeds the expected benefits that would result in the event of no U.S. defense:

$$p(C + R) > (1 - p)(B)$$

(4) Generating this belief -- or more exactly, raising this estimated probability -- is therefore the object of U.S. deterrence policy.

(5) The requirements for generating this belief are that the principal communist powers perceive a U.S. "signal" of intent to defend and find the signal "credible."

(6) The signal requires appropriate declaratory and action moves by the United States, communicating intent.

(7) Credibility requires that the U.S. appear to become committed to carrying out its intention to defend the attacked nation. (pp. 59-60)

George and Smoke suggest that it is the last three propositions that have pre-occupied most of the original theory, and game theory, while not the only possible exposition of the rational decision paradigm, has been appealed to by both proponents and critics of deterrence theory as best illustrating the major axiomatic propositions and theoretical implications of "signalling" and "commitment". (See Schelling, 1960, vs. Rapoport, 1964.)

Game Theory as Rational Analysis

Game theory posits two or more unitary, self-serving,

rational players, pursuing partially or totally incompatible goals, with perfect information as to outcomes (or probabilities of payoffs) for certain choices. Rationality is defined axiomatically: players will always choose according to their own best interests, as dictated by the potential payoffs.

The simplest phylum of game theory is a two-player, one-move, zero-sum game with a saddle point. In such a game, one player's loss is exactly the other's gain, and vice versa. Matrix 1, below, shows the payoffs for two players, A and B, each with two choices, C and D; the payoff for the row player, A, is given first in each box of the matrix. Each player's payoff is partly dependent on the other's choice. The game is said to be stable because there is a best strategy for each player. Player B, being rational, will choose D because he is better off regardless of player A's choice. Player B is said to have a "dominant" strategy. Player A, knowing B's dominant strategy, will choose D as well. Thus DD is the saddle point or "minimax" solution to the game. The game, therefore, is prescriptive as well as deductive in that it prescribes an optimal strategy for each player.

Another slightly more complex game is similar but with no saddle point, illustrated in Matrix 2, below. This game is said to be unstable because now player B's best choice is not necessarily D; rather, it depends on player A's

Figure 1: Matrixes for two-player ("A" and "B") zero-sum and non-zero-sum games

Matrix 1

		"B"	
		C	D
"A"	C	2, -2	1, -1
	D	3, -3	-2, 2

Matrix 2

		"B"	
		C	D
"A"	C	1, -1	2, -2
	D	3, -3	-2, 2

Matrix 3

		"B"	
		C	D
"A"	C	R, R	S, T
	D	T, S	P, P

C = Cooperate
 D = Defect
 R = CC = Reward
 S = CD = Sucker
 T = DC = Temptation
 P = DD = Punishment

Prisoner's Dilemma:
 DC > CC > DD > CD

Chicken:
 DC > CC > CD > DD

choice in the move. And now player A, not being able to count on player B's choice, cannot automatically decide his choice either. So, while the logic of possible outcomes for each player can be deduced, the game offers no prescription for an optimum, as is. What is needed to save the foundations of game theory is a new "mixed strategy" of choosing randomly in a specific frequency ratio across many plays of the game, so that eventually the payoffs for each player would accrue around some "expected" maximum (Rapoport, 1960, p. 155). But, such a solution rests on a new concept of rationality, premised on expected (probable) gain across many moves, rather than on certain gain in a single move. In 1965, Rapoport and Chammah claimed that

The two-person zero-sum game without a saddle point constitutes in our opinion the limit of applicability of game theory as a normative (or prescriptive) theory. That is to say, the minimax mixed strategy can be both rationally defended and intuitively accepted as a "best course of action" in a certain sense. Once the limits of two-person zero-sum games are transcended, game theory, while remaining a powerful tool for analyzing the logical structure of conflicts of interest, loses its prescriptive power. In this realm, strategically rationalizable courses of action are frequently intuitively unacceptable and vice versa. (p. 23).

Yet it is in "this realm" -- specifically, non-zero-sum games -- that social scientists have focussed their attention. In such "mixed-motive" games, one player's loss is not necessarily the other's gain; in fact, there are outcomes that can realize gains or losses for both. Many conflict theorists argue that it is these games which better model many social situations, including

international relations. In making this argument for international relations theory, normative assumptions are made: partially independent, instrumentally rational states exist in conflict. These theorists take seriously the limits of this heuristic, in describing real global politics, but they believe that "...the assumption of rational behaviour is a productive one. It gives a grip on the subject that is peculiarly conducive to the production of theory." (Schelling, 1960, p. 4)

Game Theory and Strategy

Without a doubt, Thomas Schelling's (1960, 1966) work is the most important initial attempt to construct a body of strategic deterrence theory around the concepts of game theory. His work essentially provided a protective belt around the contradictory and Clausewitzian-mitigating hard core of deterrence. Game theory added a sense of rigour to earlier natural language formulations and added the credibility of "neutral" behavioural science to the consciously normative theories of Kissinger and others.

The two particular phyla of non-zero-sum games that Schelling and others used extensively were "Prisoner's Dilemma" and "Chicken." As in the previous games, the choice of each player is independent of the other, but the outcomes for each are now interdependent. Moreover, there is no central authority to enforce threats or to

communicate in advance the choice of either player to her opponent. As with all phyla, these games are defined by the relative values of each outcome. Both games posit a choice between cooperation (C) and defection (D), as illustrated in Matrix 3, above. If both players cooperate, each is rewarded (R). If one player cooperates and the other defects, the former loses and gets the sucker payoff (S) while the latter gets the best payoff by acting on his temptation to exploit (T). If both players defect, then each is punished (P).

In the Prisoner's Dilemma, the best payoff possible comes with the temptation to defect (DC); the second best payoff comes from mutual cooperation (CC); worse than either of these is the punishment each player receives for mutual defection (DD); but worst of all is the sucker payoff (CD). The dominant strategy for both players is defection since it results in a payoff as good as or better than the opponent, regardless of his choice. If, however, each player chooses the dominant strategy, both players are worse off (through mutual punishment) than if they had cooperated. At the heart of the problem is the inability for the players to coordinate their choices and cooperate. Therefore, what is rationally prescribed for each player turns out to be irrational for the two of them. This is the dilemma. In the context of deterrence theory, Schelling posited this dilemma as the "Reciprocal Fear of

Surprise Attack": the fear that an adversary intends to launch a surprising and disarming attack (sucker payoff) tempts each player to pre-empt (mutual punishment). His conclusion was that such a situation must be avoided; hence each side should deploy, and encourage the other side to deploy, invulnerable second strike counter-value weapons:

A weapon that can hurt only people, and cannot possibly damage the other side's striking force, is profoundly defensive: it provides its possessor no incentive to strike first." (1960, p. 233)

This was the essence of Brodie's initial mandate, what came to be called "Mutual Assured Destruction" (MAD), and it represents the "strategy of stable conflict" upon which most arms control theory and policy has been premised.

In the game of Chicken, the relative payoffs are the same except that the outcomes for sucker and punishment are reversed in relative value (rank), with mutual defection being the worst payoff for each player ($DC > CC > CD > DD$). Hence, there is no longer any dominant strategy for either player, even though there are still incentives for both cooperation and defection. The larger the incentives to cooperate (to avoid mutual punishment), the larger also the temptation for each to defect and exploit the other's fear of being punished. The game is divinely unstable. In the context of deterrence, to cooperate by backing down from a confrontation seems the best course of action in light of the possible mutually devastating consequences of both parties standing firm. In such a case, joint survival is

an incentive for cooperation. But there are also Hobbesian "temptations" to, in Kissinger's words, "run risks" and competitively exploit the other player's fear of punishment to realize both national ends and international avoidance of all-out war. The player that makes the concession in the face of this intimidation is called "chicken," hence the name of the game.

How are these concessions induced? Schelling's solution involved *credible* deterrent or compellent threats which commit one player to the choice of defection so as to intimidate the other player. The credibility of such threats is, of course, problematical in high conflict, where there is potential for mutual suicide, later if not sooner. There is, then, a "stability-instability" paradox:

If an uncontrolled war would lead to mutual destruction, then neither side should ever start one. But this very stability allows either side to use limited violence because the other's threat to respond by all-out retaliation cannot be very credible. (Jervis, 1984, p. 31)

Schelling's answer was to move beyond the objective "payoff" situation (which might be conceived as the "existential" deterrence relationship) and develop a series of tactics designed to broaden the concept of strategic rationality. For instance, a player could make a threat that "leaves something to chance." Here, although punishment is not certain, "if the threatened party fails to comply, the final decision is not altogether under the threatener's control." (1960, p. 188) These threats are

said to be more credible. An example of such a threat is the threat of inadvertent war, or the potential risk of limited war escalating out of control: "The supreme objective may not be to assure that it stays limited but rather to keep the risk of all-out war within moderate limits above zero." (1960, p. 193) Limited violence also allows for the special use of military force to "hurt" rather than destroy. (Schelling, 1966)

Another example of "manipulating the shared risk of war" is through brinksmanship, where the brink is not

...the sharp edge of a cliff where one can stand firmly, look down, and decide whether or not to plunge. The brink is a curved slope that one can stand on with some risk of slipping, the slope gets steeper and the risk of slipping greater as one moves toward the chasm. (Schelling, 1960, p. 199; see also Schelling, 1966, 99-105)

Finally, credibility may be enhanced by communicating tacitly, over several plays (e.g., several superpower confrontations), a reputation for standing firm, thus attempting to manipulate the adversary's perception of 1) one's consistently non-rational stubbornness, or 2) one's own naivete of the risks involved in the game. (Schelling, 1966, 116-125)

In both the Prisoner's Dilemma and Chicken, game theory cannot itself prescribe a rational and defensible strategy; they each present fundamental paradoxes in rational choice, given the axiomatic definition of rationality. Schelling's prescription for communicating defensive intent, by

avoiding destabilizing first-strike weapons and encouraging invulnerable retaliatory capabilities, emphasized the goal of avoiding Armageddon through the avoidance of situations whose logic resembled that of the Prisoner's Dilemma. And his tactics of fighting for less than intrinsic interests, destroying crisis communication links, feigning anger and irrationality, burning bridges, rocking the boat, using little black-boxed roulette wheels, and throwing the steering wheel out the window, focussed on the contradictory goal of pursuing national interests with theoretically extraneous mechanisms of commitment and communication to deal with the credibility problem. They were, in Schelling's own words, a "rigorous" re-interpretation of old devices where

...ancients exchanged hostages, drank wine from the same glass to demonstrate the absence of poison, met in public places to inhibit the massacre of one by the other, and even deliberately exchanging spies to facilitate transmittal of authentic information. (1960, p. 20)

While Schelling is only an exemplary, and certainly not exclusive, representative of the second wave, his central focus -- credible options for exploiting potential force -- has survived as the *raison d'être* of strategic studies, and behavioural scientists who accept the strategic world view have lent their intellectual legitimacy to the nuclear deterrence status quo. Even if prescriptions for commitment and threat have evolved over the years (with a shift in emphasis from MAD to "flexible response," and most

recently, to war-fighting strategies), they all derive from the strategists' professional concern with exploiting deterrence theory to pursue national ends other than avoiding Armageddon (see Freedman, 1981). Furthermore, this project is steeped in the Realist's normative priority of maintaining the decentralized Westphalian state system, regardless of the new destructive capacity of these states' military forces. It is, in Gwynne Dyer's words, an attempt at "keeping the old [Westphalian] game alive" (Dyer, 1985).

Strategists continue to except this second wave characterization of the game as an immutable given. To wit, Brams (1985) concludes that

There are good reasons why this [superpower] conflict, and the rancor and recriminations it has generated, will not simply disappear. Generally speaking, they are not rooted in inexorable historical forces, the obtuse choices of leaders, or rank misperceptions that the leaders harbor of each other. Rather, I believe, it is the unforgiving nature of the games the superpowers play that better explains their continuing conflict... (p. 145)

And if there is any doubt about the symbiosis, or at least benign complementarity, between strategic studies and the contemporary national security culture, U.S. Secretary of Defence Casper Weinberger illustrates the case in his account of present "U.S. Defense Strategy":

Our strategy is simple. We seek to prevent war by maintaining forces and demonstrating the determination to use them, if necessary, in ways that will persuade our adversaries that the cost of any attack on our vital interests will exceed the benefits they could hope to gain. The label for this strategy is

deterrence. It is the core of our defense strategy today, as it has been for most of the postwar period. Moreover, this strategy is working. It has succeeded in winning for the United States and our allies four decades of peace with our primary adversary -- a period nearly twice as long as the time-span between World War I and World War II.

To be effective, deterrence must continue to meet four tests:

Survivability: our forces must be able to survive a preemptive attack with sufficient retaliatory strength to threaten losses that outweigh gains. **Credibility:** our threatened response to an attack must be a form that the potential aggressor believes we could and would carry out. **Clarity:** the action to be deterred must be sufficiently clear to our adversaries so that the potential aggressor knows what is prohibited. **Safety:** the risk of failure through accident, unauthorized use, or miscalculation must be minimized. (Weinberger, 1986, pp. 676-677)

CRITIQUE OF THE RATIONAL ACTOR

Outside of strategic studies, social scientific criticism of the second wave has been ubiquitous (K. Deutsch, 1968; George & Smoke, 1974; Green, 1966; Jervis, 1979; Jervis, Lebow & Stein, 1985; Morgan, 1977; Rapoport, 1964; Snyder & Diesing, 1977). These criticisms have been directed at both the empirical and normative weaknesses of the second wave and the inadequacy of consequent policy. Most of these critics belong to the rather diverse interdiscipline of political psychology, and I will not detail the entire spectrum of their often persuasive revisions. My emphasis here is on the general consensus that the rational decision paradigm is inadequate and inappropriate, either for explaining the nature of deterrence, or for prescribing policy to bolster it. Such

a critique must begin with Rapoport's (1964) claim that strategists often fail to distinguish between three modes of analysis in the rational paradigm: formal, empirical, and normative.

The Political Irrelevance of Game Theory

The formal mode is purely deductive; it rests on a set of internally consistent axioms relating hypothetical sets of variables such as values to be sought or defended and situational constraints on alternative options. In game theory, the formal mode makes only one normative commitment, namely to the instrumentally rational maximization of self-interest. Rapoport (1964) and Green (1966) have elaborated on the ethical implications of such a discourse, but it is important to note that game theory's formal mode has no direct relevance or reference to the social world. Knowing about the logic of particular games can only be useful if there are rules of correspondence to analogous games or logics in the real world; and as with any social theory, these rules cannot derive from the theoretical axioms themselves, but rather, must require the social theorist to make some critical normative assumptions. It is the strategists' assumptions that are a main target of their critics.

First, assumptions are made about what might be called the "behavioural predispositions" of players. Whether this

is a political or psychological question depends on the level of analysis chosen (a decision that theorists rather than theory must make). The value-preferences necessary for game theory imply motives of the players, but the theory sheds no light on how to determine these motives, their causes, or their dynamic nature. In deterrence theory, adversaries are assumed to have aggressive intentions toward merely defensive, status quo deterrers, but no concern is paid to why these threatening intentions exist or how they can be changed. Thus, Green (1966, p. 110) claims that "...game theory tells us how to theorize about an extended game [of deterrence] but not how to recognize whether we are in one," and Jervis (1979, p. 292) claims that deterrence theory follows the Realist tradition of ignoring transformation of hostility, offering no explanation of "...how to change an adversary or to determine whether changes have taken place." (see also K. Deutsch, 1968) Such intentions are important to determining which game is being played: Jervis (1976) distinguishes between "deterrence" (Chicken) and "spiral" (Prisoner's Dilemma) models of superpower conflict solely on the basis of differing perceptions of Soviet intentions (aggressive or defensive), a decision rational analysis cannot make. Deterrence theory, then, forces real life scenarios of dynamic and complex motives into static games without any intellectual rationale for why the games must

be played. Rationale must come from a theorist's normative commitments.

The Psychological Irrelevance of Game Theory

Of course, most theorists recognize this limitation of the rational paradigm, even if they do not appreciate the implications for a normative theory (see below). Nevertheless, strategists still claim that, given a certain environment (e.g., hostile states in anarchy), game theory offers an explanation of the psychology of diplomatic-military decision-making. Game theorists operating in the empirical mode have conducted extensive experimental research on how actual people play games such as the Prisoner's Dilemma. The resulting theory, in Rapoport's (1964, p. 6) words

accomplishes its aims if it can say (and support the statement with empirical evidence) something like this: "People make decisions, as if they were guided by the following such and such decision rules..."

Any conclusions derived from these models allow for, but do not themselves articulate, psychological explanation. In Rapoport and Chammah's (1965) experiments using the Prisoner's Dilemma, they make this point clear. In constructing a theory of decision from the statistical regularities of subjects' responses, they seek to derive mathematical models of how players "learn" across extended plays, but

It is from the parameters of these models, not from any a priori assumptions about the psychological determinants of the situation that we shall derive our psychological interpretations.(p. 30)

The limitations are even more formidable outside the laboratory, where the experimental parameters cannot be incontrovertibly determined. Steinbruner (1974) has pointed out that when the value-preferences cannot be determined independently of consequent behaviour, the decision of actors is explained tautologically. The hard core of rational decision -- the integration of value trade-offs, the estimation of outcomes given uncertainty, the harmonious decision-making of collective actors, and the temporal, instrumental learning of these actors -- is assumed rather than verified by behaviour. The hegemony of rational analysis in everyday explanation is illustrated by the frequent use of behaviour as an indicator of intentions (values-preferences), e.g., the Soviet-Warsaw Pact decision to accumulate large conventional forces is seen as a clear indication of aggressive intent. But, even to distinguish between instances of "cooperation" and "defection" requires exogenous propositions regarding the values of each player, and players' "correct" evaluations of outcomes. In Gergen's terms, game theory requires "motive language," via theorists' concomitant subjective evaluations of the situation under study, in order to be socially intelligible. (Thus "Prisoner's Dilemma" seems an appropriate label for the frustration Schelling emphasizes

in a Hobbesian world of self-help, but not for the opportunity for cooperation which non-Realists like Rapoport and Axelrod point out; see below.) Such language may have legitimacy in the laboratory, where the experimenter can be reasonably sure she has manipulated these values and outcomes, but it is always contentious in describing behaviour in the real world. This is why game theory is so often criticized for ignoring change in a player's values, or perception of payoffs, particularly when values and subsequent payoffs may be partly determined by the other player's behaviour. (Green, 1966, p. 111)

In fact, psychologically interesting analyses focus on retreats from instrumental rationality. Schelling's emphases on theoretically extraneous factors of deception, commitment, posturing, and so on, vitiated the players' rationality by implying a need to stifle the opponents' information base and subsequent calculation of risk. For example, the strategists' rationale for communicating and upholding commitments to less-than-vital areas with conventional deterrence (e.g., Vietnam) was, at least in part, to bolster the credibility of nuclear deterrence to protect vital interests (such as extended NATO deterrence in Europe), thus implying that reputation for credibility in each case is linked. (This was one of Kissinger's main concerns.) But this merely avoids the irrationality of upholding nuclear deterrent threats once deterrence fails.

As Morgan (1985) argues,

Deterrence theory was developed to improve our understanding of the burdens of statecraft in the nuclear age. And the fact is that where nuclear weapons might have to be involved in upholding a commitment, there are no intrinsic interests of sufficient value to make that commitment inherently credible. (p. 131)

Earlier, Morgan (1977) makes the persuasive argument that in a world of perfect rationality, deterrence would be irrelevant:

If the problem is to prevent an attack, the decision maker will have all the necessary information about the opponent's intentions, capabilities, and cost/benefit calculations. Given sufficient forces, our perfectly rational deterrer can prepare a retaliation severe enough to tip the opponent's cost/benefit ratio against attacking. The trouble with all this is that in such a world the deterrer must face a truly rational opponent. Such an opponent presumably knew all along that if he made preparations to attack he would end up being deterred by the prospective retaliation -- the idea of attacking would therefore have been abandoned relatively early in his calculations, and the occasion for practicing deterrence would never have arisen.

If our would-be deterrer lacks the capabilities to ever do enough retaliatory damage to deter an attack he would, being perfectly rational, abandon this in favor of some other alternative -- capitulation, bribery, or the like -- at a very early stage in his deliberations. He would never attempt to practice deterrence, and wisely so because under those circumstances the perfectly rational opponent would not be deterred...

You will not feel the need for deterrence without some initial doubts or mistakes as to the enemy's intentions, plans, and capabilities. He will not have planned attack and yet now be deterred without having initially misjudged your intentions, plans, and capabilities. Deterrence accompanies this sort of frailty and ignorance, not perfect wisdom and foresight.

Laboratory gaming has explicitly dismissed perfect rationality. Rapoport and Chammah (1965, p. 29) report that although game theoreticians have been unable to solve

the paradox of mutual defection (the dominant strategy),

Ordinary mortals...when playing the Prisoner's Dilemma many times in succession hardly ever play DD one hundred percent of the time... Evidently the run-of-the-mill players are not strategically sophisticated enough to have figured out that strategy DD is the only rationally defensible strategy, and this intellectual shortcoming saves them from losing.

Even within formal theory, the iterated Prisoner's Dilemma forces a re-definition of rationality, and the most celebrated example is Axelrod's (1984) book on *The Evolution of Cooperation*. He opens his analysis with a question that was decidedly not the focus of the second wave: "Under what conditions will cooperation emerge in a world of egoists without central authority?" (p.3) Using the iterated Prisoner's Dilemma, he shows that players can tacitly communicate to each other, willy nilly, by their behaviour across previous plays of the game. The strategy each player chooses, therefore, must take into account the opponent's previous choices. Axelrod deduces that cooperation can emerge if the "shadow of the future" is long enough; that is, cooperation is a stable dominant strategy for each player if and only if the "discount parameter" (the degree to which the future is valued as a function of T, R, P, and S) is large enough. Thus, if both players are concerned in the present about how their immediate choices will affect their potential gains in the future, they may risk being a sucker (CD) and forego short-term temptations (DC) by cooperating. This challenge

to the competitive analysis of Schelling is derived from formal theory, but it begs a host of psychological questions about the development of trust, over which Schelling's static games could not even speculate.

So, upon close scrutiny, all of the important elements of deterrence -- the opponents' intentions, the components of credibility, the manipulations of misperception, the assessment of risks, the importance of past behaviour -- turn out to be exogenous to the second wave's static rational paradigm. While sometimes recognizing these shortcomings, Schelling and other strategists argue that game theory is methodologically useful as a structure for analyzing real world phenomena. But this is not literally true -- it is merely an awkward attempt to describe and explain the real world in the simplified, abstract, and often incongruous language of instrumental rationality. Such an attempt requires the theorist to make a whole assortment of crucial political and psychological assumptions which are not held open to scrutiny or debate. Thus, in Green's (1966, p. 124) words, "...game theory has merely provided us with a mathematical technique for explicating the meaning of our assumptions," mimicking Gergen's (1982, p. 77) charge that any psychological theory "...primarily furnishes intelligibility for what is already known." Whether or not game theory is used explicitly, Rapoport (1964) suggests that

Overwhelmingly...strategic analysis postulates "actors" whose only psychological traits are those which the strategist finds convenient to endow them with or those which the strategists, in their professional involvement, imagine themselves to possess. Nowhere in strategic literature does one find searching questions concerning the role of real psychological factors. In their place, one finds mostly formalized assumptions or standard cliches. It seems, therefore, that the strategist's wistful longing for psychological know-how will not be fulfilled. (pp. 123-124)

The Normative Inadequacy of Game Theory

Some of the implications of these flaws for prescribing security policy should be obvious; others are more difficult to discern. Recall that the third mode of rational analysis is normative: here, game theorists are concerned with determining optimal or "best" choices, usually via strategies (decision rules); these strategies outline how players ought to act, given the logic of instrumental rationality. Recall as well that game theory is very weak in prescribing acceptable and rational solutions in non-zero-sum games, leaving aside all of the problems alluded to above in relating any theoretical strategies back to the real world. For example, does cooperation entail appeasing an evil nemesis or merely realizing the security dilemma both players are in?

Divergent morals can be drawn from this weakness. For Schelling and other strategists, the moral was that players must retreat from rationality in order to pursue conflicting goals (e.g., national policy and war

avoidance). These problems presented themselves in the real world, even if rational analysis could not solve them. What the rational paradigm provided was an explicit formulation of a nasty Hobbesian world of states in anarchy. But since the rational paradigm could neither explain nor prescribe, it merely provided a rhetorical vehicle for entrenching the strategists' psychologically naive and politically controversial neo-Realist agenda of deterrence as the best way to manage (rather than reduce or resolve) conflict in this anarchy, with all of its givens: the existence of rational but intractably aggressive adversaries (e.g., the Soviet Union), the emphasis on military threats as the sole method of influence (negative threats rather than positive inducements), the need to bolster the credibility of larger threats by making good on smaller ones, the legitimation of a never-ending arms race to avoid the appearance of capitulation, and the call for sophisticated strategies of brinksmanship to exploit the instability of the nuclear age through intimidation.

But the rational paradigm's normative failure to provide acceptable solutions in non-zero-sum games (mutual defection in the Prisoner's Dilemma; risking sudden death to "win" in Chicken) provides the first legitimate moral, and it is this moral which Rapoport has expressed since the first strategists appeared on the scene (Rapoport, 1960, 1964, 1967). In Rapoport's view, whether or not

strategists consciously use rational analysis in their designs of strategy, the rational paradigm illustrates the logic of the strategic mind. The intellectual failure of this paradigm illustrates the hazards and pitfalls of strategic logic in the conduct of contemporary global politics. If, for example, the Prisoner's Dilemma is a model for an ongoing arms race, or recurring games of Chicken, between the superpowers, then the irrationality of a growing global war machine is seen as resulting from individually rational decision-making. The only acceptable solution must be to re-define rationality in terms of the systemic dynamic of international relations, with a new call for "collective rationality" to replace antagonizing national self-interest. This revelation -- derived from the paradox of strategic rationality -- would illustrate the inadequacy of the second wave in the nuclear age, and that "A better strategy must start with the realization that each side's civilization can be protected only by the other's cooperation." (Jervis, 1984, p. 29)

This mandate for cooperation can be illustrated within the language of game theory. In normative mode, Axelrod suggests that

Since the Prisoner's Dilemma is so common in everything from personal relations to international relations, it would be useful to know how best to act when in this type of setting." (p. 27)

He invited experienced game theorists to submit strategies in the form of computer programs to be played against each

other in a computer tournament of the iterated Prisoner's Dilemma. He found that the best overall strategy, played against all others including unconditional defection, was also the simplest: Anatol Rapoport's *TIT FOR TAT*. Rapoport's strategy is based on straightforward reciprocity; it chooses cooperation on the first move and thereafter chooses whatever the other player chose on the previous move. After an analysis of all the submitted strategies, Axelrod cites four reasons for *TIT FOR TAT*'s success. First, it is nice; it is never the first to defect. Second, it is provokable; it retaliates with defection immediately after being exploited. Third, it is forgiving; its memory reaches back only as far as the most recent move. Finally, it is clear; being the simplest of all strategies, it is easily recognized by its opponents.

This is an explicit refutation of the strategists' instrumental rationality, while still addressing the international security dilemma that Realists claim to be the reason for war-prone international relations (Jervis, 1978). Not surprisingly, then, Axelrod's new theoretical impetus for cooperation has been used by Realists to lend further legitimacy to classic diplomacy (see the October 1984 *World Politics* issue). Obviously, the strategies for superpower cooperation promoted by these theorists frequently contradict the second wave strategies for deterrence. But Axelrod's formal theoretic analysis does

not really settle the dispute over strategy. As in the second wave illustration of deterrence, game theory merely provides an elegant vehicle for illustrating the logic of cooperation, the pursuit of which is defined *a priori* by the theorist's normative commitment. This is why Jervis (1976) sees the distinction as deriving from different perceptions of Soviet intentions. (Are they motivated by aggressive designs that must be deterred, or are they motivated by defensive security concerns that must be alleviated through international cooperation?) Further, Morgan distinguishes between "immediate" deterrence, which assumes an imminent crisis (modelled by the game of Chicken), and "general" deterrence, which assumes only international anarchy (modelled by the Prisoner's Dilemma). The point is that determining the game states are playing, and thus the strategies to be employed, requires theoretical givens as to players' intentions, time-frame, and priorities. The rational paradigm appears to be silent, therefore, both on important psychological processes and on crucial political assumptions.

The Reification of Strategic Analysis

There is still a larger problem. While the deterrence agenda propounded by the strategists relied on political and psychological givens, it was the hegemony of this world view in state leaders that provided, in Gergen's terms, "a

willing cartel". Therefore,

The intellectual imperialism of deterrence theory is not just an academic fact, but a political act... The false attribution of expertness to an intellectual elite, which has in effect passed a test of political acceptability, narrows rather than enlarges the channels of influence. (Green, 1966, p. 275)

And by placing their faith in the logic of deterrence and the assumption of high conflict, strategists fell into what Rapoport (1964) calls the "zero-sum trap":

Recall that although prisoner's dilemma is a non-zero-sum game and is readily perceived as such, it is strategically indistinguishable from a zero-sum game in which both players have a dominating strategy... This is the zero-sum trap. Strategic thinking seals this trap, because it seeks one strategic "solution," but there is no strategic escape from the trap. (p. 107)

The second moral of the rational paradigm, then, is that any attempt to use a myopic behavioural theory of inter-nation influence as a technological tool for instrumental knowledge -- knowledge as to how to exploit the non-use of force in pursuing national objectives -- tends to reify the logic of that theory, and thus further legitimize the status quo. The strategists' gaming becomes a self-fulfilling prophecy, ignoring the inescapable conclusion that "A rational strategy for the employment of nuclear weapons is a contradiction in terms." (Jervis, 1984, p. 19)

MODIFYING THE RATIONAL ACTOR

Several alternative theories of inter-state influence

have been offered by the critics to replace the empirical and normative inadequacies of the second wave. While these alternatives vary considerably in methodology and theoretical content, I will briefly discuss two dominant traditions. Both approaches draw on work in experimental and social psychology to attempt to offer a new *behavioural* decision theory of international influence.

Abandoning Deterrence Theory: The Social Psychology of Conflict

The first general alternative comes from the long history of social psychological research on conflict. Much of this work has used experimental gaming as empirical support, and the normative commitment of researchers varies considerably. (Both Schelling and Rapoport have published in conflict theory's seminal *Journal of Conflict Resolution*.) Consequently, the theoretical emphases also vary, from the strategic theory of Schelling, to theories of bargaining and negotiation, to theories on the development of trust and cooperation. In terms of an alternative to the second wave, some of these theorists shift focus away from the competitive features of conflict (strategies and tactics leading to one party to gain advantage over another) toward the potentially cooperative features of conflict (strategies and tactics for constructively resolving conflict) (Deutsch & Shichman,

1986).

Morton Deutsch is exemplary of experimental researchers in this latter vein. Some of his crucial ideas on national policy were presented very early (Deutsch, 1961). In his view, the stable conflict applauded by strategists ignored such destabilizing psychological elements as misperception, nervousness in crisis, and reciprocal hostility. Later, he distinguished between "destructive" and "constructive" conflicts, and he implied that the strategists' deterrence resulted in the former (Deutsch, 1969; see also 1973). Such a conflict tends to expand and escalate independent of initiating causes. Deutsch cited three inter-related processes in destructive conflicts: 1) the drive for competitive advantage, forcing the situation into zero-sum terms (this is essentially Rapoport's zero-sum trap); 2) misperceptions and biased perceptions leading each party to perceive its own actions in the context of benevolence and the other's in the context of aggressive intent (attribution theory); and 3) the gradual unwitting commitment to escalation in order to justify previous action (cognitive consistency). (He further cited U.S. involvement in Vietnam as a classic destructive conflict.) His solution was to transform destructive conflicts into constructive ones by strengthening the cooperative elements in the relationship. What is important, however, is that such a transformation would require a shift in the

perceptions of each party; and this might be achieved through conflict management tactics of influence to replace an atmosphere of suspicion with an atmosphere of trust (see also, Deutsch, 1958; Deutsch & Shichman, 1986).

One of the best known programs for such a shift (from competitive to cooperative conflict management) was Charles Osgood's (1962) proposal for "Graduated and Reciprocal Initiatives in Tension-reduction" or GRIT (see also Linskold, 1978; Mitchell, 1986). Osgood saw the Cold War not as a result of an inevitably hostile conflict between two giants (the strategists' neo-Realist view, calling for deterrence), but rather, as a spiralling conflict of misperception due to the "Neanderthal" world view of state decision makers. Such a mentality derived from some of the same processes discussed by Deutsch: egocentrism, stress due to mutual hostility, stereotypy, psycho-logic, and so on. His answer was to take small, unilateral, cooperative initiatives to reduce international tensions, without eroding the robust retaliatory deterrent capacity necessary for national security. Such initiatives need not be only in the area of arms control, but the main goal was to significantly reduce the levels of armaments. The behaviour engendered would be an "arms race in reverse" with a slow de-escalation of superpower tension. But, again, such a program would require a world view quite different from that of the strategist. Whether or not

lowered international tensions (however defined) would facilitate such a perceptual and attitudinal transformation, an equally supportable hypothesis is that such armaments reflect real political tensions, rather than merely creating artificial psychological ones. The simple causal model offered by Osgood does not decide the issue unequivocally (Jervis, 1976).

Both Deutsch's and Osgood's prescriptions essentially called for positive influence tactics, emphasizing cooperation through trust development and tension reduction. Their proposals were drawn from their psychological expertise, an expertise which seriously doubted the simple and misleading psychological hypotheses of the second wave. I have certainly not done justice to the many interesting and important ideas these psychologists have provided and promoted. Once again, however, it is difficult to distinguish between their empirical conclusions and their normative commitment. For example, if social psychological research offers an alternative view of the Soviet threat (the American misperception of aggression, the inherent insecurity of hostile conflict, and so on), then the interpretation of the international environment, and therefore the normative questions to be addressed, will also be different. As the work on gaming illustrates, the difference between Schelling's emphasis on competitive conflict management and

Deutsch's emphasis on cooperative conflict resolution does not merely suggest an empirical moral (there may be one), but also reflects a difference in beliefs and values of the two theorists. Both are searching for a rational international relations, and both admit that there are irrational aspects to the nuclear age of mutual vulnerability; but their assessment of the problem to be solved rationally (pursuit of national interests via deterrence for Schelling; pursuit of stable international peace via conflict resolution for Deutsch) frames their respective rational prescriptions. Thus, it is not surprising that Axelrod's research on *TIT FOR TAT*'s success insinuated an alternative to deterrence which is similar to *GRIT*, without recourse to experimental psychology. The more sophisticated version of conflict behaviour, while perhaps providing psychological insight, does not settle the political question as to whether deterrence ought to be used. And this has been at the heart of the normative debate between strategic studies and peace research.

Saving Deterrence Theory: The Third Wave

The other major direction political psychologists have taken in refuting the second wave is to offer a "third wave" of deterrence theory, judging the appeal of deterrence on its own viability, rather than merely on alternative paths to security. (Jervis, 1979) Rather than

garnishing empirical support or dissent from the laboratory of experimental gaming, third wave scholars look to historical accounts, often using analytic tools drawn from cognitive and social psychology, to verify or refute the claims of strategic deterrence theory (Lebow, 1981; Snyder & Diesing, 1977). The most recent and ambitious attempt to make psycho-history the final arbiter in deciding the relevance and merits of deterrence is Jervis, Lebow and Stein's *Psychology and Deterrence* (1985). This leading edge of political psychology,

...represents an attempt to use history to shed light upon the most fundamental assumptions of deterrence theory and to test their validity in a broad variety of geographical, cultural, and temporal circumstances...to draw on propositions and concepts from major areas of psychology in order to understand how deterrence works -- or fails to work -- in practice. We have chosen this approach because deterrence is fundamentally a psychological theory. It is based on a series of "hidden" assumptions about the relationship between power and aggression, threat and response, and the ability of leaders to influence the calculations and behavior of their would-be adversaries. Until recently, these assumptions had to be accepted or rejected as matters of faith as there was little evidence that could be mustered in support of or opposition to them. This is no longer true. A growing corpus of empirical research... permits us to evaluate the psychological underpinnings of deterrence and to conceptualize about them in a more sophisticated manner. Needless to say, insights of this kind have important policy implications. (p. vii-viii)

Their mandate is clear. They attempt to fill in all of the psychological gaps that second wave rationality left unanswered -- the self-defeating effects of misperception, the intentions of states which determine when deterrence strategy is and is not merited, the use and misuse of

history and commitment in assessing credibility, and so on -- by constructing a less parsimonious theory of how state leaders actually make decisions and how state actors actually behave. While the rational paradigm may prove useful in providing a "baseline of expectations" for nomothetic explanation and prediction, third wave theory addresses the all important idiographic failures which have, at the very least, killed a lot of people, and which could, at the very worst, bring global catastrophe. (pp. 6-11)

Both in the extended case studies (the outbreak of World War I, 1914; the six years preceding the October War of 1973; the crisis leading to the Falklands War, 1982) and in the frequent brief illustrations, they provide an account of the "motivated" and "unmotivated" biases in the perceptions and calculations of political and military leaders, necessarily drawing on a loose information-processing paradigm of state behaviour. With this methodology, they reach a host of conclusions that are counter-intuitive to the strategists' portrait but rather unsurprising to more traditional accounts of international relations.

For example, Jervis and Lebow each show that once a threat is perceived, the intricate signalling prescribed by strategists is often inappropriate:

...on the one hand, complex and sophisticated strategies of coercion usually fail because they make

excessive demands on the state's ability to understand its environment and to coordinate its policy instruments, but, on the other hand, statesmen are slow to realize this and so many attempt policy of this type. Such complex policies are consistent with standard deterrence theory, and so the fact that states adopt them is partial confirmation of the theory. But the failure of these policies, and the refusal of decision makers to see that they are too complicated to work cannot easily be explained by it. A nice example is provided by the U.S. attempt to make force support diplomacy in Vietnam. (p. 7-8)

Everyone is thought to understand, so to speak, the meaning of fierce guard dogs, barbed wire and 'No Trespassing' signs. In practice, however, this may not be so. Statesmen, moreover, frequently adopt complex and finely calibrated strategies of coercion that make unrealistic demands on their adversaries' interpretative abilities. (p. 205)

These abilities, in turn, suffer from perceptual biases, domestic negotiation, and the projection of reputation (pp. 7-8, 27-28, 204-206). Stein also shows how misperception led to both the success and failure of deterrence, when Egyptian leaders first under-estimated and then over-estimated their military strength. Thus rational calculation is neither necessary nor sufficient for deterrence.

Lebow finds that even when four necessary and sufficient conditions of second wave deterrent signals obtained, aggression was nevertheless not deterred. This challenge to the second wave stems from the fact that

...aggression may be less a function of opportunity and more of perceived need...policy makers, at least in brinkmanship cases, are more responsive to internal imperatives than they are to external opportunities. (p. 183-184)

Stein's chapters on Israeli attempts to deter Egypt

similarly suggest that while deterrence may work for a time (as it did in the Middle East), it cannot (and did not) obtain indefinitely when the status quo is unacceptable to the deterree.

Returning to Morgan's critique of the strategists' linked-credibility argument in favor of protecting, and perhaps fighting over, peripheral areas, he challenges the received view that states defend lesser commitments to bolster their reputation vis-a-vis intrinsic commitments. Rather, it is more likely that

...nuclear powers may uphold lesser commitments because *they are the only ones they can certainly decide to fight for....* This converts a concern for reputation from a rational extension of the art of commitment into a pretense used to hide a pervasive insecurity over what to do if one's most important commitments are challenged. A vigorous response to attacks on lesser commitments might even be an index of this insecurity... (p. 134)

Finally, the most devastating critique of the second wave challenges the *a priori* of rationality -- namely the perception of threat and the determination of an adversary's intentions:

Judging others' intentions is notoriously difficult. Any number of methods can be used, all of them fallible. Statesmen can concentrate on what they believe is the other's objective situation if they believe that the external environment is the most important source of behavior; they can look to its past behavior if they think national character or stable domestic attributes are critical; they can study the goals, beliefs, and personalities of their opposite numbers if they think idiosyncratic characteristics matter. Of course, all three kinds of variables may and probably do influence foreign policy, but statesmen are no more able than scholars to construct useful hypotheses based on an excessive number of factors. (p.

14)

Jervis claims that "The single most important unmotivated bias which influences several aspects of deterrence is that people's thinking is heavily theory-driven," (p. 18) leading, for example, to the attribution bias of representativeness and availability:

Hitlers are very rare, but when a state acts in a way that resembles this model, statesmen are likely to infer that the state is very aggressive. They will not modify the judgment that stems from resemblance by a consideration of the base rates. (p. 24)

As well, there may be domestic-internal motivated biases to attribute bad intentions to other states, the attribution being a rationalization rather than a rational assessment:

The policy comes first, often for reasons that are politically illegitimate or psychologically painful to recognize, and the justification follows, reversing the normal [i.e., strategical] order in which beliefs about other states precede and lead to the foreign policy. (p. 25)

Thus, what may be a rational decision given specific military interests (e.g., the cult of the offensive) is actually a "motivated" bias leading to irrational state behaviour such that "Observers who look at the external situation, the state's interests, and even its previous behaviour will, then, be led to incorrect conclusions." (p.

27) Lebow concludes that

The demonstrable insensitivity of adversaries to each other's signals and interests and the clashing perceptions they so often have of each other's intentions and motivations point to what is perhaps the central failing of deterrence theory. This is the assumption, on which everything else depends, that adversaries relate to each other in terms of a common

frame of reference. (p. 229)

... Our analysis of adversarial relationships in the study indicates that the expectations that the two sides have of each other may bear little relationship to reality... Both sides, moreover, may also prove insensitive to each other's signals for a variety of political, cultural, or other reasons. In such circumstances, recurrent deterrence episodes may not facilitate greater mutual understanding. Experiences may actually hinder real learning to the extent that it encourages tautological confirmation of misleading or inappropriate lessons. (p. 231)

The essence of this alternative deterrence theory points to the absence of any universal and intelligible rationality in the decision-making and behaviour of states and their leaders, and to the hazards of nevertheless reifying a particular rationality in order to legitimize policy. The games states play are often delusionary, the rules are frequently abandoned, and the rationale for playing often derives from interests and attitudes far removed from either the Realist rationale of international anarchy or the neo-Realist rationale of adversaries' rational aggressive designs. Thus, it is the lack of any such real world rationality that makes orthodox second wave deterrence theory inadequate as an explanation of international behaviour and potentially hazardous as a normative device for conflict management.

The third, wave alternative to strategic deterrence theory is essentially an agenda for classic diplomatic efforts at (once again) manipulating adversaries' behaviour: "A more realistic goal of conflict avoidance may not be in denying an adversary the opportunity to act

but rather in minimizing its perceived need to do so." (p. 184) Lebow posits five incentives for aggression, but laments that four of the five (intra-elite competition, political weakness of leadership, weakness of a state's political system, and fragility of the state's existence) are not subject to any external (i.e., diplomatic) influence. This leaves only "the perceived need to forestall or compensate for a dramatic adverse shift in the political strategic balance of power," (p. 189) as a target for external manipulation, through a mixed strategy of deterrence and reassurance. Jervis and his colleagues suggest that externally-induced aggression results both from rational exploitation of opportunity (the neo-Realist deterrence argument) and from the unstable nature of anarchy (the Realist "security dilemma" argument), and that a comprehensive strategy for conflict management would have to involve (albeit sometimes contradictory) strategies to address both causes.

These strategies would, in turn, need to be informed by the instrumental knowledge culled by political psychologists:

The reality of international relations is less ordered, less comprehensible, more contradictory, and more unpredictable than deterrence theory admits. A more realistic attempt to describe the causes of international conflict and to offer guidance to statesmen for coping with their manifestations must take all these disturbing realities into account. To do so, it must sacrifice the elegance of abstract theory in favor of a more elaborate array of hypotheses based on careful empirical research. (p. 232)

Thus, the second wave scholars' prescriptions, which

mistakenly assumed rational international behaviour, would be replaced by the third wave scholars' rational assessment of the non-rational components of international behaviour, deriving from both the non-rational situation of anarchy, and the non-rational decision-making of state leaders. But Lebow is rather pessimistic about the prospect of such a program:

Deterrence, which, relatively speaking, is easy to implement, may nevertheless not be a very effective strategy of conflict management, because it does not address the most important [i.e., psychological] sources of aggression. On the other hand, efforts to alleviate the kinds of insecurities that actually encourage or even compel leaders to pursue aggressive foreign policies do not seem very likely to succeed.
(p. 192)

Dilemmas of Political Psychology: Radicals and Traditionalists

The third wave analysis, and its conservative agenda for reform, has been labeled "traditionalist" by Stanley Hoffmann (1986, see also Blight, 1986/87), who contrasts this trend in political psychology with those similar to the "radical" and "therapeutic" agendas implied by Deutsch and Osgood's work on constructive conflict resolution (a major concern of peace research as well). Hoffmann, a traditionalist himself, sees the former as the only relevant and potentially effective route for social science in reforming international behaviour and advancing the cause of peace. Jervis and his colleagues, agreeing with

Hoffmann that "politics is wholly psychological" also agree with him that

...there seems to be something implacably constant about the logic of behavior of sovereign actors in the international state of nature....[and] the student of present-day world politics who turns to Thucydides finds in his masterpiece, certainly not relief, but enlightenment and shock. The empires and nations of today seem to be playing the same ballet: the music may be different; the choreography hasn't changed. (Hoffmann, 1986, p.2)

In this view, therefore, the game is taken as a given (at least the game defined by careful and often empirical scholarship), and any normative commitments on the part of scholars, must be directed at reforming the way in which states play the game rather than conceiving of alternative games to be played. In essence, the traditionalist is merely the latest breed of strategist, who, while much more sophisticated in method of analysis and sensitivity to danger, is still attempting to find instrumentally rational solutions to save the Westphalian system of states, derived from the instrumental knowledge culled through careful analysis of history. The delineated areas of research

...aim at using the tools, theories, and insights of a variety of disciplines in order to dredge the huge pond that lies in between the very simple, stark landscape described by the theory of interstate behavior -- the logic of strategic-diplomatic conduct -- and the jungle of past and present international and national events. (Hoffmann, 1986, p. 17)

It is safe to say that neither traditional nor radical political psychology has had influence on the policies and conduct of the state to the degree that second wave

strategists have. But the only noticeable attention paid by strategic studies has been a new concern for the traditionalist focus on miscalculation and misperception in crises (Allison et al., 1985). They pay at least lip-service to the idea that when playing games of Chicken, things may go wrong. But the new "owl agenda" of dos and don'ts for nuclear policy give qualified support to the status quo of deterrence: maintaining NATO first-use and supporting modernization of the U.S. strategic triad, while avoiding provocative acts and signals, destabilizing strategies and weapons. Dismissing the debate over games (the "hawk" game of deterrence; the "dove" game of the security dilemma) as relevant but probably unresolvable, the concern is with safer strategies for whatever game the state wishes to play. Not surprisingly, today's state is only sympathetic to the intellectual discourse which uses the language and logic of the state as it behaves today.

REJECTING THE RATIONAL ACTOR

The Poverty of a Behavioural Deterrence Theory

I have surveyed the main components of the evolution of deterrence theory, profiling a long and arduous attempt to find an intellectual method to 1) determine the nature of deterrence, and 2) construct a set of appropriate strategies for international security given this nature of international behaviour. The first wave culminated in the

revelation that nuclear diplomacy might be characterized in psychological terms. When the second wave attempted to use rationality to define the psychology of deterrence, it turned out that rationality dissolved when addressing both unstable and robust deterrence relationships. The strategic critics charged that this dissolution left wide gaps for normative and analytic maneuver. Two alternatives were then offered, with radical agendas calling for an abandonment of the Realist games of Westphalian power struggles, and with traditional agendas calling for a new rational method for analyzing the lessons of history and prescribing new state strategies derived from these lessons.

Three underlying themes can be delineated here: First, while early scholars attempted to offer an "objective" analysis of international behaviour, later critics saw that strategic studies encompassed a host of normative commitments, a world view, which was integrated into the language of rational analysis, leaving crucial political and psychological questions unanswered. Second, the influence of theoretical work on deterrence has been directly related to the degree to which such work uses the language and logic of the state; either theory was used as a rhetorical device for legitimizing and reproducing the status quo, or it was dismissed as irrelevant to the real world of international politics. Finally, the paradigm of

rational decision, in the face of overwhelming empirical refutation by strategic revisionists and critics, remained hegemonic in the minds of policy makers, who in turn were often informed by the second wave.

The normative cul-de-sac of the traditional third wave of deterrence theory, and its eclipse of the radical alternatives to deterrence, result from the futility of constructing a behavioural theory of state influence to be used as an instrumental tool for state strategy. This futility derives from the fallacy of the privileged subject. It was argued in the previous chapter that knowledge about human social behaviour cannot derive from a universal method of objective inquiry, but rather, must necessarily contain the contingent and particularist interests, beliefs and values of the theorist. The intelligibility of social theory relies on this subjectivity. It was further argued that the theorist's analysis cannot remain socially benign, since the social act of analysis itself can and does effect its object of study.

The third wave scholars discussed in this chapter illustrated these points with regard to the second wave. The rational analysis and policy prescriptions offered by the initial strategists encompassed a host of political assumptions and psychological sleights-of-hand. As well, the inappropriate and dangerous behaviour of state leaders

was "theory-driven," and much of this theory derived from the second wave's compatible normative commitment to the pursuit of national self-interest. But by the traditionalist account, ending the receptivity of state decision makers to the inadequate reliance on strategic deterrence is a question of politics, not social science: the games may be dangerous but there is little we can do to change them. This conclusion implies that while second wave theorists were deluded, the third wave has reclaimed privileged analysis. Only by attributing to themselves the privileged truth-through-method that second wave theory lacked, can third wave theorists rationalize a conservative agenda.

Such a construction misses the most important moral to be illustrated by the rational paradigm: any games that are played, whether in the laboratory or in global politics, derive from the terms of reference of the players. As pointed out in the first chapter, there are no transcendental and intractable laws of behaviour "out there." As illustrated by the strategic critics discussed in this chapter, there is no singular rationality by which such behaviour can be analyzed or prescribed. There are only historically situated and socially negotiated conventions, informed by the languages (of diplomacy, of rationality, of behavioural scientific method, and so on) which players and adjudicators employ, which are in turn

derived from specific normative commitments (to cooperation, to security, to incremental reform, to pursuit of national interests, to preservation of the Westphalian system, and so on). The danger lies in claiming universal and immutable status to these conventions. Such claims, at the very least, narrow the possibility for changing the game, and at the very worst, result in legitimizing and reifying the games being played. As McGwire (1985) has suggested, deterrence has become a dogma which has constrained rather than broadened potential change in the international security agenda.

Toward an Emancipatory Theory of Deterrence

Of course, behavioural theory cannot and will not be abandoned by social science, and to the extent that critics of strategic deterrence manage to inject more caution and sensibility into the agenda of nuclear diplomacy, it can certainly be applauded. But social science must do more than enlighten decision makers as to what is. It must also realize its rhetorical power in illustrating and perhaps participating in what might be. The target of such a program cannot be the decision making elites; the evolution of deterrence theory points to the futility in attempting to transform the logic of the state via expert "off-the-shelf" instrumental knowledge. Rather, such an emancipatory program must be directed towards the general

public, to challenge the universalist claims of any one discourse or theory and those state leaders who proclaim it.

In this view, any theory is merely a more or less persuasive account, relying on a host of other foundational beliefs. A certain *Weltanschauung* was implied by second wave deterrence theorists, and challenged by strategic critics. But when radicals implied the need to change this world view, to change the normative commitment of state decision makers, traditionalists argued that the logic of the national security state was more or less intractable. This may be true for those in service to the state. But social science can challenge this intractability on the part of theorists, with its ethical and pragmatic inappropriateness for the nuclear age, by showing how it impinges and constrains their intellectual product. We need a theory, itself a "more or less persuasive account, of the essential subjectivity and relativity of any theories, and in particular, the Westphalian commitments imbedded in behavioural theories of deterrence. One possible program to this end is presented in the following chapter.

CHAPTER 3: Deterrence Theory as Lay Epistemology

A Recapitulation

In the first chapter, I compared and contrasted two versions of behavioural science. The orthodox version envisions behavioural science as attempting to inform social policy by making it a more rational instrument to achieve desired social ends. Thus, the content of behavioural science expertise must, itself, consist of instrumental knowledge, specifically the ability to explain and perhaps predict real world events via deductive-nomological inference. This expertise, the orthodox version suggests, can be derived from a specific epistemology, borrowed in large part from the natural sciences. This epistemology ascribes legitimacy to a specific form of knowledge by positing a privileged subject (the scientist) who can objectively observe empirical phenomena and construct theories which posit incorrigible covering laws (usually causal in content) which govern these phenomena. As a consequence of this epistemology, the various orthodox methods used in behavioural science all give pre-eminence to theories and constructs which are formal in definition (with operational concepts, axiomatic logic, and so on) and potentially falsifiable with empirical tests (direct measurement, statistical analysis, and so on).

The second chapter, however, illustrated many of the tensions within this tradition. Second wave deterrence theorists chose the method of rational analysis, exemplified by game theory, for the program of strategic studies, in an attempt to rationalize national defence policy. But critics challenged this program on grounds consonant with some of the challenges offered in the first chapter: Second wave theorists, it was implicitly charged, were working within a paradigm whose axiomatic foundations required a host of normative assumptions and commitments not held open to scrutiny. Furthermore, in second wave attempts to construct deductive-nomological theory, critics pointed to the lack of psychological and political insight into the deterrence phenomenon. The tautological nature of rational analysis, it turned out, vitiated the search for theoretically interesting and empirically testable hypotheses about international influence, contributing mere rhetorical support to existing defence policy. The third wave attempted to save deterrence theory by sacrificing parsimony in return for greater insight, subjecting both the second wave theoretical claims and some of the normative assumptions to empirical tests, through an ostensibly privileged analysis of historical cases. If the second wave failed, third wave theorists were clearly attempting to abide by the teleological, epistemological and methodological ideals of orthodox behavioural science,

using rational tools to define rationally, and perhaps mitigate, the consequences of frequent non-rational behavioural phenomena.

My critique takes as its normative starting point the view that, despite all attempts, the behavioural scientific program of deterrence theory has not successfully challenged what has become a dangerous and costly national security agenda in any important practical sense. This failure, in my view, has been due to the futile attempts at using an instrumentally rational method of inquiry to derive instrumentally rational expertise for the consumption of state decision makers. But what are the alternatives? Third wave theorists toss up their hands and claim that, for the most part, the fault lies in the realm of political action, which is not amenable to social scientific manipulation. This view is the basis upon which they criticize radical research agendas calling for a change in the rules of the Westphalian game and its international security dilemma. In essence, traditionalists want to keep intact much of the normative *hard core* of the strategic agenda.

The present critique suggests that the fault lies not just in the limitations of behavioural science, considerable though they may be, but primarily in the failure to use our rôle as manufacturers of knowledge to the best rhetorical advantage in changing the existing

order of things. Because of the conservative bias of orthodox behavioural methodology and epistemology, with its anchors of interpretation firmly grounded in historically and politically established conventions (such as, in this case, the Westphalian nation-state and its seemingly immutable diplomatic logic), the possibility of a truly emancipatory role -- to explore what is *not*, but what *might* be -- has been all but ruled out by these same orthodox epistemological and methodological commitments.

This chapter offers a possible alternative social scientific theory of deterrence which, while still grounded in the behavioural research tradition, takes the deterrence theorist, rather than deterrence, as its object of study. This alternative is based on the Theory of Lay Epistemology proposed by Arie Kruglanski. Two fundamental characteristics recommend Kruglanski's theoretical framework for the present task. First, his theory is a general theory of knowledge acquisition, providing social psychological insight into the interaction between affective and cognitive antecedents of peoples' theorizing about the world around them; that is, it explores the normative hard core of values, beliefs and finite alternative visions necessary to all knowledge-seeking behaviour. Second, Kruglanski's work takes into account several of the challenges to orthodox behavioural science already outlined, particularly the nonjustificationist view

of the fact-value fallacy, and the implications of science as rhetoric. This metatheoretical view is indeed supported by his own theoretical challenges to any notion of a privileged account of truth. In essence, his theory is a post-rationalist, social constructionist theory of knowledge acquisition, and as such, may provide the kind of rhetorical emancipation called for in an effective critique of deterrence theory.

ARIE KRUGLANSKI'S THEORY OF LAY EPISTEMOLOGY

Kruglanski's theory involves both a critical reassessment, and a theoretical synthesis, of existing social psychological theories concerned with the nature of human attribution and social judgment (Kruglanski, 1980, 1986, in press; Kruglanski & Ajzen, 1983; Kruglanski & Klar, 1987). Therefore, a proper account of his work must begin with at least a cursory outline of the traditions preceding him. While a comprehensive discussion of this large literature would be inappropriate here, I will highlight the major theoretical assumptions and conclusions of the research usually referred to as attribution theory, consistency theory, and human judgment theory.

Kruglanski's Critique of Prevalent Theories of Human Inference

Psychological inquiry into how people make sense of the

social world around them has a long history, but in contrast with many other fields in psychology, this research has remained for the most part within the cognitive paradigm (Fiske & Taylor, 1984). In particular, the classic texts on human attribution and social judgment (broadly concerned with how people formulate and reformulate their beliefs about, and attitudes toward, the behaviour of themselves, others, and social events) have usually adhered to an informational paradigm (Heider, 1958; Jones & Davis, 1965; Kelley, 1967; Tversky & Kahneman, 1973, 1974; Nisbett & Ross, 1980). The alternative motivational paradigm, focussing more exclusively on the nature of attitudes and general questions of affect, began with the early cognitive consistency theories (Festinger, 1957; Heider, 1958; Abelson & Rosenberg, 1958; Abelson et al, 1968), and the general work on persuasion (Hovland, Janis, & Kelley, 1953; for a contemporary review, see Petty & Cacioppo, 1981).

These two research traditions, the informational and the motivational paradigm, have been seen conventionally as theoretically incompatible because of their contrasting image of human inference and judgment:

The consistency model projects a largely irrational image in which a person's judgments and inferences are strongly biased by needs, wishes and desires. Attribution theory projects a highly rational image in which the lay person is likened to a sophisticated scientist or a dispassioned seeker after truth. (Kruglanski & Klar, 1987, p.1)

But while the informational approach in general, and attribution theory in particular, has been the dominant approach since the late 1960s, the concern with affect and motivation has "crept into" the foundations of even the earliest work, with discussions of what motivates people to infer in the first place, and how these affective dynamics, in turn, bias the inference process away from some rational ideal causing human error in judgment and attribution. Kruglanski's revisionist account of the existing literature is best illustrated in his critique of these two concepts, "bias" and "error," and in his later appeal to the implications of this revision in attempting to systematically incorporate the motivational dimensions of attribution and judgment into a synthetic theory of human inference.

Kruglanski and Ajzen (1983) reflect upon three major assumptions of most work on bias and error in human judgment: First, this work assumes that there exist known criteria of valid inference; this assumption implies that the investigator is in the privileged position of having an objective standard by which to measure actual human inference, invoking either ostensibly accurate normative models (such as statistics) or direct objective verification. In contrast, Kruglanski and Ajzen appeal to a nonjustificationist philosophy of knowledge to suggest that "one is never objectively justified in holding a given.

2

of/de

2

MicroD

proposition as true or in making quantitative (e.g., probabilistic) estimates of the extent to which it is." (p. 17) In principle, formal conceptualizations (such as those used in normative research methods) differ from informal ones (such as those the experimental subject might use) only in their linguistic precision, and not in the degree to which their veridicality is estimable. Therefore, there can never exist a privileged subject who can measure another's judgments or behaviours against "the facts" since these facts, however acquired, are in principle mere conjectures.

The second assumption which Kruglanski and Ajzen see as underlying the existing literature follows from the first, namely that motivational and cognitive psychological factors bias human judgments away from "valid" inferences. But since they refute any universalist claims to privileged inference, they take issue with the conventional definitions of bias and error. They see both terms as necessarily being subjectively determined. In their view, bias refers to a subjectively-based preference for a given inference or conclusion over alternatives, but this is necessarily true for both the investigator and the investigatee. In other words, all knowledge is biased. It follows that if bias is unavoidable, any meaningful notion of error must become subjective and tentative in nature, referring to the subjective "...experience a person might

have following an encountered inconsistency between a given hypothesis, conclusion, or inference, and a firmly held belief." (p. 19) In principle, while bias is unavoidable in human inference, error is no more or less likely than truth.

Finally, Kruglanski and Ajzen characterize the current state of attribution and human judgment theory as adhering to the notion of "epistemic pluralism," that is, the assumption that lay persons use a wide variety of disparate strategies -- schemas, attributional principles, heuristics, scripts, intuitive theories, and so on -- in inferring a host of causal and non-causal attributes and relationships in the social world. Speaking specifically in reference to the work on human judgment, Kruglanski and Ajzen attribute this pluralism to the decidedly empirical nature of the research, caused by the absence of an overall theoretical framework to subsume these disparate strategies under one concept. It is, of course, this last assumption which Kruglanski is specifically committed to overturning with his own theory. But another possible explanation of the existing pluralism may derive from their first two criticisms: epistemic pluralism may reflect the disparate biases of the researchers themselves, informed by specific experimental contexts. In this sense, then, Kruglanski's work represents an attempt to synthesize theoretically the existing empirical research, while at the same time

abandoning some of the premises upon which this research was originally grounded.

As already mentioned, the other major dilemma of existing research on human attribution and judgment to which Kruglanski's work is addressed is the issue of motivation in human inference:

...the motivational dimension as such has not been systematically incorporated with the more informational aspects of attribution. Thus, it has not been made explicit whether motivational influences signify a separate route to attributions, juxtaposed to the informational route, or whether motivational and informational factors are parameters of the same process whereby all attributions are rendered. (Kruglanski, 1986, p. 4)

His theoretical resolution to this problem, elaborated below, is consonant with his earlier attack on the notion of the privileged subject. In his view, most theorists' commitment to a rational and objective method for measuring lay bias and error encouraged these same theorists to avoid the question of non-rational and subjective motivations in the inference process. In contrast, Kruglanski's perspective implies that "the motivational dimension constitutes an inseparable part and parcel of all attributional activity," and indeed all human judgment. (Kruglanski, 1986, p. 41-2) Thus, once the commitment to a search of the paragon of rationality is abandoned both in method and in theoretical content, the issue of motivational (as opposed to purely cognitive) biases is less problematic. From this new perspective, human

inference reflects neither the hyper-rationality of early attribution theories, nor the qualified rationality of human judgment theories, nor the irrationality of cognitive consistency theories. In fact, Kruglanski has attempted to incorporate all of these theoretical traditions, with certain adjustments and qualifications, into his general theory of knowledge acquisition, discussed below.

Kruglanski's Theoretical Alternative

A definition of knowledge. Kruglanski's Theory of Lay Epistemology is concerned with the general psychological process through which individuals acquire knowledge about the social world. His general conception of this process is conditioned by a similarly general and all-inclusive definition of knowledge. In Kruglanski's account, a person's knowledge at any point in time is comprised of an indeterminate number of propositions. These propositions can be characterized in two dimensions: content and confidence. The contents of propositions are of a potentially infinite variety (either causal or non-causal in nature), and may be labelled as thoughts, beliefs, feelings, and so on, depending on the specific context. (Note that Kruglanski does not see a clear or meaningful distinction between affective and cognitive dispositions.) Furthermore, propositions can be subscribed to with more or less confidence, depending on the antecedent premises from

which they are derived; in this sense, "facts" are merely propositions in which the knower has more confidence than in the case of "hypotheses" (again, consistent with a nonjustificationist epistemology).

The epistemic sequence. All knowledge-seeking behaviour is proposed as a two-step process: hypotheses generation and hypotheses validation. The generation of mutually exclusive hypothetical propositions is teleologically functional in that it serves some particular intrinsic or extrinsic interest or interests of the knower; that is, hypothesis generation is conditioned by "epistemic motivations" (to be elaborated below). The hypothesis generation stage is also conditioned by cognitive capacity: limitations of memory, availability or accessibility of alternative propositions, attention span, momentary salience, and a host of other dynamics examined in the research on human judgment, learning and memory.

The second stage in the sequence, the validation of hypotheses, is seen as a universal invocation of standard deduction of an *if-then* nature. A person validates generated propositions by deducing them from other premises to which the individual already subscribes. This suggests that for the hypothetical proposition to be considered at all, it must have some subjective relation to existing knowledge propositions. Crucial to Kruglanski's account, however, is the claim that in principle there are an

infinite variety of possible evidential categories from which the knower may choose to validate a hypothetical proposition. In this way, astrologers, inductivists and social scientists all validate new propositions with the same if-then linkage, even if the content of this linkage differs radically across their disparate epistemological traditions. To illustrate, the astrologer might believe that if "the moon is in the Seventh House" then "there may be trouble in store for Capricorns;" the inductivist might believe that if "many swans are observed to be white" then "all swans are white;" and the social scientist might believe that if "the experimental variance differs statistically from the control condition" then "the experimental stimulus has had an effect on response." All of these particular categories of evidence are nevertheless used deductively in the validation process. Finally, the confidence a person will have in any validation will be a function of the confidence in (commitment to) the existing premises appealed to in the deduction. The degree to which the astrologer, inductivist, and social scientist will have confidence in their conclusions will be conditioned by their confidence in astrology, induction, or statistical methods, respectively.

"Freezing" and "unfreezing" of the process. The vehicle through which Kruglanski incorporates the motivational dimension into his theory, and the element of most import

for the present discussion, is in his account of when and how knowledge-seeking behaviour stops (or starts) and knowledge is acquired (or revised):

...the deductive validation of hypotheses has no unique or natural point of termination. In principle, it should be always possible to come up with further alternate hypotheses compatible with the same body of evidence.

The absence of a natural termination point for the epistemic sequence poses the following dilemma for a theory of knowledge acquisition: On the one hand the sequence could continue incessantly without us ever being able to confer the status of firm fact on any of our hypotheses. On the other hand, at most times we do seem to possess numerous items of definite knowledge, that is, things we "know for sure." This suggests that, pervasively, the epistemic sequence does come to a halt somehow. (Kruglanski, 1986, pp. 6-7)

Kruglanski refers to this starting and stopping of the epistemic process as the "unfreezing" and "freezing" of knowledge, respectively, and in this way he accounts for both the crystallization of new knowledge and the modification of established knowledge. This dynamic is conditioned by the two determinants of generation mentioned above, namely cognitive capacity and epistemic motivations. Kruglanski distinguishes three epistemic motivations. First, a "need for non-specific structure" reflects a desire for any answer, regardless of content; this might be the case, for example, in a situation where there is a need to make a decision. Second, a "need for specific structure" reflects the knowledge-seeker's strong preferences for a particular conclusion; this would promote freezing when existing knowledge is congruent with the

knowledge-seeker's desires, and unfreezing in the opposite case. Third, a "need for ambiguity" might reflect a fear of commitment to an erroneous judgment, a preference for a lack of structure, and so on; in general, such a motivation would encourage unfreezing.

While the three motivations differ in their desired end-states, they all fulfill a similar epistemic role, namely to effect the freezing or unfreezing of knowledge, and therefore, the initiation or halting of the epistemic sequence. Furthermore, there is in principle an infinite variety of potential origins for each of the three epistemic motivations. These antecedent motivations, along with particular limitations on cognitive capacity, lead to the unavoidable bias in all knowledge acquisition, including the subsets of judgment and attribution.

Implications for an Analysis of Deterrence Theory

In one sense, Kruglanski's work represents the most recent account of social knowledge acquisition, within the mainstream of social cognitive psychology. As seen earlier, much of the antecedent research upon which his own construction is built has been explicitly appropriated by both the "traditional" and "radical" research on conflict. One of the most recent illustrations of this can be found in Larson's (1985) landmark political psychological analysis of the *Origins of Containment* where she draws

exclusively on five cognitive theories of social judgment and attribution (Hovland's persuasion paradigm, cognitive dissonance theory, attribution theory, Bem's work on self-perception, and the schema paradigm championed by Schank, Ableson and others). One implication of Kruglanski's work, then, might be to bolster this traditional political psychology, providing, for example, the well-integrated model of motivated and unmotivated biases which Jervis et al. (1985, p. 18) have not yet found:

As they seek to judge whether others are a threat, statesmen are subject to both motivated (that is, affect-driven) and unmotivated (that is, purely cognitive) biases. But a theory that shows how these are combined, is still beyond our grasp.

In fact, Klar, Bar-Tal and Kruglanski (1986) have already offered one such analysis of an "epistemology of conflict" to suggest some of the social cognitive antecedents to conflict and its potential resolution. Such work can be applauded for the same reasons cited earlier for political psychology.

In the remainder of this chapter, however, it is my intention to explore what I believe to be the more profound implications of Kruglanski's theory, namely his explication of the unavoidable subjectivity of all knowledge acquisition, including social scientific theorizing about the social world, and in particular, theorizing about international security. From this perspective, his theory

provides a behavioural scientific account of the *hard core* of any research program.

Some of these implications are spelled out in his analysis of the psychological underpinnings of behavioural science (Kruglanski, 1986). He claims that "the process whereby scientists and lay persons arrive at their 'knowledge' is essentially the same," and, in concert with Gergen, suggests that

...according to present theory, deduction alone is insufficient to account for the inference process. In addition, it is necessary to specify the various cognitive and motivational factors affecting the generation of alternate hypotheses...and the search for relevant information. Finally, in order for an idea to become part of scientific knowledge it must be accepted not just by an individual scientist but by a whole community of scientists. (p. 19)

Thus, Kruglanski illustrates the interdependence between analytic and normative commitments which was implied in the first chapter's critique of behavioural science, and which was a recurring theme in the second chapter's genealogy of deterrence theory. My objective in the following section is to re-interpret the most influential wave of deterrence theory -- the second wave and its legacy of strategic studies -- in light of Kruglanski's attack on the privileged subject.

DETERRENCE THEORY AS A LAY EPISTEMOLOGY

What are the historical, political and psychological antecedents which delimit the research paradigm of

strategic studies? Or, in Kruglanski's terms, what are the cognitive and motivational causes for selective "freezing" and "unfreezing" which condition second wave theorists and their intellectual products? To understate the case, this is a difficult issue to address. It has been a major theme of my whole discussion that all knowledge acquired about the social world, including specific theories about that world, are essentially social constructions which can never be firmly established and protected from further revision, re-interpretation, or out-right rejection. Any such social constructions serve a variety of contingent interests on the part of the theorist, and derive from his or her own normative commitments. The following discussion is no exception.

My object of study is "the Strategist," as first portrayed in Rapoport's (1964) polemic. This is not the artificial "strategic man" posited by second wave theorists, nor is it a label for a certain group of people. It is, rather, a particular role which may be seen to condition the epistemic process within the discipline of strategic studies. It is to the Strategist's epistemological status that I now turn.

The Hard Core: Realism

As I have frequently alluded in the previous discussion, the major value orientation of strategic studies is

primarily influenced by the general tradition known as Realism. In other words, when the Strategist is generating alternative hypotheses about international security, he is constrained by the Realist's cognitive and epistemic-motivational biases.

What characterizes the *propositional* content of Realism, or more precisely, the particular strain of Realism that frames the context of strategic studies? A sketch of some of the Strategist's fundamental propositions might derive from a selective synthesis of two specific Realists, Hobbes and Clausewitz:⁶

Proposition 1: The international environment is characterized by decentralized power in the hands of states in anarchy. Without a central authority, the mores and conventions found in civil society do not obtain in the relations among states.

Proposition 2: The resulting state of war, all against all, leads to the security-dilemma, with each actor (state) attempting to ensure its own survival by unilaterally reserving the right to use force in an environment of self-help and natural right.

Proposition 3: War and the threat of war are merely rational instruments of state policy, in pursuit of survival in the international state of war. Preparations for waging war are the best protection against war: "Si vis pacem, para bellum."

This is only a partial list, and other more complete accounts can be found (see especially Rapoport, 1964, 1967a). Both Hobbes and Clausewitz wrote in the pre-nuclear Westphalian heyday. It has been argued by classical Realists (e.g., Bull, 1977) that in the most peaceful period of the nineteenth century, the crucial role of the balance of power was actually conditioned by a certain set of international conventions, notably Grotian international law, which comprised an "anarchical society." Furthermore, some Realists have pointed to both the theoretical and normative inadequacy of Clausewitzian formulations in the nuclear age (e.g., Jervis, 1984; Kennan, 1983; Morgenthau, 1964). Indeed, in the previous chapter we saw the way in which nuclear weapons posed significant problems and contradictions for the first wave of deterrence theory.

But it is the truncated strain of Realism, as summarized in these propositions, which forms the hard core of irrefutable "facts" of strategic studies. Recall that for Kruglanski, "facts" are merely hypothetical propositions for which the knower is more confident. Recall as well that propositions need not be exclusively analytical; they may be thoughts, beliefs, feelings, and so on. From Kruglanski's perspective, we need not, and probably cannot, disentangle analytical and normative propositions. Nevertheless, the question of most concern here is how the

Strategist's knowledge reproduces and sustains itself, despite some serious contradictions?

The Protective Belt: Rational Analysis

To understand the endurance of the Strategic mindset, and of its product, deterrence theory, we must speculate on the Strategist's epistemic sequence. As for all other knowledge acquisition, this is a two-step process of hypothesis generation and validation. While the generation of specific mutually exclusive hypotheses is potentially infinite, the Strategist is constrained by cognitive capacity and epistemic motivations.

The Strategist's cognitive capacity is delimited by the content categories, outlined above, and by the rational paradigm, from which he generates and validates his various hypotheses on deterrence. Thus, intellectual theorizing fulfills a fundamental psychological need to abstract and simplify what is otherwise an infinite, contingent, and disordered world of propositions. In particular, any hypothesis for achieving security via "collective rationality" (as Rapoport and Axelrod suggest) cannot be accommodated by the rational paradigm as the Strategist uses it (that is, for discerning instrumentally rational strategies). Furthermore, any of the more problematic hypotheses about the tensions in the concept of deterrence, as some second wave critics and classical Realists have

suggested, are similarly vitiated by the rational paradigm. Therefore, the only hypotheses the Strategist can survey are those which place the state actor as the agent of security, within an international environment of self-help, hence leading to the Strategist's cognitive obsession with credibility and commitment. Just as the liberal micro-economist starts with certain simplifying assumptions about individual consumers and producers (a free market, perfect competition, rational self-interest), so too does the Strategist start with similar assumptions about states (dangerous anarchy, adversarial aggression, stable conflict, national pursuit of national security), in order to see how far these assumptions get him in "the production of theory."

This cognitive constraint is sustained by the Strategist's paradigm of validation, and his teleologically functional epistemic motivations. This is how the strategic mindset is frozen or crystallized. Validation, in the strategic case, involves an epistemic commitment to the rational paradigm. If a generated hypothesis -- about the robust or unstable nature of deterrence, for example -- can be modelled by a game-theoretic formulation, it can be "validated" by calculating the individually dominant strategy for each hypothetical player. It is the Strategist's commitment to the rational paradigm which further explains his confidence in subscribing to the

"facts" of the international security dilemma, the imperatives of deterrence, and the national strategies for bolstering this deterrence. Given the twin *a priori* premises of the Realist security dilemma and the rational paradigm of validation, the Strategist is confident that his games are the most appropriate reflections of real world imperatives.

One way to interpret the move from the Realist first wave to the Strategist second wave may be in terms of the epistemic motivation for "non-specific structure." Seeing the inherent contradictions in nuclear deterrence, the defence establishment unfroze its Clausewitzian heritage and sought means by which it could reassert order in an increasingly disorderly and insecure international arena. Strategic studies provided "solutions" for control and manipulation of international behaviour, and thus became an indispensable part of the state.

But because the state was the primary sponsor of strategic studies, the resulting order had to preserve the state. Therefore, the Strategist's most important epistemic motivation may be the "need for specific structure." That is, the Strategist is only interested in hypotheses which can accommodate the Westphalian structure of global security. Any conclusion drawn from the validation of hypotheses will only be accepted (frozen) when and if it fulfills the Strategist's telos:

preservation of the Westphalian state. Even if he can generate hypotheses which suggest the need for collective rationality, he will continue to generate alternative hypotheses, within his rational paradigm, until he finds a state-preserving option. Thus, as we saw in the previous chapter, the Strategist will ignore the contradictions of Chicken and Prisoner's Dilemma games, focussing on tactics which can preserve the state's autonomy in achieving security.\

Alternative Epistemologies

The Strategist's epistemology contrasts with, among others, the epistemology of the Peace Researcher. Rather than attempting to delineate an archetype (which does not exist in the diverse field of peace research), it may be helpful to suggest a few propositions which the Strategist does not recognize.

When Strategists look at the world, they do not see the twentieth century demise of the Westphalian order. They do not see sovereignty's irrelevance to the "structural violence" in the Third and Fourth Worlds which comprise a politically and economically enslaved "periphery," dependent on the more industrially advanced "core" societies; nor do they see the lack of effective state regulation and influence in the nevertheless continuous non-violent conflict management and resolution between

various non-state actors and agencies which comprise the "complex interdependent" Western capitalist societies. Indeed, most classical Realists exhibit remarkable silence on the "low politics" of the world political economy.

Even in the security domain, Strategists have had great difficulty in interpreting and explaining the non-state origins of collective violence. For example, Strategists constantly submit proper nouns as the prime movers behind terrorism (Qaddafi, Khomeini, Syria, Iran) without considering that such acts reflect the otherwise latent conflict between those who are recognized (e.g., states such as Israel) and those who are not (e.g., non-state actors such as the PLO, or mass movements such as Islamic internationalism). Furthermore, since Strategists are concerned with conflict management rather than conflict resolution, they refuse to recognize the overwhelming contemporary evidence against the effectiveness of military force in resolving conflict. In particular, they are insensitive to the "radical" agenda for cooperation and conflict resolution. Finally, they do not see the individual state military machines as a collective, self-reproducing, and symbiotic war system whose only purpose is to maintain the Westphalian global order.

What these alternative propositions have in common is that they do not take the Westphalian normative order as an unassailable given. The knowledge of the peace researcher

is conditioned by other imperatives such as social justice, non-violent conflict resolution, and perhaps a hope for some form of collective human emancipation.

These two brief characterizations -- of the Strategist and the Peace Researcher -- may appear to be simplified or caricatured. But my intent has been to highlight the social-psychological components which help to explain the endurance of a particular strategic mindset, without recourse to other political explanations. Seen in Kruglanski's lay epistemic terms, the manufacturing of deterrence theory emerges as merely one possible program of knowledge acquisition. Like all other knowledge, it is conditioned by the particular constraints on the epistemic process. It is these constraints which frame and sustain the contentious normative commitments of the Strategist.

Nevertheless, if the Strategist's knowledge claims have been reduced to the level of lay epistemology, this does not mean that such claims have been rejected. The implication here, and throughout my discussion, has been that attempts to challenge deterrence theory via an appeal to facts is likely to prove ineffective. Strategists have their facts and their epistemologies, just as peace researchers have theirs. Moreover, since beliefs, desires, interests, and ethics are all merely other knowledge propositions, any dialogue cannot appeal to extra-epistemic ideals. This is the Kuhnian problematique: There is no

rational recourse for choosing among paradigms. The present lay epistemic account of the necessary and enduring normativity of the Strategist's paradigm, and its production of deterrence theory, implies that any emancipatory disarming and supplanting of the Strategist's agenda must invoke strategies for "unfreezing" the crystallized knowledge of strategic studies. And these strategies must be motivated by the underlying non-rational ethical debate between Strategists and Peace Researchers, namely that of conscience (Rapoport, 1964).

CONCLUSION: Social Science as Political Activism

My concern throughout this discussion has been with contemporary modes of thinking within the behavioural sciences in general, and within strategic studies in particular. The signal intellectual product of strategic studies has been the behavioural program of deterrence theory. As surveyed in the second chapter, this program is informed by certain *a priori* assumptions. As the third chapter illustrated, it is the commitment to these normative and epistemological assumptions which in turn defines the Strategist. Some behavioural scientists have criticized the orthodox second wave of deterrence theory, but their critique has been limited by their own adherence to the same methodological and epistemological commitments of orthodox behavioural science, as outlined and re-appraised in the first chapter. I have attempted to move beyond the debate between "radical" and "traditional" political psychology by challenging the epistemological legitimacy of deterrence theory and the normative foundation upon which it is sustained.

As with any social scientific theory, deterrence theory is a social construction which can claim no ultimate foundation in immutable social logic. Its endurance is better explained not by its intellectual privilege, but by its rhetorical support for the international security

status quo. Orthodox behavioural science has often been an accomplice to this conservative trend through an insensitivity to its own historicity. This is best seen in the traditionalist third wave of deterrence theory: using some of the most sophisticated behavioural scientific knowledge, the result is a normative cul-de-sac, with a highly restrained appeal for patience, tolerance, and incremental reform of a highly inadequate global order.

For a variety of reasons, alternatives such as those offered by radical political psychologists have had close to nil effect. For the most part, this has been due to the politics of the state: for the same reasons that strategists are afforded the highest credentials, peace research is largely ignored by those in power. But as long as alternatives are couched in the same rational, instrumental language of orthodox behavioural science, they are not likely to achieve any status of their own, outside of a few university colloquiums and peace groups.

One possible way out of this intellectual and political isolation has been my final focus. Kruglanski's theoretical work highlights some of the main concerns of previous reflections on the status of social science, supporting the most important conclusion: that social science is fundamentally a form of political action. His social psychological account supports a social constructionist view of knowledge, and the concomitant

possibility of emancipation from existing social orders. And his theory suggests how all knowledge is, at its roots, a struggle to assert a specific set of norms.

If behavioural science can help us to understand the thoroughly subjective, historicist, and socially negotiated nature of all theorizing, from the strategist to the peace researcher, then we may be able to debate the merits of competing paradigms -- of divergent lay epistemologies -- not by some inappropriate epistemological criteria laid down by positivist philosophy, but rather, by focussing on the ethical desirability of these paradigms' implications. Stripped of its epistemological clothes, deterrence theory would appear to be highly inadequate and morally suspect. A new debate would require participants to "unfreeze" their own thinking, to entertain new conceptions of peace, security, human progress, and social justice.

As stated in the first chapter, this should be a welcome conclusion for all behavioural scientists. The implication is that the surface structure of experiments, model-building, and empirical observation is actually secondary to the deep structure of conflict between competing ideas of what ought to be. This is what peace education must be about. We must bring the discourse back to the underlying social and political questions which beg to be asked in the contemporary world.

Notes

1. The term, paradigm, has been used to connote various ideas in the social sciences, and there appears to be little consensus on what defines a scientific paradigm. Thus, some observers might refer to the "behavioural" paradigm and the "rational analysis" paradigm as specific and coherent methodological, epistemological and ontological commitments presumed to exist across several disciplines; others might refer to the "informational," "motivational" and "psychoanalytic" paradigms as competing methods of theory construction within the discipline of psychology; and still others might refer to the "perceptual," "informational" and "developmental" paradigms talking past each other within the tiny sub-discipline of the psychology of reading. Although he was certainly not the first to use the term, some have blamed this lack of consensus on either the ambiguity of Kuhn's original statement, or on the misreading by his audience. I suggest, however, that Kuhn's original conception of paradigm (as discussed here), deriving from specific Gestalts, can and does occur at any of these theoretical levels, even if the resulting paradigms are interdependent. Thus, the strategic studies, by the present definition, attempts to work within a specific epistemological and methodological paradigm (behaviouralism) and within a specific ontological paradigm (states' struggle for power as a first cause in international relations).

2. The term, normative, is used here to denote several implications. Normative models or theories (such as statistical inference, rational analysis, classical realism) all attempt both explanation and prescription with the use of some ideal as a criterion. Thus statistical inference compares real world events with the "expected" outcomes derived from non-empirical mathematical models; rational analysis (see below) compares an actor's behaviour with the "expected" behaviour of an infinitely knowledgeable and self-serving prototype; and classical Realists compare real world states or state decision-makers with their often ambiguous ideal of how the rational and prudent state or statesman achieves state security.

In the present view, the validity of these ideal types are not subject to verification or falsification in terms of "correspondence to truth," but are rather the *a priori* telos to which the theory is addressed. These normative ideals are essentially informed by the theorist's beliefs, desires, or specific intellectual or political interests or values. In the present discussion, normative commitment will refer to any or all of these subjective antecedents.

3. For heuristic value, I will use Jervis' conception of "waves" of deterrence theory throughout my discussion. I will not, however, sustain any argument regarding the historical accuracy of this progression. As I will make clear below, third wave critics appeared on the scene much earlier than Jervis suggests, and it can be argued that second wave strategists are still very much the rule rather than the exception. Jervis' conception of three waves is better seen as a flawed but useful normative differentiation rather than a pre-Kuhnian scientific cumulation.

4. Realism and Neo-Realism: In contrast to the philosophical definition of realism as a specific ontological or epistemological commitment, the present concern is with the theoretical and methodological contrast of these two traditions within the International Relations literature. In the present discussion, both terms will be used to denote a normative commitment to the maintenance of the state (McKinlay & Little, 1986), and the view that a state's security derives from the power it can wield in an anarchical environment of competing states. (The link to the much older philosophical definition derives from the claim that IR Realists are primarily concerned with what is, namely the contemporary dominance of competing states in a struggle for power.) The two traditions diverge, however, in considering the nature of this anarchical environment, international insecurity, and the conception of power.

For classical Realists (e.g., Thucydides, Morgenthau, Aron, Bull, Kennan, Hoffmann), power is a dialectical and often elusive concept, and although a primary determinate is military capability, other historically situated determinants such as diplomatic skill, prudential statemanship, and perhaps even international opinion are considered important to explaining and prescribing state behaviour. In considering the inherent insecurity of states, classical Realists place most of the blame on the "war-proneness" of international anarchy, which forces states to compete for power in intractable but often irrational systemic anarchy. For many classical Realists, the multipolar Balance of Power that endured through the nineteenth century is an exemplary model for assuring stability and security in anarchy.

For the more recently dominant neo-Realists (e.g., Waltz, Gilpin, and virtually all strategists) power is determined solely by the military capability a state can use, or threaten to use, in order to defend vital interests or seek confrontation in other less vital areas. The insecurity of individual states is ascribed to the inherent

search by major state actors to increase their power or domination over others, via increased military capability and its use. Because they see international anarchy in absolutist, Hobbesian terms, they generally reject proposals for collective security, and favour instead the unilateral threat or use of force to maintain stability, and consequently see advantage in a bipolar, as opposed to multipolar, security regime.

Recently, critics of neo-Realism have characterized the method of neo-Realists as specious and conservative (eg, Ashley, 1984), but their method, as implied by the present discussion, follows from their normative commitment. The dialectical conception of power is not readily accessible to empirical analysis, whereas the a-historical, positivist conception of power is. Nevertheless, some behaviouralists (e.g., Jervis, 1984) may be considered classical Realists, particularly by their historically reasoned accounts of the role of diplomacy and non-military determinants in the dynamics of the international security dilemma.

5. Rationality: This term has such wide currency in both scientific and lay discourse, that its specific meaning in this discussion must be clarified. By rationality is meant any normative commitment to reason (either the psychological reasoning of the individual or the artificial reasoning ascribed to a collective actor) as the cause and explanation of behaviour. Contemporary theoretical explanations are usually couched in instrumental terms, rational behaviour being defined as the appropriate means for achieving a desired end. Rationality is usually considered independent of any criteria for assessing the worthiness of the ends being sought. But in appealing solely to rationality as the explanation of behaviour, there is no *a priori* account taken of these ends, other than by reference to means. Thus, the use of rationality as an ideal is tautologous. Nevertheless, by the present definition, any normative commitment is a tautology. Rational analysis (discussed below) is merely a formalized symbolic account of a specific tautology.

6. For an excellent interpretation of the Hobbesian and Clausewitzian heritage which informs strategic studies, see Klein (1986).

References

- Abelson, R. P., et al. (Eds.) (1968). *Theories of cognitive consistency*. Chicago: Rand McNally.
- Abelson, R. P., & Rosenberg, M. (1958). Symbolic psycho-logic. *Behavioral Science*, 3.
- Alker, H. R., & Biersteker, T. J. (1984). The dialectics of world order: Notes for a future archeologist of international savoir faire. *International Studies Quarterly*, 28, 121-142.
- Allison, G. T., Carnesale, A., & Nye, J. S., Jr. (Eds.) (1985). *Hawks, doves, and owls: An agenda for avoiding nuclear war*. New York: W.W. Norton & Co.
- Apfelbaum, E., & Lubek, I. (1976). Resolution versus revolution? The theory of conflicts in question. In L. Strickland, F. Aboud, & K. Gergen (Eds.), *Social psychology in transition*. New York: Plenum Press.
- Ashley, R. K. (1984). The poverty of neorealism. *International Organization*, 38, 225-261.
- Axelrod, R. (1984). *The evolution of cooperation*. New York: Basic Books.
- Ball, T. (1976). From paradigms to research programs: Toward a post-Kuhnian political science. *American Journal of Political Science*, 20, 151-173.
- Banks, M. (1984). The evolution of international relations theory. In M. Banks (Ed.), *Conflict in world society*.
- . (1985). The inter-paradigm debate. In M. Light & A. Groom (Eds.), *International relations: A handbook of current theory*.
- Blight, J. G. (1986/87). The new psychology of war and peace. *International Security*, 11, 175-186.
- Brams, S. J. (1985). *Superpower games*. New Haven: Yale University Press.
- Brodie, B. (1946). *The absolute weapon*. New York: Harcourt, Brace.
- . (1959). *Strategy in the missile age*. Princeton: Princeton University Press.

- _____. (1978). The development of nuclear strategy. *International Security*, 2.
- Brown, H. I. (1983). Incommensurability. *Inquiry*, 26, 3-29.
- Bull, H. (1967). *The anarchical society: A study of order in world politics*. New York: Columbia University Press.
- Bundy, M. (1983). The Bishops and the bomb. *New York Review of Books*, June 16.
- Chomsky, N. (1982). *Towards a new cold war*. New York: Pantheon Books.
- Craig, G. A., & George, A. L. (1983). *Force and statecraft: Diplomatic problems of our time*. New York: Oxford University Press.
- Deutsch, K. (1968). *The analysis of international relations*. Englewood Cliffs, N.J.: Prentice-Hall.
- Deutsch, M. (1958). Trust and suspicion. *Journal of Conflict Resolution*, 2, 265-279.
- _____. (1961). Some considerations relevant to national policy. *Journal of Social Issues*, 17, 57-68.
- _____. (1969). Conflicts: productive and destructive. *Journal of Social Issues*, 25, 5-41.
- _____. (1973). *The resolution of conflict: Constructive and destructive processes*. New Haven: Yale University Press.
- _____. (1983). What is political psychology? *International Social Science Journal*, 35, 221-236.
- Deutsch, M., & Shichman, S. (1986). Conflict: A social psychological perspective. In M. G. Hermann (Ed.), *Political psychology*. San Francisco: Jossey-Bass.
- Dyer, G. (1985). *War*. Toronto: Stoddart
- Eckhardt, W. (1983). Pioneers of peace research V: Anatol Rapoport: Apostle of collective rationality. *International Interactions*, 10, 213-245.
- Festinger, L. (1957). *A theory of cognitive dissonance*. Evanston, Ill: Row, Peterson.

- Fiske, S. T., & Taylor, S. E. (1984). *Social cognition*. Reading, Mass: Addison-Wesley.
- Freedman, L. (1981). *The evolution of nuclear strategy*. London: MacMillan Press.
- George, A., & Smoke, R. (1974). *Deterrence in American foreign policy*. New York: Columbia University Press.
- Gergen, K. J. (1973). Social psychology as history. *Journal of Personality and Social Psychology*, 26, 309-320.
- _____. (1982). *Toward transformation in social knowledge*. New York: Springer-Verlag.
- _____. (1985). *Psychological inquiry in an age of social epistemology*. Paper presented at the Boston Symposium in the Philosophy of Science, Boston, MA.
- _____. (1986). *On the social construction of knowledge*. Paper presented at Inaugural Conference of the Centre for Advanced Studies in Humanist Social Scientists, Kitchener-Waterloo, ON.
- Green, P. (1966). *Deadly logic: The theory of nuclear deterrence*. Ohio State University Press.
- Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.
- Hoffmann, S. (1986). On the political psychology of peace and war: A critique and an agenda. *Political Psychology*, 7, 1-21.
- Hovland, C. I., Janis, I. L., & Kelley, H. H. (1953). *Communication and persuasion*. New Haven, Conn: Yale University Press.
- Jones, E. E., & Davis, K. E. (1965). From acts to dispositions: The attribution process in person perception. In L. Berkowitz (Ed.), *Advances in experimental social psychology*, Vol. 2. New York: Academic Press.
- Jervis, R. (1976). *Perception and misperception in international politics*. Princeton: Princeton University Press.
- _____. (1978). Cooperation under the security dilemma. *World Politics*, 30, 167-214.

- _____. (1979). Deterrence theory revisited. *World Politics*, 31, 289-324.
- _____. (1982/83). Deterrence and perception. *International Security*, 7, 3-30.
- _____. (1984). *The illogic of American nuclear strategy*. Ithaca: Cornell University Press.
- _____. (1986). Cognition and political behavior. In R. Lau & D. Sears (Eds.), *Political cognition*. Hillsdale: Lawrence Erlbaum Associates.
- Jervis, R., Lebow, R. N., & Stein, J. G. (1985). *Psychology and deterrence*. Baltimore: Johns Hopkins University Press.
- Kahan, J. P., et al. (1983). *Preventing nuclear conflict: What can the behavioral sciences contribute?* RAND Note.
- Kahn, H. (1960). *On thermonuclear war*. Princeton: Princeton University Press.
- _____. (1962). *Thinking about the unthinkable*. New York: Horizon Press.
- Kelley, H. H. (1967). Attribution theory in social psychology. In D. Levine (Ed.), *Nebraska symposium on motivation*. Lincoln, Nebraska: University of Nebraska Press.
- Kelman, H. C. (1981). Reflections on the history and status of peace research. *Conflict Management and Peace Science*, 5, 95-110.
- Kennan, G. F. (1983). *The nuclear delusion*. New York: Pantheon Books.
- Kiesler, C. A., & Lucke, J. (1976). Some metatheoretical issues in social psychology. In L. Strickland, F. Aboud, & K. Gergen (Eds.), *Social psychology in transition*. New York: Plenum Press.
- Kissinger, H. A. (1957). *Nuclear weapons and foreign policy*. New York: Harper.
- Klar, Y., Bar-Tal, D., & Kruglanski, A. W. (in press). On the epistemology of conflicts: Toward a social cognitive analysis of conflict and conflict resolution. In W. Stroebe, A. W. Kruglanski, D. Bar-Tal, & M. Hewstone (Eds.), *Social psychology of intergroup and international conflict*. New York: Springer-Verlag.

- Klein, B. S. (1986). *Strategic discourse and its alternatives*. Paper presented to the International Studies Association, March, 1986.
- Knorr, K., & Rosenau, J. (Eds.) (1969). *Contending approaches to international politics*. Princeton: Princeton University Press.
- Kruglanski, A. W. (1980). Lay epistemo-logic -- process and contents: Another look at attribution theory. *Psychological Review*, 87, 70-87.
- _____. (in press). The making of science and making it in science: On the psychological underpinnings of research methodology. In A. W. Kruglanski, *Basic processes in social cognition: A theory of lay epistemology*. Plenum Press.
- _____. (1986). *Motivational bases of attributions*. Unpublished manuscript.
- Kruglanski, A. W., & Ajzen, I. (1983). Bias and error in human judgment. *European Journal of Social Psychology*, 13, 1-44.
- Kruglanski, A. W., & Klar, Y. (1987). A view from a bridge: Synthesizing the consistency and attribution paradigms from a lay epistemic perspective. *European Journal of Social Psychology*, 17, 211-241.
- Kuhn, T. S. (1970). *The structure of scientific revolutions*. Chicago: University of Chicago Press. First published in 1962.
- Larson, D. W. (1985). *Origins of containment: A psychological explanation*. Princeton: Princeton University Press.
- Lawrence, P. K. (1985). Nuclear strategy and political theory: A critical assessment. *Review of International Studies*, 11, 105-121.
- Lindskold, Svenn. (1978). Trust development, the GRIT proposal, and the effects of conciliatory acts on conflict and cooperation. *Psychological Bulletin*, 85, 772-793.
- MccGwire, M. (1985). Deterrence: the problem -- not the solution. *SAIS Review*, 5, 105-124.
- McKinlay, R. D., & Little, R. (1986). *Global problems and world order*. Madison: University of Wisconsin Press.

Mitchell, C. R. (1985). Conflict, war and conflict management. In M. Light & A. Groom (Eds.), *International relations: A handbook of current theory*.

_____. (1986). GRIT and gradualism -- 25 years on. *International Interactions*, 13, 59-90.

Morawski, J. G., & Goldstein, S. E. (1985). Psychology and nuclear war: A chapter in our legacy of social responsibility. *American Psychologist*, 40, 276-284

Morgan, P. M. (1977). *Deterrence: A conceptual analysis*. Beverly Hills: Sage Publications.

_____. (1985). Saving face for the sake of deterrence. In Jervis et al. (Eds.), *Psychology and deterrence*.

Morgenthau, H. J. (1964). The four paradoxes of nuclear strategy. *American Political Science Review*, 58, 23-35.

Nisbett, R. E., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, N.J.: Prentice-Hall.

Osgood, C. E. (1962). *An alternative to war or surrender*. Urbana: University of Illinois Press.

Petty, R. E., & Cacioppo, J. T. (1981). *Attitudes and persuasion: Classic and contemporary approaches*. Dubuque, Iowa: W. C. Brown.

Rapoport, A. (1960). *Fights, games, and debates*. Ann Arbor: University of Michigan Press.

_____. (1964). *Strategy and conscience*. New York: Harper and Row.

_____. (1967a). *Introduction to Clausewitz' On war*. London: Penguin Books.

_____. (1967b). Escape from paradox. *Scientific American*, July, 50-56.

_____. (1984). Preparation for nuclear war: The final madness. *American Journal of Orthopsychiatry*, 54, 524-529.

Rapoport, A., & Chammah, A. M. (1965). *Prisoner's dilemma: A study in conflict and cooperation*. Ann Arbor: University of Michigan Press.

- Rorty, R. (1982). *Consequences of pragmatism*. Minneapolis: University of Minnesota Press.
- Russett, B. M. (1963). The calculus of deterrence. *Journal of Conflict Resolution*, 7, 97-109.
- Sampson, E. E. (1978). Scientific paradigms and social values: Wanted -- a scientific revolution. *Journal of Personality and Social Psychology*, 36, 1332-1343.
- _____. (1981). Cognitive psychology as ideology. *American Psychologist*, 36, 730-743.
- Schell, J. (1982). *The fate of the earth*. New York: Alfred A. Knopf.
- Schelling, T. C. (1960). *The strategy of conflict*. Cambridge: Harvard University Press.
- _____. (1966). *Arms and influence*. New Haven: Yale University Press.
- Smith, M. B. (1983). The shaping of American social psychology: A personal perspective from the periphery. *Personality and Social Psychology Bulletin*, 9, 165-180.
- Snyder, G. H., & Diesing, P. (1977). *Conflict among nations*. Princeton: Princeton University Press.
- Steiner, B. H. (1984). *Using the absolute weapon: Early ideas of Bernard Brodie on atomic strategy*. ACIS Working Paper #44.
- Steinbruner, J. D. (1974). *The cybernetic theory of decision*. Princeton: Princeton University Press.
- Strickland, L. H., Aboud, F. E., & Gergen, K. J. (Eds.) (1976). *Social psychology in transition*. New York: Plenum Press.
- Swingle, P. (1986). *Conflict resolution*. Paper presented at the Quest for Peace in Our Day conference, Ottawa.
- Tetlock, P. E. (1986). Psychological advice on foreign policy: What do we have to contribute? *American Psychologist*, 41, 557-567.
- Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5, 207-232.

- _____. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124-1131.
- Vasquez, J. (1979). Colouring it Morgenthau. *British Journal of International Relations*, 18, 210-228.
- Waltz, K. N. (1959). *Man, the state, and war*. New York: Columbia University Press.
- _____. (1979). *Theory of international politics*. London: Addison-Wesley.
- Weinberger, C. W. (1986). U.S. defense strategy. *Foreign Affairs*, 64, 657-697.
- White, R. K. (Ed.) (1986). *Psychology and the prevention of nuclear war*. New York: New York University.
- Wohlstetter, A. (1959). The delicate balance of terror. *Foreign Affairs*, 37.

END

07.07.189

FIN