

Beyond Lakatos: New Approaches to Appraising Progress in International Relations Theory

By Hiroaki Abe (ha2233@columbia.edu)

Paper for Mini APSA, Department of Political Science, Columbia University
Spring 2010

Introduction

Theorists of International Relations (henceforth IR) do not often enough ask themselves how well they are doing in what they are doing. And although the construction of theory is inherently a more reflexive exercise than the practice of applied research, (though these two dimensions necessarily imply each other), insofar as theorists as much as anyone else tend to get caught up in the practice of what they are doing – they just as much need to step back occasionally and assay the fruits of their work to question its meaning. In this light, the practice of becoming reflexive about one’s work opens up a new dimension (i.e. the meta-level) to its practice, just as ‘learning how to learn’ opens up a new dimension to the practice of learning.¹ Therefore, this added awareness that comes from turning around one’s analytical tools and analyzing the process of theorizing in its own right is the valuable fruit that comes from the study of theoretical development in IR. And it is with the aim of contributing to the store of reflexive knowledge in the discipline that I have undertaken the inquiries that follow.

Despite the meta-theory of IR being under-theorized, an implicit understanding of the nature of theoretical development in the discipline is nevertheless held by many of its practitioners. This is as it should be, for no observation or fact is ever theory-neutral, and theories at one or another level of specificity are always already built into the act of seeing. The implicit understanding of the development of IR theory held by most is that of an overall progress and cumulative development, but such an understanding remains pre-reflexive most of the time and is seldom questioned. This sense of progress is of course intimately linked with the rationalist spirit of our times. Hans Morgenthau, in *Scientific Man vs. Power Politics*, writes that the prevalent spirit of Western thought since the turn of the eighteenth century has been dominated by a rationalist mode of thinking which “expresses itself philosophically in the belief

in continuous progress and in the trivial optimism for which life dissolves into a series of little hurdles which, one after the other, increasing skill cannot fail to overcome.” (Morgenthau, 1946: 207)² And this deep-set belief in progress is also reinforced by the increasing technical mastery of man over his environment. As the historian J.B. Bury has noted:

“This uninterrupted series of technical inventions, proceeding concurrently with immense enlargements of all branches of knowledge, has gradually accustomed the least speculative mind to the conception that civilization is naturally progressive, and that continuous improvement is part of the order of things.” (Bury, 1932: 332)

Nevertheless, the implicit understanding of IR theory as subject to a path of progress and cumulative development does become explicit at various moments in disciplinary discussions, such as in the forum on the progressiveness of neo-traditional approaches to realism in the December 1997 issue of the *American Political Science Review*,³ and in the book resulting from a conference on meta-theory in IR, entitled *Progress in International Relations Theory (PIRT)* and published in 2003.⁴ Looking through these papers, the reader gets a sense that most of these authors believe theoretical research in IR to be progressive, or hold normative views of progress as an important (perhaps *the* important) goal for work in the field to attain. Nevertheless it does seem a problem that the concept of disciplinary progress used in most of these accounts is insufficiently robust, being based as they are overwhelmingly on Imre Lakatos’s notion of *progressive problem-shifts* in scientific research programs – without being attentive enough either to critiques of progress narratives in the philosophy of social science, or to alternate characterizations of progress in the philosophy of science. I intend in what follows to rectify the

current over-dependence of IR meta-theory on Lakatosian methodology by taking up the suggestions of two of the contributors to the 2003 *PIRT* volume, who suggest the need for multiple frames in appraising IR theory (Vasquez) and the need to reintegrate post-Lakatosian developments in the philosophy of science back into appraisals of the field of international relations (Bennett).

In this paper, I will ask the question of whether theoretical development in IR is best understood through a Lakatosian framework. In the opening section of this paper, I put forward an account of the Lakatosian paradigm in the context of some of the debates between Lakatos and his main predecessors and interlocutors in the philosophy of science. In the second section, I discuss how well the Lakatosian paradigm manages to epitomize the empirical development of IR when compared to some of the other alternative frameworks which are available in the philosophy of science. Although the Lakatosian rejection of naïve falsification is shown to be mirrored in the way that IR theory develops, his notion of theories as being superseded by other theories which perfectly encompass earlier ones is found to be problematic. Furthermore, I discuss the common misinterpretation of Lakatos which reads him as providing appropriate tools with which to choose between competing theories. It emerges upon closer analysis that Lakatos is not primarily interested in the question of theory choice and its rationality, being in fact quite happy to keep a historical tally of the relative progressiveness of competing research programmes. To the extent that he has normatively oriented concerns, Lakatos is intent to specify the conditions under which knowledge growth can be facilitated to the greatest extent possible.

Lakatos and his predecessors

In order to understand the ideas of Imre Lakatos in their proper context, it will be necessary to frame them against the background of the debates undertaken with his interlocutors – out of which Lakatos’s ideas were formulated. In this first section, we will look briefly at the ideas of Karl Popper, Thomas Kuhn, and Stephen Toulmin before giving a more extensive account of the Lakatosian method.

Karl Popper was one of the most important influences on the work of Imre Lakatos. This is why any evaluation of the ideas of Lakatos must start by looking at the connections between the ideas of these two philosophers of science. Popper’s early work in the philosophy of science argued against a view of science current at the time and associated with an inductive approach, which saw science as a process of confirming theories to be true. Instead, Popper argued that the truth content of general laws cannot be known for certain, and therefore he took it that the only way to ensure that science moves closer to the truth (and achieves verisimilitude) is to falsify our hypotheses and correct our errors. Therefore, Popper writes, we should:

“[L]ook for theories which, however fallible, progress beyond their predecessors; which means that they can be more severely tested, and stand up to some of the new tests...we for our part are satisfied that the rationality of a theory lies in the fact that we choose it because it is better than its predecessors; because it can be put to more severe tests; because it may even have passed them, if we are fortunate; and because it may, therefore, approach nearer to the truth.” (Popper, 1963: 336)

Falsification through theory testing is one of the principal pillars of the Popperian position, as are his injunctions that theories should: i) be able to explain away anomalies that their predecessors could not, ii) throw up novel facts that their predecessors did not predict, and iii) conserve the success of previous theories by accounting for everything that their predecessors could account for. It is this Popperian view of the growth of knowledge in science that inspired the work of both followers and critics to build on and/or challenge his insights.⁵ But this view of the growth of knowledge is also a rather paradoxical one, insofar as it posits a dual process whereby wrongly held views are overthrown through falsification on the one hand, but where the knowledge content of earlier theories is nevertheless preserved in later theories despite an ongoing process of conjecture and refutation. It should be remembered that Popper's acceptance in his later thought of revolutionary overthrow as a necessary component of the growth of knowledge can, at least in part, be attributed to the influence on Popper of Thomas Kuhn and his theory of scientific revolutions. As Popper has written on the role of revolutions,

“[P]rogress in science, although revolutionary rather than merely cumulative, is in a certain sense always conservative: a new theory, however revolutionary, must always be able to explain fully the success of its predecessor. In all those cases in which its predecessor was successful, it must yield results at least as good as those of its predecessor and, if possible, better results.” (Popper, 1981: 94)

This view of progress as cumulative growth in the face of continual revolutions in knowledge cannot be easily reconciled with a simple linear picture of the evolution of science. But here we must keep in mind that Popper's main goal was to build a normative method for the

growth of knowledge and a theory of rational choice between theories. As Diesing has well put it, the central task for Popper, is to answer the question of “What should science do to get better, to improve itself?” (Diesing, 1992: 30) From this point of view, Popper’s philosophical project is mainly about proposing a normative method for what science *ought* to be in order that the growth of knowledge is more closely attained. This normative goal for Popper takes precedence over both the descriptive task of characterizing existing scientific practice, and the analytical task of abstracting the logical essence of science out of its varying practices, (Diesing, 1992: 30). This is a point that is not often enough remembered today.

The notion of falsification championed by Popper has proven to be an important methodological injunction for researchers in the social sciences, and even the authors of recent methodological textbooks (such as that of King, Keohane, and Verba’s *Designing Social Inquiry*) make a number of specific injunctions about how social science should be done that clearly owe much to the Popperian view. These injunctions center round the notion that social science theories should be falsifiable. To this end, King *et al.* suggest that theories should be capable of generating as many observable implications as possible, (to facilitate falsification), and that theories should be as concrete as possible. Furthermore, they suggest that social science research (and presumably also social science theories) should make a progressive contribution to the state of research in a particular field by guarding against duplication and aiming to contribute to the cumulative development of scientific knowledge. Although King *et al.* do not explicitly define what they mean by scientific progress, we think it is safe to assume that they have in mind something like the Popperian view of it, which we have just elaborated.

In contrast to the optimistic view of Popper, which sees there to be an ongoing process of cumulative knowledge growth and self-correction, Thomas Kuhn’s intuition in proposing his

own model of the development of scientific knowledge was to correct what he saw to be the disjuncture between Popper's optimistic views and the actual historical process of scientific development. Thus Kuhn argues that the process of scientific development is largely a discontinuous one whereby scientific revolutions overturn previously ascendant scientific paradigms and bring with them new problems and explanations which cannot be directly compared with their predecessors. For Kuhn, theories that are established through scientific revolutions are based on different sets of motivating problems and fundamental assumptions from those of their predecessors. This is the reason why they cannot be directly compared, and are therefore *incommensurable*. He writes:

“To the extent...that two scientific schools disagree about what is a problem and what a solution, they will inevitably talk through each other when debating the relative merits of their respective paradigms. In the partially circular arguments that regularly result, each paradigm will be shown to satisfy more or less the criteria that it dictates for itself and to fall short of a few of those dictated by its opponent.” (Kuhn, 1996: 109-110)

One implication of Kuhn's ideas about scientific paradigms, is that a particular scientific discipline is normally dominated by one particular paradigm, where this paradigm – whether it be the bodies of research built around the theories of Newton, Lavoisier, Maxwell, Einstein, or some other – holds a hegemony over an entire field in terms of predetermining the types of questions worthy of investigation, the types of experimental techniques to be used, and the types of theories and explanations acceptable as answers to empirical problems. The diversity of paradigms is therefore something that is manifest diachronically, but seen through a synchronic

lens such diversity is scarce as Kuhnian disciplines are dominated by single overwhelmingly accepted paradigms. Nevertheless, according to Kuhn, the hegemonic nature of paradigms within scientific disciplines is something that actually facilitates the attainment of progress within individual paradigms. This is the case because working within the limits of *normal science*, (i.e. the state of affairs where a single paradigm is dominant in a field), allows researchers to take fundamental assumptions for granted in their work, concentrating instead on “extending the knowledge of those facts that the paradigm displays...by increasing the extent of match between those facts and the paradigm’s predictions, and by further articulation of the paradigm itself.” (Kuhn, 1996: 24)

For Kuhn, it is the lack of hegemonic paradigms (i.e. the lack of disciplinary consensus) within certain disciplines that hinders their ability to make progress. This is also another way of saying that certain disciplines have not yet managed to consolidate themselves and enter the stage of becoming normal sciences. And in fact Kuhn thinks that it is nothing more or less than the very lack of such progress (defined semantically – i.e. conventionally) that determines whether a field has achieved the state of becoming a science.⁶ But besides making the obvious point that stagnant disciplines are not often considered to be sciences, there are also good institutional reasons as to why, for Kuhn, pre-paradigmatic disciplines have difficulty in achieving progress. When there is lacking a disciplinary consensus and thus where there exist a number of different paradigms competing to answer a set of common problems, Kuhn thinks that the assumptions about first principles that can be taken for granted when practicing normal science cannot be made – and thus it is often necessary for researchers to build intellectual structures from the bottom-up. The lack of such a disciplinary consensus in a pre-paradigmatic field also spreads knowledge outward breadth-wise rather than allowing it to accumulate depth-

wise – due to the relative freedom for individual researchers to choose which paradigm they wish to work in, and which questions they wish to answer. On the other hand, Kuhn infers that in the normal sciences deepening knowledge may facilitate progress, but intellectual novelty and the creativity that is made possible by the random variation and auto-didacticism that are prevalent in pre-paradigmatic fields will necessarily suffer.⁷

Kuhn's notion of theory choice, in contrast to the rationalistic views of Popper, is an evolutionary one where theories are chosen through a process of natural selection and where theories progressively better articulated and specialized to solve problems are left standing over time. This problem solving emphasis, is also the mechanism that for Kuhn explains the retention of knowledge even across scientific revolutions. As he puts it, "because the unit of scientific achievement is the solved problem and because the group knows well which problems have already been solved, few scientists will easily be persuaded to adopt a viewpoint that again opens to question many problems that had previously been solved." (Kuhn, 1996: 169) This is perhaps the one moment where Kuhn appears closest to Popper in his views, insofar as he – like Popper – attempts to maintain a picture of the cumulative growth of knowledge despite the occurrence of scientific revolutions.

Importantly, the process of the evolutionary development of scientific knowledge does not for Kuhn tend towards a telos, or goal. Thus, once again in contrast to Popper, Kuhn is critical of the notion that scientific theories progressively tend closer and closer to the truth.⁸ Drawing an analogy with Darwin's evolutionary views, Kuhn believes that there does not need to be the guarantee of a final end-point for us to be able to discern progress in science. Instead, progress can still be recovered from a thoroughly relativistic view of the development of science.

This is because incremental progress can still be discerned from any vantage point in time. As he writes:

“[N]othing that has been or will be said makes it a process of evolution *toward* anything...we are all deeply accustomed to seeing science as the one enterprise that draws constantly nearer to some goal set by nature in advance. But need there be any such goal? ...does it really help to imagine that there is some one full, objective, true account of nature and that the proper measure of scientific achievement is the extent to which it brings us closer to that ultimate goal?” (Kuhn, 1996: 170-171)

As things stand, the views of both Popper and Kuhn about the cumulative growth of knowledge are in one respect internally contradictory, and they can be seen to be afflicted by the same problem from different directions. Popper, on the one hand, acknowledged that scientific revolutions are an important part of the evolution of science – a corollary of his views regarding the growth of knowledge via falsification. But although the notion of a scientific revolution necessarily implies a certain degree of knowledge loss, (as it is not possible that all the knowledge content of earlier theories is retained by later ones), Popper nevertheless wished to retain a commitment to the view that knowledge is cumulative in an absolute sense across time. Kuhn, on the other hand, was faced with the opposite problem. He was committed to a view of the primacy of scientific revolutions over the progress evident during periods of normal science, and also the notion of incommensurability existing between pre-revolutionary and post-revolutionary paradigms. Nevertheless, (and perhaps partially influenced by Popper), Kuhn wished to retain some sense of the cumulative growth of knowledge over time.

Stephen Toulmin presents us with a way of seeing the evolution of knowledge which to a certain extent reconciles the internal tension evident in the views of Popper and Kuhn. Reacting against both Popper's assumption of a deep continuity between scientific theories, and the Kuhnian assumption of a deep underlying discontinuity, Toulmin developed an evolutionary model of the development of scientific knowledge which – rather than assuming either a necessary continuity or discontinuity at the heart of the model – gives a more nuanced account of scientific change as consisting of both elements of continuity and elements of variability. He writes that:

“[F]rom one historical cross-section to the next, the actual ideas transmitted display neither a complete breach at any point – the idea of absolute ‘scientific revolutions’ involves an oversimplification – nor perfect replication, either. The change from one cross-section to the next is an *evolutionary* one in this sense too: that later intellectual cross-sections of a tradition reproduce the content of their immediate predecessors, as modified by those particular intellectual novelties which were selected out in the meanwhile – in the light of the professional standards of the science of the time” (Toulmin, 1967: 466)

Toulmin's critique of Kuhn addresses the notion of revolutionary science that Kuhn established. Although he agrees that Kuhn was right to focus on discontinuity due to the misleading empirical picture implied by Popper's view of knowledge as consisting of an absolute continuity and growth,⁹ Toulmin points out that the notion of an absolute breach of continuity that is implied by the concept of incommensurability is unsatisfactory because there is always a certain degree of continuity in the structure of knowledge before and after scientific revolutions and across paradigms. Nevertheless, Toulmin argues that this continuity cannot be discerned

without a change in perspective. He thinks that the notion of absolute discontinuity that Kuhn associates with scientific revolutions results from an insufficiently large sample size, (the population here consisting of all cases of scientific revolutions), which leads to the mistaken impression that the breach implied by a revolution is in fact more radical than it really is. Thus, despite what might seem to be differences between the philosophical conceptions of Kuhn and Toulmin, we are led to discern an important continuity between their views insofar as Toulmin's evolutionary perspective is in a sense the view from a great height of phenomena that Kuhn is looking at from ground level. This difference in magnification, so to speak, allows Toulmin to see that Kuhn's revolutions are not in fact unique but instead represent a class of events, and from this distance the interweaving of continuous and discontinuous elements also comes into view.

The three perspectives that we have looked at so far, (i.e. those of Popper, Kuhn, and Toulmin), represent some important basic positions regarding the nature of scientific change. When we abstract out the basic logic from these philosophies, we can see that Popper represents a view of change as continuity and increase, whereas Kuhn presents a view of change as discontinuity and constancy,¹⁰ and Toulmin takes scientific change to be continuous and constant. If we extrapolate from the theories that we have looked at so far, we can see that there are two basic dimensions of a theory of change, and all of these theories stake a position along each of these dimensions. The first dimension is that of the nature of change, which so far appears to be captured by the continuous ↔ discontinuous continuum, whereas the second dimension represents its directionality along the continuum of increase ↔ constancy ↔ decrease. We can perhaps, for the sake of completeness, speculate that continuity can be subdivided into three forms, (line, circle, and spiral), whereas discontinuity can be seen to consist of a series of points

which – when seen from afar – meld into the nature of a line. But to digress at length on these matters would take us too far from the matter at hand.

All theories of scientific change, it would seem to me, are composed of a value along each of the two dimensions that I have just outlined, and *progress* defined in these terms is, I would argue, a subset of those values which combine continuity with increase. Certainly the Popperian notion of progress is incompatible with either a mere constancy or decrease of knowledge content across theories, and discontinuity would undermine any standard by which we could measure directionality and thus must also be denied. However, Kuhn's views might at first glance seem to imply that progress can exist in the face of both discontinuity and the lack of increase. As we have already discussed above, this impression is mistaken. Insofar as Kuhn tends to limit discussions of progress to developments that take place within individual paradigms in normal science, it would be most consistent to interpret him as denying that progress is sustained across the gulf created by revolutionary change. As far as Kuhn's views about the growth of scientific knowledge are concerned, we have seen that they are conflicted, but Kuhn ultimately hews close to the Popperian line which associates scientific progress with a cumulative growth in knowledge. Therefore the development of knowledge is not only continuous but growing under conditions of normal science, and it is only when we attempt to apply the term 'progress' to inter-revolutionary shifts – an application for which it was not intended – that Kuhn is pushed into seemingly contradictory views.

It is now necessary to discuss these preceding views and their relationship with the ideas of another philosopher of science who has speculated on the nature of scientific change, (and who is the main topic of this section), Imre Lakatos. Lakatos's *methodology of scientific research programmes* (henceforth MSRP) is a development of many of Popper's themes and at

the same time incorporates some of Kuhn's strongest critiques of the Popperian view. The fundamental idea is that scientific theories are embedded within a number of competing research programmes, which are either progressing or degenerating. Firstly, we should notice that the basic unit which is subject to change is no longer scientific theories per se, but the research programme – which consists of a 'hard-core' and a 'positive heuristic.' This shift in Lakatos's MSRP of the basic unit of change from the individual theory to the research programme, is an important critique of the earlier Popperian and Kuhnian emphasis on single theories as the primary unit of analysis for understanding theoretical change. Lakatos emphatically holds that a logically related series of theories, or *scientific research programme*, is the most appropriate unit of analysis from which progress or its absence should be detected.¹¹

Furthermore Popper's 'naïve' falsification criterion is criticized, for according to Lakatos a theory is never simply falsified – especially if there is lacking a better alternate theory that can take its place. Thus within a research programme, "a theory can only be eliminated by a better theory, that is, by one which has excess empirical content over its predecessors," (Lakatos, 1981: 118), and for Lakatos the excess content criterion comes to play a more important role than that of falsification in determining scientific change. As he has written, "progress is marked by instances of verifying excess content rather than by falsifying instances." (Lakatos, 1981: 118) In a sense, Lakatos's renewed emphasis on the verification of excess content over the falsification of content counts equally as a critique of Popper's falsification and of Kuhn's view of scientific revolutions and the overthrow of old conceptual paradigms. Instead of there being an overthrow and a reinstatement, as both Popper and Kuhn imply in their respective ways, for Lakatos the process of continuous growth is underlined much more strongly and the need for creative

destruction is made redundant because no theory is ever thrown away before a replacement has already been found.

Theories within a Lakatosian research programme are related to each other because they all accept a conventionally defined *hard core* consisting of assumptions that are shared and not normally open to challenge during the course of conducting ordinary research. A negative heuristic directs criticism away from the hard core of a research programme, thus protecting the hard core. In this notion of a negative heuristic, one can discern the debt that Lakatos owes to Kuhn. For Popper, all parts of a theory were up for grabs and open to question, but it was Kuhn who contended that normal science more often entails the effort to prove theories right (and then more right) than it does the effort to prove theories wrong. Thus the negative heuristic of Lakatos is very much akin to what the practice of normal science looks like for Kuhn. The practitioners of a research programme are also guided by a positive heuristic, which for Lakatos entails building a protective belt of auxiliary hypotheses around this hard core, and attempting to amend the protective belt in such a way that anomalies and criticisms are foreseen and then turned victoriously into confirmatory examples for the research programme under attack.

In other words, a successfully operating positive heuristic ensures that a research programme is continually able to turn anomalies into supporting instances, creating excess empirical content from out of these anomalies and thus maintaining its momentum and ensuring progressive shift. However, once the positive heuristic no longer operates properly the questions arising from anomalies come more and more to penetrate the hard core – the momentum of the research programme is lost and it starts to fall into a process of degenerating drift. We should here acquaint ourselves with Lakatos's precise views of scientific progress, so let us see what Lakatos has to say about the progressive and degenerative shift of research programmes:

“A research programme is said to be *progressing* as long as its theoretical growth anticipates its empirical growth, that is as long as it keeps predicting novel facts with some success (*‘progressive problemshift’*); it is stagnating if its theoretical growth lags behind its empirical growth, that is, as long as it gives only *post hoc* explanations either of chance discoveries or of facts anticipated by, and discovered in, a rival programme (*‘degenerating problemshift’*). If a research programme progressively explains more than a rival, it ‘supersedes’ it, and the rival can be eliminated (or, if you wish, ‘shelved’).” (Lakatos, 1981: 117)

Thus, for Lakatos, a research programme is progressive if subsequent theories in the series into which it is embedded are not only theoretically progressive, (meaning that a subsequent theory has some excess empirical content over its predecessor), but also empirically progressive as well, (where empirical progress consists of the corroboration of this excess content, which takes the form of novel facts).

However, interestingly enough, despite the numerous ways in which Lakatos departed from the ideas of his predecessors, when we evaluate MSRP according to the two key dimensions of all theories of change we outlined above, we can see that Lakatos’s conception of scientific progress is closest to that of Popper in its essentials. This is the case insofar as MSRP posits both a fundamental continuity in the development of research programmes, and a progressive increase in the content of scientific knowledge. We have already seen how Lakatos takes it that the verification of excess theoretical content is more important than its falsification, but his notion of sophisticated falsification – rather than being a straightforward refutation of the views of Popper – should more properly be seen as its denouement. We have discussed above how Popper’s falsification sat in tension with his view of there being an absolute growth of

knowledge, and Lakatos's notion of sophisticated falsification in fact smoothes over the latent tension in Popper's thought, allowing there to emerge a more consistent picture of progress in scientific knowledge as an absolute increase. Furthermore, the Lakatosian picture of multiple research programmes existing at once – where forward-moving momentum shifts from one research programme to another as they respectively progress, stagnate and regress much like athletes running a long distance race – is a view which does away with the need for diachronic discontinuity between paradigms necessitated by Kuhn's view of long periods of progressive normal science interspersed with short bursts of revolutionary science. Therefore MSRP recovers a fundamental continuity in the growth of scientific knowledge that had been put into question by Kuhn.

The appraisal of the progressiveness of research programmes cannot for Lakatos be an activity that is solely guided by measurement against a set of evaluative criteria alone, (contra Popper), nor is it possible merely by looking for cumulative growth in a sequence of theories within one research programme, (contra Kuhn). This is because evaluation also necessarily implies a two-pronged comparative process through which both the degree of progress within particular research programmes, and the *relative* progressiveness of different research programmes when compared with each other, is appraised. Therefore this proposed shift from the evaluation of theories by comparing them with a set of facts to the evaluation of research programmes via a comparison of their relative problem solving effectiveness, is another important aspect in which the ideas of Lakatos move beyond the work of his predecessors. The comparative method of MSRP also makes explicit the fact that within most disciplines there are more than one research programmes (equivalent to paradigms for Kuhn) at any one time, and these research programmes are not incommensurable insofar as they attempt to explain empirical

problems that they see in common as being problems. It is these competing research programmes that must be compared with each other if we are to be able to evaluate whether they are undergoing progressive or degenerative change when compared to their rivals.

Ian Hacking, one of Lakatos's most perceptive interpreters, has argued convincingly that Lakatos was attempting to install a theory of the growth of knowledge as the centerpiece of his philosophy of science, substituting it in the place of a representational theory of truth. As he has put it, for Lakatos "the point is not that there is knowledge but that there is growth." (Hacking, 1981: 129) Therefore, this view of MSRP implies that even where a progressive problemshift is taking place, there is no way that one can tell whether the research programme is tending closer to the truth or not, contrary to what Popper's view of verisimilitude might lead us to believe. This is a somewhat controversial assertion that puts Lakatos views close to those of Kuhn, who (as we have seen) also wishes to do away with the need to ground an account of paradigmatic development on the assumption that theories progressively tend towards truth over time. But despite holding problematic implications, we believe that this is indeed a correct reading of Lakatos.

Confirmation of this interpretation is shown in passages such as those in which Lakatos writes that "with sufficient resourcefulness...any theory can be defended 'progressively' for a long time, even if it is false." (Lakatos, 1981: 117) In other words, here Lakatos is asserting that the progressiveness of a research paradigm is not necessarily connected to its being true or false. Furthermore, not only is progressiveness decoupled from verisimilitude, but degenerating problemshifts are no longer necessarily terminal states but oftentimes temporary ruts which a research programme can transcend by recovering momentum and progressivity at a later date. Thus research programmes are never definitively falsified, and Lakatos states that:

“One *must* realize that one’s opponent, even if lagging badly behind, may still stage a comeback. No advantage for one side can ever be regarded as absolutely conclusive. There is never anything inevitable about the triumph of a programme. Also, there is never anything inevitable about its defeat.” (Lakatos, 1981: 118-119)

The drift of this argument leads us to the even more contentious conclusion that Lakatos was not interested in the normative evaluation of research programmes at all, but rather was happy merely to track the relative progress or regress of research programmes in the past. In other words, Lakatos offers no help in deciding which research programmes we *ought* to choose in the present. Rationality here becomes merely an answer to the question of whether the most progressive research programmes (i.e. research programmes where the growth of knowledge through the prediction of new facts is occurring) were actually chosen from amongst the contenders in actual historical instances. Hacking implies that Lakatos’s theory is essentially one that replaces *ought* with *is*, that in other words it is a theory which offers us a method “for examining and sorting past sequences of theories to see whether they are degenerating or progressive,” and he goes on further to say that Lakatos’s MSRP “provides no forward-looking assessments of present competing scientific theories.” (Hacking, 1981: 132, 133)

One important caveat here, is that though the question of scientific rationality is assimilated into the historical question of whether progressive research programmes were actually chosen by practitioners in a field at a particular point in time, nevertheless Lakatos ultimately retains an implicit normative focus similar to that of Popper insofar as he is concerned to stipulate the conditions in which a growth of knowledge in science can occur. In other words,

though the normative question of how progressive (or true) theories can best be chosen is of no interest to him, Lakatos is nevertheless interested in the normative question of how theories can be constructed such that the likelihood of knowledge growth in science can best be ensured. To this question, the answer that Lakatos proposes is a rather self-evident one: to make sure that theories are constructed such that they have excess empirical content over the extant theories in a field. This practice increases the probability that progressive problemshifts, (and not degenerative ones), will characterize the development of scientific research. There is evidence for such views in the rule for the acceptance of theories that Lakatos puts forward, which states that “for the sophisticated falsificationist a theory is ‘acceptable’ or ‘scientific’ only if it has corroborated excess content over its predecessor.” In other words, if a prior process of elimination takes place such that only ‘acceptable’ theories with excess content are considered to be viable replacements for their predecessors, then the growth of science would have been ensured – not at the moment of theory choice, but at an earlier point before such choices are even made.

Lakatos and Theories of IR

Having now given an account of Lakatos's philosophy of science and the other earlier theories which contributed to the development of his approach, let us next turn to the task of evaluating whether theoretical development in IR is best understood through the Lakatosian method of MSRP. In order to accomplish this task, we will critically examine some applications of the Lakatosian framework to the evaluation of IR theory. Two motivating interests will guide the following enquiries. Firstly, I am interested in understanding why Lakatosian philosophy of science (over the theories of his competitors) has attracted such a strong interest among theorists of IR, and secondly, I am interested in understanding what the limitations are of current applications of the Lakatosian MSRP to IR theory. In the following discussion I will focus in particular on two sources that I believe to be representative of the range of applications of the Lakatosian framework in IR: the forum on the progressiveness of neo-traditional approaches to realism in the December 1997 issue of the *American Political Science Review*¹²; and the book entitled *Progress in International Relations Theory (PIRT)*, which resulted from a 1999 conference on meta-theory in IR.¹³

Before anyone can properly assess whether Lakatosian progress is an apt characterization of disciplinary development in IR, it must first be ascertained exactly what the appropriate unit of analysis is that progress may or may not be an apt characterization of. Should the unit of analysis be the individual theories of IR, groups of IR theories collected under the banner of either a research paradigm or a research programme, or some other appropriate unit?

Kenneth Waltz's answer is that the proper units of analysis for evaluating disciplinary progress are the individual theories of international relations. We should of course be clear about what exactly Waltz takes a theory to be. As he puts it:

“A theory depicts the organization of a realm and the connections among its parts. The infinite materials of any realm can be organized in endlessly different ways. Reality is complex; theory is simple. By simplification, theories lay bare the essential elements in play and indicate necessary relations of cause and interdependency – or suggest where to look for them.” (Waltz, 1997: 913)

Waltz, criticizing what he calls Lakatos’s “displacement” of the problem of theory testing onto that of testing research programmes, asserts that the task of evaluating theories takes priority over the task of evaluating groups of theories. The reasons he gives are as follows: first, research programmes are secondary to theories because they are usually built around powerful theories and not the other way around; second, the problem of evaluating theories on their own must be tackled whether or not research programmes are successfully compared; and third, the positive evaluation of a theory is usually a necessary step for a theory to become successful and spawn a research programme – all of which for Waltz goes to show that the evaluation of theories is prior to that of research programmes. (Waltz, 2003)

In other words, when Waltz asserts that theories are the appropriate level of analysis on which questions of disciplinary progress should be evaluated, he is in agreement with the emphasis that Popper and Kuhn place on individual theories as being the primary units of appraisal. When even such an eminent theorist as Waltz disagrees with Lakatos about the main units of analysis in appraising the theories of IR, we are entitled to ask whether there are any strong reasons why the Lakatosian emphasis on groups of theories as the unit of should be retained.

It seems to us that the main advantage of the Lakatosian focus on research programmes is twofold. Firstly, when compared to the Kuhnian notion of paradigms, Lakatos's picture of the co-existence of multiple research programmes provides a way of describing the empirical terrain of IR theory that both *prima facie* seems descriptively accurate, and which allows IR to retain for itself status as a mature science. Contained within Kuhn's concept of the paradigm, as we have seen above, is the implicit notion that in a normal science the paradigm generally holds a position of disciplinary hegemony. However, when a discipline has not yet managed to streamline itself and become a normal science, Kuhn takes it that the co-existence of multiple and fundamentally divergent paradigms is the norm.

IR theory on the one hand is characterized by an irreducible heterogeneity, such that the Kuhnian image of the hegemonic paradigm does not tally well with the important typologies of the disciplinary terrain. For instance, the 'levels of analysis' approach to categorizing the theories of IR posits there to be a complementarity between the theories that locate their explanatory variables at different levels, (the individual, the state, and international system, respectively), rather than there being a competitive and zero-sum relationship between them. As Waltz has written in his early work, the inclusion of factors at each level of analysis is necessary for a satisfactory analysis to be made.¹⁴ Even in his later work, Waltz defines the systems approach that he advocates to consist of an explanation of "how the structure of the system affects the interacting units and how they in turn affect the structure." (Waltz, 1979: 40) However, a caveat must be added, which is that empirical generalizations made at one level of analysis cannot merely be combined with those at another level in an additive fashion, but as Singer notes "a prior translation from one level to another must take place." (Singer, 1961: 91) Nevertheless, if interdependence (and not independence) is held to characterize the relations

between different classes of theories and ‘levels’ of analysis, this clearly cuts against the logic (or is it a metaphor?) of statistical control where it is necessary to establish the independent effect of each individual level/theory against that of the others. And it would seem that an operative logic of appraising theories equivalent to statistical control is necessary for the explanatory force of individual IR theories to be established and ultimate ‘winners’ chosen – as the Kuhnian approach would require.

The other, and perhaps more widespread typology whereby theories of IR are divided into ‘isms,’ (such as for instance neo-realism, neo-liberalism, and constructivism), also more closely resembles the Lakatosian image of multiple research programmes inhabiting a disciplinary space than it does other alternative accounts. These contrasting research programmes are characterized by conflicting epistemological assumptions which cannot be transcended,¹⁵ and thus the likelihood of a disciplinary consensus emerging on the basis of the final victory of one of these research programmes seems highly unlikely. At the same time, the existence of multiple research programmes that disagree on fundamental assumptions might just be a symptom of the fact that IR theory is in fact an immature discipline that has not yet managed to become a normal science. This, however, would seem to cut against the disciplinary self-identity that IR has cultivated for itself of having become, (or almost succeeded in becoming) a true science. In this regard, the Lakatosian framework both allows for multiple research programmes to co-exist at the same time as it legitimizes this fragmented state of affairs as being a healthy one for a mature science.

There is another more analytical reason as to why the displacement from theories upwards to groups of theories as the key unit of analysis makes good sense. This has to do with the fact that common analytical elements are necessary as a pivot upon which comparisons, and

therefore any evaluations of progress, can be made. In any comparative analysis, common elements that link the objects of comparison are necessary for a stable basis for comparison to exist. In the case of philosophies of science that posit progress to be a relation between individual theories, there must be a basis for comparing different theories, which would seem to necessarily involve disaggregating theories to look for internal components which can be compared and through which a measure of “how much more excess content one theory contains than another” might be weighed. However, not only would this task seem *prima facie* to be exceptionally difficult in terms of measurement, (the ‘knowledge content’ of a theory seems hard to measure in quantitative terms), but the transposition upwards to groups of theories seems preferable to the one downwards to disaggregated theory components because comparisons of a more intuitive kind can be undertaken between the theories of which a research tradition is composed.

There remains the question of what exactly the analogue of a Lakatosian research programme is in the field of IR theory. Let us remind ourselves that in order for there to be a Lakatosian research programme, there must be a group of theories that are tied together by their commitment to a hard core of commonly shared assumptions. There have, of course, been numerous suggestions as to what the exemplifications of Lakatosian research programmes are in the field of IR theory. For instance, Keohane and Martin (2003) argue that institutional theory constitutes a research program,¹⁶ Moravcsik makes the claim on behalf of liberal IR theory,¹⁷ and Schweller asserts that neoclassical realism constitutes a progressive research programme in the Lakatosian mould.¹⁸ However, for all these claims, Bennett has stated frankly that “there is no non-arbitrary means of achieving a consensus on what constitutes the hard core and outer belt of any particular research program.” (Bennett, 2003: 462) This implies that the task of identifying

Lakatosian research programmes in the field of IR theory cannot but end up as a descriptive endeavor rather than an analytical one. Furthermore, as Schweller has noted:

“[T]here is little consensus on what constitutes an IR research program...does anyone reasonably believe that a fieldwide consensus can be reached on these issues? With the possible exception of Waltzian neorealism and John Mearsheimer’s offensive realism, none of the field’s various ‘isms’ represents a fully articulated theory of international relations. They are, instead, vaguely conceptualized ‘perspectives’ or political philosophies rooted in and guided by beliefs about ‘what matters’ and what requires explanation in international affairs.” (Schweller, 2003: 313)

If this is in fact the case, it is indicative that Keohane and Martin have so much difficulty in deciding whether neo-liberal institutional theory constitutes an autonomous research programme or is merely a “half-sibling of realism” (Keohane and Martin, 2003: 81). This is due to institutional theory sharing all of the key assumptions of Waltzian neo-realism except the implicit notion of the informational content of an international system, which they see as being a variable and not a constant. But how do we decide when a theory belongs in one research programme or another when most of the core assumptions are common to both? Lakatos gives us no clear answers here. Perhaps a more intuitive example of the Lakatosian research programme in IR theory, instead of the ‘isms’ such as neo-realism and neo-liberalism, are groups of theories sharing not merely a set of assumptions self-defined by the theory’s proponents, but groups of theories more broadly defined to share ontological assumptions in common. As Moravcsik has proposed, it might be reasonable to “speak of a rationalist research program in IR, with realist, institutionalist, and liberal ‘paradigms’ as leading elements.” (Moravcsik, 2003: 202) In either

case, the isolation of cases of the central unit of analysis proposed by a philosophy of science seems for us to be a non-trivial problem, in contrast to the attitude adopted by many of the authors in the *PIRT* volume, which for us amounts to the laissez-faire view that even if examples of research programmes cannot be found in IR, a) it does not matter because the theoretical groupings in IR are somehow self-evident, or b) a future consensus will eventually emerge as to what counts as a research programme in the field of IR theory. If analogues of the unit of analysis that a philosophy of science proposes cannot be found in a straightforward manner, this would seem at the very least to suggest a limitation in the applicability of that theory for the case at hand.

The problem of deciding what constitutes valid examples of research programmes in IR seems to correspond to a question that posed an important problem for Waltz, namely how the term ‘theory’ should be defined.¹⁹ However, the initial distinction that Waltz sets up in this regard between theories as the end product of an inductive process and theories as constructed pictures of autonomous domains that guide our observations – is not the most pertinent distinction for our analogy between Waltz and Lakatos. Instead, we must turn to the other possible view of what a theory is – the view that theories are a means of interpreting the world, and not of explaining it. This is the view that Waltz associates with the tradition of political theory, where theory “is concerned more with philosophic interpretation than with theoretical explanation.” (Waltz, 1979: 6) The problem of deciding between defining theories as explanatory devices on the one hand or as interpretive devices on the other, leads directly to the broader question of whether entities such as ‘theories’ and ‘research programmes’ should be defined from a purely external (or *etic*) point of view, (where the analyst decides what should or should not count as a case), or whether the self-description of a purported theory or research programme

by its author should also play a role in its definition, (i.e. an *emic* approach). This issue has not been adequately tackled in the discipline, and we must deal briefly with it here.

The debate about the respective merits of *etic* and *emic* approaches to social research has a long tradition in the social sciences. On the one hand, for the tradition of interpretive research the meanings of social actors are taken to be the primary form of data upon which researchers should base their findings, whereas for behaviorists and others who emphasize the explanation of action through causal models – the primary task of social research involves transcending the meanings of individuals to grasp large-scale regularities in social action.²⁰ In the case at hand, paradoxically, the social actors in question are the theorists of IR themselves – but this need not deter us from seeing the analogy between these two scenarios. The question here is whether we must choose one or the other of these poles, or whether we should decide that such a choice would unnecessarily conflate two approaches that each hold value for us in illuminating an important aspect of reality.

By way of example, in the case of research programmes, an *etic* perspective might allow us to identify groups of theories in IR, such as Moravcsik's 'rationalist research programme,' held together by logically defensible criteria such as a commitment to causal theories, derivation of testable hypotheses, parsimony in theory construction, and so on. But nevertheless any grouping of theories made by such means will necessarily conflict with some theorists' self-understandings of what school of thought their theory belongs to, (for instance most strands of neo-realism, much of neo-liberalism, and some theories in the constructivist school would seem to come under the rubric of a 'rationalist' research programme), and different interpreters would likely have different perspectives as to how widely inclusive such categories should be – opening up the specter of indeterminacy and confusion. On the other hand, *emic* approaches might focus

in on 'natural kinds' such as the 'isms' or even narrower approaches such as research on the democratic peace – but few of these natural kinds seem to comfortably conform with the requirements for being considered as research programmes, let alone theories. For instance, the democratic peace paradigm has often been called a statistical correlation in need of a theory, due to the sparseness of the hypotheses and predictions about the international system that can be drawn from its key contentions. How can the democratic peace perspective plausibly claim to be a research programme when even its status as a genuine theory is in doubt? In any case, the irreducibility of etic and emic approaches would seem to be a necessary fact that must be accepted, and we believe that any future attempts to specify Lakatosian research programmes in IR should attempt to combine both etic and emic indicators in a single multidimensional scale.

Let us move on now to ask how well Lakatos's notions of sophisticated falsification and progressive problemshift manage to capture the nature of theoretical development in IR. Lakatos notes that a theory embedded in a series of theories constituting a research programme is 'falsified' in the sophisticated sense "when it is superseded by a theory with higher corroborated content." (Lakatos, 1970: 118) How often, in fact, are theories in IR superseded by other theories that possess higher corroborated content? We are faced here with the problematic fact that theories in IR are rarely superseded to the extent in which theories in the natural sciences are superseded by their successors.

To take one example, what has been labeled as 'neoclassical realism' by Randall Schweller and others, has recently gained considerable interest in IR not by proposing a new theory intended to supersede what is considered the most sophisticated contemporary presentation of realism, Waltz's *Theory of International Politics*, but by in fact merely 'bringing back in' aspects of older realist theories dealing with the domestic sources of foreign policy as a

means of supplementing neo-realism's weaknesses and thus providing additional explanatory power.²¹ As Schweller has put it, neo-classical realism "while not abandoning Waltz's insights about international structure and its consequences...have added first and second image variables to explain foreign policy decision making..." (Schweller, 2003: 317) Although neo-classical realism is asserted to be a progressive research programme by its proponents, it would seem more accurately to be described as a degenerating one, based on its discovery not of novel facts but of facts already predicted by the theories of classical realism that neo-classical realism self-consciously adopts.

But what exactly does the fact that prior knowledge (contained in earlier theories) is frequently re-incorporated back into later theories tell us? For one thing, this observation seems to contradict the notion of cumulative knowledge growth across sequences of theories, because if it were true that the knowledge content of earlier theories is perfectly encompassed by later ones, there would be no need for such re-incorporation. On the same note, if it were the case that all later theories perfectly encompassed earlier ones, this would be equally problematic for it would come to seem as though no such thing as a genuinely 'new' theory were possible – all later theories becoming ad hoc adjustments of earlier ones. Perhaps this paradox of transitions in theory sequences is in part a necessary consequence of the incomplete nature of parsimonious theories which try to minimize theoretical redundancy at the same time as maximizing explanatory power. Whatever its source, it is clear that there is a fundamental tension between the idea of sequences of theories where later theories supersede earlier ones and the idea of cumulative knowledge growth where the knowledge content of earlier theories is absolutely retained in earlier ones.

This seems also to be related to a notable characteristic in social science more generally, which is that intellectual revolutions analogous to those in the natural sciences do not seem to occur very frequently. It is perhaps a corollary of the relative absence of revolutions in social science that theories in social science are rarely abandoned completely, and the phenomenon we have just seen of certain parts of an older theory being brought ‘back in’ is something we can witness over and over again in the social sciences. Why might this be the case?

One part of the reason, as we have already seen, is that social science theories in fact very rarely manage to encompass all of the knowledge content of their predecessors, contrary to Popperian and Lakatosian assumptions.²² There is often a loss in knowledge content when new theories are adopted, and this goes some part of the way to explaining why older theories in the social sciences are rarely definitely superseded. We might look, for instance, to the resurgence of interest in questions of collective preference formation and collective identity by those working on developing institutional theory and the rationalist paradigm more generally.²³ This has been a response to one of the main perceived weaknesses of the rationalist paradigm, namely the under-determined nature of the preferences of real-world actors when game theoretical models are applied to empirical scenarios. (Milner, 1998) As it happens, some thinkers of the older paradigms, (such as classical liberalism and psychological anthropology), had already dealt at some length with issues to do with preference formation and collective identity. And predictably enough it is to these approaches that rationalists first turned when attempting to fill-in the lacunae associated with their research tradition. In any case, it must be acknowledged that one of the main limitations of the rationalist approach – its under-theorization of questions of collective identity – had arisen from a loss of extant knowledge generated by earlier thinkers.

Another aspect of this absence of definitive obsolescence of social science theories is a seeming paradox, which is that some of the older theories of politics, society, and so on, persist even when the object of analysis in social science (i.e. society) changes at what can be presumed to be a faster rate than the fundamental entities posited by the hard sciences – which are assumed in their essence to remain the same across time. However, the paradoxical nature of this fact becomes more explicable when we see that this discrepancy is partially an artifact of a misleading comparison. Historical process and the constantly changing nature of social reality necessitates that contemporary accounts from past ages embedded in the theoretical, historical, and other narrative writings of that age are often the only means we have of unlocking the phenomenological (i.e. lived experiential) reality of a distant age. Because the task of understanding (and not merely explanation) constitutes an important part of the mission of the social sciences, the study of societies, individuals, and ideas distant from us not only geographically but also historically is an essential task without a clear parallel in the natural sciences. This seems to go some way to explaining why the study of theories and paradigms of past ages does not seem to be integral to the practice of natural science in the same way that it is for the social sciences.

Returning to our discussion of the seeming lack of paradigmatic revolutions in the social sciences, one tempting response might be to conclude that this characteristic is due to the social sciences having failed thus far to become real sciences. This perspective, which utilizes the Kuhnian distinction between normal science and pre-paradigm science, would take the lack of revolutions to be indicative of the absence of a disciplinary consensus in a pre-paradigmatic discipline, because when there exists no such disciplinary consensus then it is logically virtually impossible for revolutions – which invariably attempt to overthrow core conventions embodied

in a disciplinary consensus – to arise. One prescription here might be to attempt to model social science more closely around the image of the natural sciences, aiming for increased disciplinary consensus and the formation of hegemonic paradigms to facilitate progressive development. This, however, is not a route that I expect to be particularly productive. The main problem, of course, is that a disciplinary consensus that is legislated into existence through conscious intervention is not the same thing as a disciplinary consensus that emerges as an unintended consequence of internal developments in a field. The desired form of evolutionary development that is associated with the latter non-intentional process, it seems to us, will not arise through the artificial induced consolidation of a discipline.

In IR, the view which suggests that social science should model itself more closely on the natural sciences first gained wide currency in the field with the onset of the so-called ‘behavioral revolution’ which swept IR in the 1950s and 1960s, when the qualitative methods and textual analysis associated with the older traditions of diplomatic history and international law were gradually supplanted by an increasing emphasis on the study of observable behavior and empirically testable hypotheses. This influx of new methods sparked a backlash on the part of scholars who were more comfortable with the traditional approach, and the debate that ensued came to define one of the most significant cleavages in the field during the 1960s.²⁴ However, even in this early phase of the so-called ‘second debate’ when empirical methods were eagerly adopted by many in the field, it was not exactly clear what the core of this ‘behaviorism’ consisted of. As Knorr and Rosenau put it in an early review article, amongst the proponents of the scientific approach “consensus is lacking...for some, science means abstract theory, for others it connotes testable hypotheses, for still others it is synonymous with the collection of quantified data.” (Knorr and Rosenau, 1969: 13)

Of course, the difficulty of pinning down exactly what constitutes the essence of the scientific method as it has been practiced was hardly a newly discovered problem. Toulmin had argued that what is commonly held to be one of the cornerstones of the scientific method, *prediction* as the means of explanation, is in fact a misleading characterization insofar as prediction denotes a family of related but nevertheless distinct practices which cannot be conflated into a singular practice without distortion. Firstly, he alerts us to branches of science which have produced powerful theories without any predictive power, such as the evolutionary paradigm in biology originating from the insights of Darwin. This leads to the conclusion that prediction, if we must retain it as the cornerstone of our definition of science, should refer not only to making intelligible the future, but also “the ability to infer the occurrence of any event in question – whether it has already happened, is happening now, or is going to happen in the future.” (Toulmin, 1961: 27)

Secondly, Toulmin points out the paradox that some historical branches of science, such as Babylonian astronomy, achieved a significant degree of predictive power without achieving a requisite understanding of the phenomena being predicted. Furthermore Newton’s physical theories, one of the most powerful scientific theories of modern times, was stunningly successful in developing our understanding of physical phenomena such as celestial and tidal motions, but nevertheless made actual forecasts that were incorrect in many cases. Toulmin concludes that it is the pre-theoretical ideals of order underlying scientific paradigms that most properly characterize the scientific enterprise, not commonly held notions such as ‘science as explanation,’ and the like. He writes of prediction that “forecasting...is a craft or technology, an application of science rather than the kernel of science itself...a novel and successful theory may

lead to no increase in our forecasting skill; while alternatively, a successful forecasting-technique many remain for centuries without any scientific basis.” (Toulmin, 1961: 36)

As we can see, even the question of what constitutes the core of scientific method in the natural sciences is open to contention. Thus, instead of advocating a premature assimilation of social science into natural science, it would seem a wiser response to acknowledge the divergent forms of disciplinary development in the social and natural sciences as being indicative of substantive differences in the subject matter and methodology of these respective fields. Perhaps one fruitful comparison here is the differential nature of change undergone by written and unwritten constitutions. James Bryce, in depicting the differences between written constitutions and unwritten constitutions, has alluded to the fact that written constitutions – due to their codified nature and formalized amendment rules – are often less able to resist revolutionary pressures than are unwritten constitutions, better placed as they are to absorb similar pressures due to their decentralized nature. He writes that:

“A Rigid Constitution, if the legal means provided for altering it cannot be used for the want of the prescribed legal majority, resists the pressure. It may of course resist successfully, but if so, probably after a conflict which has shaken the State and excited hostility to it in the minds of a large part of the people. It may, however, if the assailing forces are very strong, be broken, and if so, broken past mending. A Flexible Constitution, however, being more easily and promptly alterable, and being usually a less firmly welded and cohesive structure, can bend without breaking, can be modified in such a way as to satisfy popular demands, can escape revolution by the practical submission of one of the contending forces in the particular dispute, that submission being recognized as a precedent which will be followed, even though it has not been embodied in any law or other formal document.” (Bryce, 1901: 147-148)

If this depiction is an acceptable one for capturing not only the distinction between written and unwritten constitutions, but more importantly the differences between disciplinary development in natural science and social science respectively, then it is perfectly explicable why revolutions are almost absent from the historical development of the social sciences. In other words, due to the more rigid and narrowly defined nature of theories in the natural sciences, most attempts at revision will necessitate throwing out the old theory altogether, rather than merely shifting the parameters of the old theory. However beneficial such a process might be for disciplinary progress, it would seem to entail a lot of intellectual waste as well.

Thus the overly negative evaluation of the decentralized and contentious disciplines in the social sciences might stand in need of correction. As a matter of fact, when compared to their more streamlined cousins in the natural sciences, such characteristics as diffuse structure and low degree of consensus might prove to be a relative boon that allows the disciplines of the social sciences not merely to avoid the all-or-nothing conflagrations that can occur in disciplines unified under a hegemonic viewpoint – but enable them also to draw on the full gamut of past theories for their analytical insights, without being restricted to merely the most current perspectives.²⁵

Therefore, sophisticated falsification might be a rather misleading way to characterize disciplinary development in a social science field such as IR. Not only are theories in research traditions seldom superseded definitively by their successors, but later theories – far from always possessing corroborated novel content as Lakatos's theory would require – often do not manage to encompass the collective knowledge-content of their predecessors and thus pillage insights from older theories to fill in gaps where knowledge-content has been lost, in the process creating

an eclectic synthesis. And there might be another more intrinsic reason why Lakatosian sophisticated falsificationism is not an appropriate way to view social phenomena. If a newly arrived physical theory predicts new particles to exist, it is not believed that these newly postulated particles have come into existence together with the emergence of the new theory. In contrast to the nature of the fundamental elements of the physical realm, (especially as conceptualized by the hard sciences – physics, chemistry, and so on), which are assumed to be unchanging and constant, the fundamental entities that must be explained in the social world in fact change with the passage of historical time and the changing forms of social organization. Thus, when the fundamental entities being observed change, a new theory will necessarily have to deal with new fundamental entities, and this would surely mean that the Lakatosian requirement that new theories must account for all the observations that previous theories have accounted for and more, is an impossible requirement to satisfy in such a circumstance. Although this can be assumed to hold regardless of context, the difficulty of constructing cumulative theories of the social world would be especially pertinent in times of momentous social change when the basic constituent entities (such as for instance the state, or the family – both of which have undergone significant historical change) and modes of organization of a society or system, are in flux and changing rapidly.

There is, however, one aspect of Lakatos's notion of sophisticated falsification which does ring true as a description of the empirical development of IR theory, and that is its rejection of naïve falsification and its assertion that theories are never falsified by mere observation. This is also the key sense in which Waltz believes political scientists can benefit from studying Lakatos's approach. As he has cogently put it:

“Lakatos’s assaults crush the crassly positivist ideas about how to evaluate theories that are accepted by most political scientists. He demolishes the notion that one can test theories by pitting them against facts.” (Waltz, 2003: xii)

In rejecting the naïve falsification of Popper, Lakatos made a two-pronged attack against both dogmatic verification and falsification as adequate theories of scientific method. As he put it, “scientific theories are not only equally unprovable...but they are also equally undisprovable,” (Lakatos, 1970: 103), and this is due to the fact that the relation between theory and fact is indeterminate as neither verification nor falsification of theory can be adduced merely from a simple comparison of facts and theories. Verification is disproved because more than one set of facts can accord with any particular theory, and falsification is disproved because more than one theory can explain any set of facts.²⁶ A similar argument holds in the case of probabilistic theories, and Lakatos notes that “no finite sample can ever disprove a universal probabilistic theory.” (Lakatos, 1970: 102) This state of affairs leads to the necessity for a more complex three-way relation between a set of facts, and two or more competing theories. Rationality can therefore no longer be the description of a simple relation between facts and theories, but must rather come to be seen as a more complex relationship between theories and other theories mediated by facts.

So how does Lakatos’s rejection of naïve falsification tally with the history of theoretical development in IR? Let us look at a couple of examples. Firstly, and as we have seen already, theories in IR (and in the social sciences more generally) are rarely definitively superseded, but are often integrated back into the contemporary body of theory in a cyclical fashion some time after their presumed obsolescence. This would already seem to suggest the limitations of

falsification as a model for theory-choice in IR. Secondly, if we look at a particular case – for instance the end of the cold war – as a potentially important set of facts with which predictions deriving from specific theories of IR can be compared, we find that such events actually do not falsify any theories in a simplistic manner. Rather the interaction between fact and theory in such situations is more complex. The end of the cold war presented the structural realism of Waltz with a dilemma, in so far as some predictions deriving from the Waltzian framework are confirmed and other predictions are disconfirmed by the facts of this case, a common outcome of the comparison of facts and theories. Nevertheless, the degree of disconfirmation that is evident and the survival of neo-realism as a major theoretical approach in the field despite this would seem to suggest that Lakatos’s rejection of falsification – and equally his advocacy of tenacity in the face of intellectual challengers together with his injunction to treat budding research programmes leniently – consists of lessons that have already been heeded by practitioners in the field, though primarily as a consequence of their pre-reflexive intuitions. Let us look briefly at this case here.

William Wohlforth in his important paper, “The Russian – Soviet Empire: A Test of Neorealism,” attempts to apply one implication of Waltzian neo-realism to the case of the evolution and decline of the Soviet Empire. He argues essentially that the vicissitudes of the rise and fall of the Russian and Soviet Empires closely reflects the power of deep anarchical structure (through the mechanisms of socialization and competition) to mould the units of an international system into becoming structurally similar. In one sense, the application of neo-realism to the disintegration of the Soviet Empire can be seen as a brilliant confirmatory application of a theory that is more often taken as being disconfirmed by the end of the cold war – in many ways a classic example of the Lakatosian positive heuristic at work. However, what at first appears to be

the successful application of neo-realism to this case nonetheless opens up numerous conceptual ambiguities which neo-realism struggles to resolve.

In discussing the mechanisms through which the units of a system become similar, Waltz writes that:

“The first way in which structures work their effects is through a process of socialization that limits and molds behavior. The second way is through competition. In social sectors that are loosely organized or segmented, socialization takes place within segments and competition takes place among them. Socialization encourages similarities of attributes and of behavior. So does competition. Competition generates an order, the units of which adjust their relations through their autonomous decisions and acts.” (Waltz, 1979: 76)

But, on a deeper structural level, what are the permissive causes that, for Waltz, lead to the mechanisms of socialization and competition to be operative? Two levels of structure can be seen to have relevance here. The ordering principle of the current international system, namely anarchy,²⁷ and through it the imperative to ensure for one’s own survival first and foremost, leads to the emergence of a self-help system. A threatening environment such as a self-help system in which states must compete to survive, means that states must mimic successful strategies in order to maximize their own capacities for survival. This mimicry takes place mainly along two dimensions: sovereignty as the principle of organization that states must reproduce in order to remain autonomous over their own domain; and the duplication of successful military technologies and strategies as the means through which states must ensure their continued ability to fend off attacks. Anarchy and the logic of survival that goes with it is for Waltz the most important structural principle that ensures that competition not co-operation,

homogenization not differentiation, and socialization and not toleration are the mechanisms that will remain operative in the international system.

The second dimension of structure, the character of the units, is the main level in which the ordering principle of anarchy and self-help is seen to have its consequences. The nature of the units in an anarchic realm such as the international system, is for Waltz characterized by functional homogeneity and structural similarity, and a mechanism of natural selection operates in the international environment through which successful adaptations are selected (and reproduced throughout the system over time) and unsuccessful adaptations or strategies fall by the wayside and become extinct. As Wohlforth notes, “the mechanism through which system pressures ‘shape and shove’ states in particular ways is not rational optimization of each decision problem but rather socialization, adaptation, and selection.”²⁸

Wohlforth, in pointing to the fact of the disintegration of the Soviet Union and its replacement by a handful of nation-states fully sovereign in the Westphalian mode, suggests that the socializing power of international structure (in creating like-units through a process of selection) was somehow operative in selecting out the Soviet Union as a unit dissimilar from the other units in the international system, (at least since the dissolution of the great Western European empires with the last wave of de-colonization starting from the mid-20th century), and thus under structural pressure to either adapt itself or ‘die.’ It is argued that this mechanism of negative selection would inevitably have played itself out over the long-term, and as an anomalous system unit the Soviet Union would have become subject to progressively greater degrees of structural pressure to assimilate so as to become similar to the rest of the system units. Although such structural pressures of a Waltzian anarchy would have been operative over a long time-period, thus bringing us back to the problem of timing and the need to account for the time-

lag separating structural transformation from unit transformation, Wohlforth points by way of explanation to the mechanism of path-dependence as a form of negative feedback which – whilst not ultimately overcoming structural pressures – temporarily raises the costs of moving to another equilibrium which might in fact be more efficient (and less costly). Such path dependence, (and the institutional inertia that results), could be one way of explaining the unduly long length of time during which the Soviet Union survived by forcibly maintaining its territorial integrity in the face of structural pressures pushing in the opposite direction.

But regardless of these countervailing tendencies, when Wohlforth writes that “securing an empire in a world of empires is one thing; doing so in a world of states is quite another,”²⁹ he gets to the crux of the matter, which is that as an anomalous unit in a system subject to homogenizing tendencies, the long-term survival of the Soviet Union was always in doubt. When Wohlforth writes that the USSR was “a state that radically concentrated political and economic power in Moscow; yet it was formally organized as a federation based on the national principle and it granted non-Russian nationalities far more cultural rights than they had ever had under the old regime”³⁰ he well captures some of the contradictions inherent in the nature of the Soviet Union as a system unit. It may even be that the USSR is best conceptualized as an international system in its own right, and Buzan and Little think that “viewing...the Soviet Union as [a] type of international system forces us to reassess what we mean by an international system and how an international system should be conceptualized.”³¹

However the USSR might ultimately be classified, it seems clear that the Soviet Union simultaneously displays facets which make it resemble an empire, a suzerain state-system, and a more conventional nation-state all at once. This very peculiar nature of the Soviet Union as a system unit should perhaps already have alerted us to the difficulty of analyzing the USSR in

classical neo-realist terms. It seems to us that the key question which all this raises for structural realism is that of how convincing it is to talk about the survival impulse in a self-help system of anarchy when the nature of the unit in question is a composite one. In cases such as that of the Soviet Union, the existence of an empire, regionally integrated entities, and several states and nations can all be discerned within the structure of a contiguous territory, (and assuming that each of these structures carries along with it a concomitant identity and ties of allegiance), then neo-realism does not give us sufficient guidance as to how to ascertain which principle of attachment will ultimately become wedded to the structural impulse for survival and be salient in a particular instance.

This indeterminacy of neo-realism is well illustrated with the case of Soviet behavior at the close of the cold war, when the Eastern European republics together with the Brezhnev doctrine were effectively jettisoned by the Soviet leadership in Moscow in the face of a rising tide of revolutionary ferment over the course of 1988 and 1989. Neo-realism seems on shaky ground here, and the theory seems to be consistent with a number of contradictory expectations. On one hand, if the leadership of the Soviet Union felt that the survival of the core of the USSR was at stake, (and here one would perhaps have to argue that a submerged 'Russian' ethno-national identity continued to exist throughout the Communist era), then the sacrifice of costly territorial appendages (such as the Eastern European republics) which not only did not fit within the conceptual image of the national core but in fact endangered the very existence of the core – could justifiably be undertaken and in fact be understood as rational behavior aimed to ensure the survival of the unit. On the other hand, if the territorial integrity of the Soviet Union as an imperial unit is taken as being the relevant unit of analysis, then the behavior of the Soviet leadership in peacefully allowing for the effective secession of the Eastern European republics

becomes inexplicable in neo-realist terms as the loss of territory would here entail a fundamental threat to unit survival – something not to be taken lightly according to the structural imperative of *self-help*.

Furthermore, the twin notions of socialization and competition that Waltz argues are the primary mechanisms through which the logic of anarchy expresses itself, runs into an intractable set of problems when its logic is extended to other limit cases. As one example, the notion of socialization seems to sit uncomfortably with the existence of periods of systemic change in the international system that John Gerald Ruggie and others have pointed to, such as the transition between the medieval order organized around the principle of overlapping and distributed sovereignties and the Westphalian international order where sovereignty came to take on a rigid territorial and either/or logic of dominion of a territory and subjects. How does the socializing power of structure operate in such transitional moments when the very nature of the dominant unit itself is in flux? It seems that a rigid logic of systemic socialization does not allow us to account for the rare moments of systemic transformation where new units emerge that then go on to become replicated throughout the system. The question of unit variability (in terms of differentiation in system units) and the rate at which socialization occurs is another issue that Waltzian neo-realism does not seem capable of handling on its own terms.

Yet, despite these significant limitations of Waltzian structural realism with respect to the facts that comprise the end of the cold war, it is indicative of the way in which research is conducted in IR that neo-realism was not generally taken to have been falsified by such facts. Of course, Lakatos would have been surprised by none of this. From a Lakatosian perspective, the reluctance of IR theorists to part with a damaged theory might have much to do with the fact that the alternate theories available (such as constructivism) which propose some plausible

mechanisms (i.e. the power of agency and systemic interaction to override the push and shove of structures) through which the end of the cold war could have come about, nevertheless currently fall short of established theories such as neo-realism in numerous other respects. These shortcomings come under a number of distinct categories: aesthetic limitations, (such as the relative lack of parsimony and weak ratio of theoretical proposition to explanatory power); logical weaknesses, (such as tardiness in tracing the causal chain all the way back beyond ideational factors); and methodological incongruities, (such as the relative absence of a falsifiable core, overall inability to encompass most of the explanatory power of extant theories, and so on). These failings, taken together, would for Waltz even endanger constructivism's claims to being a legitimate 'theory' per se. But in any case, the main point that we should take away from this discussion is that the evidentiary status of the end of the cold war as a test for IR theory suggests that theories in IR are never simply falsified by a set of facts.

The more complex three-way relation that Lakatos advocates between two or more competing theories mediated by a set of facts clearly requires us to transcend the Popperian epistemology which (despite its skepticism regarding verification) ultimately reduces to a correspondence theory of truth. In its place, we should erect a more nuanced epistemology centered not around an absolute relation between theory and reality anchored by truth, but a comparative relation between one theory and other theories, where the important criterion is not truth, but a quality such as 'relative explanatory power,' or the like. Both Lakatos and Waltz agree that this comparative relationship between theories is of the utmost importance, the final arbiter of the quality of a theory here being not whether it is true but whether it is 'worth taking seriously' as Waltz has often stated. However, as we have seen already, there are still many in the IR fraternity who accept the somewhat outdated Popperian epistemology of theories as being

things that are either true or false. This seems to be a sign that an adequate epistemology which consolidates and extends the fragmentary insights of Lakatos and Waltz is yet to emerge.

Conclusion

In concluding, it might be helpful to recap some important points from the discussion so far. I will do so by extracting what seems to be a coherent position from what Lakatos and Waltz have to say about the nature of theories. For both Lakatos and Waltz, no theory is provable, but neither are theories disprovable either. Therefore a theory is neither true nor false. This is due to the fact that a theory is not the reality itself, but merely a 'picture' or 'model' of a set of facts. This implies that there can be more than one picture of a set of facts. No picture is definitively true, because a picture necessarily distorts its subject matter to some extent, (no matter how accurate it is), insofar as it is a representation and not the thing-in-itself. On the same note, no picture need necessarily be false, insofar as every representation can be made to accord with a set of facts to a lesser or greater degree.³²

For Waltz, in the place of the truth criterion, there are only theories which are better or worse in terms of the explanatory power that they provide. Would Lakatos agree with Waltz on this point? Despite a difference of emphasis, we think that the answer is essentially that he does. We have already seen that Lakatos believes that a theory is only 'scientific' if "it has corroborated excess empirical content over its predecessor" (Lakatos, 1970: 116) Here 'scientific' for Lakatos is another way of saying 'useful,' and the final arbiter of whether a theory is useful or not is not whether it is true but rather the extent to which it contributes to the growth of knowledge in the discipline. Thus, we think that Lakatos would endorse Waltz's emphasis of the explanatory power of a theory as being the prime measure of its use value. However, the meaning of 'explanation,' for Waltz, does not necessarily equate to predictive power, but must be defined more broadly. We have already seen that Lakatos does not view the philosophy of science as being primarily a predictive enterprise, but rather as concerning itself more with the

retrospective task of evaluating whether theoretical shifts in the historical evolution of research programmes were rational or not, depending on whether the most progressive alternatives were chosen by practitioners at moments of paradigm change. Waltz also concurs with this de-emphasis of the future in favor of the past, and as he writes: “success in explaining, not in predicting, is the ultimate criterion of good theory...theories of evolution, after all, predict nothing in particular.” (Waltz, 1997: 916)

It certainly seems as though we can't live with progress and can't live without it. Although progress seems ephemeral as an analytic concept, it is nevertheless something that seems vital in giving us a sense of direction both in our intellectual endeavors and in our everyday lives. Perhaps it is the impulse to over-insure ourselves against the possibility of a lack of progress that causes a discipline to fold back in on itself and concentrate on progressively narrower problems, rather than attempting to develop multiple analytic domains simultaneously. In this regard, perhaps Betts's admonition to the intelligence community to become more tolerant of intelligence failures is also fruitfully directed towards the practitioners of IR. As he notes in comments that are directly relevant to our field: “although marginal reforms may reduce the probability of error, the unresolvable paradoxes and barriers to analytic...accuracy will make some incidence of failure inevitable.” (Betts, 1978: 87) The lesson here is that we should be careful what we wish for. If we let our field be driven too much by the normative search for theoretical and predictive progress at the expense of other key values, this might have the paradoxical effect of taking us further away from producing meaningful accounts of the political world that are grounded in the exigencies of finding practical solutions to concrete political problems.

¹ Watzlawick *et al.*, (1967), underline the importance of meta-communication for successful interaction, and conversely point to an association between the inability to meta-communicate and various pathologies, including amongst others, schizophrenia. They write that, “the ability to metacommunicate appropriately is not only the *condition sine qua non* of successful communication, but is intimately linked with the enormous problem of awareness of self and others.” (p. 53)

² This optimistic outlook of progressive development is held not only by practitioners in the subfield of IR, but also by many of those in the wider field of political science. Such implicit understandings are indicated by the adoption of Popperian assumptions of scientific progress by King, Keohane and Verba in a widely used textbook on the logic of research entitled *Designing Social Inquiry*, (King, Keohane, and Verba, 1994).

³ See John A. Vasquez, “The Realist Paradigm and Degenerative versus Progressive Research Programs: An Appraisal of Neotraditional Research on Waltz’s Balancing Proposition,” *American Political Science Review*, 91(4): 899-912; and the responses by Kenneth Waltz, Thomas Christensen and Jack Snyder, Colin Elman and Miriam Fendius Elman, Randall Schweller, and Stephen Walt.

⁴ Colin Elman and Miriam Fendius Elman, (eds.), *Progress in International Relations Theory: Appraising the Field*, MIT Press, 2003, (henceforth *PIRT*)

⁵ It is also these Popperian insights that underpin efforts in the social sciences (such as that of King, Keohane, and Verba, 1994) to draw up a unified logic of inference in the social sciences anchored around the core notion of causation as inferred from counterfactual conditions. Such arguments for a unified logic of social research are not new, of course, and they not only bear a family resemblance to Max Weber’s discussion of causal analysis through the construction of ideal types (Weber, 1949), but they are also analogous to an earlier attempt by Arend Lijphart when he argued for a unified logic of explanation underlying statistical (large-n) methods, comparative methods, and case studies, all of which he claims are anchored on the logic of the experimental method, (Lijphart, 1971).

⁶ There can be discerned to be circular reasoning at work here, although the circle in this case is not a vicious one.

⁷ It has been noted that revolutionary developments in thought are frequently set off either by investigators who are young, or by those with a background in other fields coming new to an area who are thus enabled to see old problems in a new light. (Kuhn, 1996; Jervis, 1976: 197)

⁸ Despite Popper having jettisoned the view that theories are confirmed to be true, he nevertheless took it that – by implication – theories left standing after the rigorous process of falsification would have the attribute of greater verisimilitude, (or closeness to the truth), because they are less *false* than their predecessors.

⁹ Toulmin, in this connection, points out that “perhaps every new generation of scientists having any original ideas or ‘slant’ of its own finds itself, at certain points and in certain respects, at cross-purposes with the immediately previous generation.” He goes on to say that it is possible to “question...whether any natural science having a serious theoretical component *ever* develops by a process of ‘accretion’ alone.” (Toulmin, 1970: 45)

¹⁰ Admittedly the ‘constancy’ here is by default, and comes about because the knowledge content of competing theories cannot be directly compared (due to incommensurability). This therefore rules out even the possibility of measuring either increase or decrease, and leaves constancy as the only viable alternative.

¹¹ Lakatos writes that an alternate theory, (*T'*), to be a viable contender to the original theory, (*T*), must have the following characteristics: “(1) *T'* has excess empirical content over *T*: that is, it predicts *novel* facts, that is, facts improbable in the light of, or even forbidden, by *T*; (2) *T'* explains the previous success of *T*, that is, all the unrefuted content of *T* is included (within the limits of observational error) in the content of *T'*; and (3) some of the excess content of *T'* is corroborated.” (Lakatos, 1970: 116)

¹² See John A. Vasquez, “The Realist Paradigm and Degenerative versus Progressive Research Programs: An Appraisal of Neotraditional Research on Waltz’s Balancing Proposition,” *American Political Science Review*, 91(4):

899-912; and the responses by Kenneth Waltz, Thomas Christensen and Jack Snyder, Colin Elman and Miriam Fendius Elman, Randall Schweller, and Stephen Walt.

¹³ Colin Elman and Miriam Fendius Elman, (eds.), *Progress in International Relations Theory: Appraising the Field*, MIT Press, 2003, (henceforth *PIRT*)

¹⁴ Waltz, in *Man, the State and War*, (1959), writes that “prescriptions directly derived from a single image are incomplete because they are based upon partial analyses. The partial quality of each image sets up a tension that drives one toward inclusion of the others.” (p.230)

¹⁵ For instance, regarding anarchy, neo-realism takes it that international anarchy is a basic fact of international life that cannot be changed and which leads inevitably to characteristic behavior such as ‘self-help’ and ‘balancing.’ Neo-realism, on the other hand takes it that anarchy is fundamental but can nevertheless be mitigated through information distributing mechanisms such as international institutions, and constructivism takes it that anarchy itself can come to take on different hues, (conflictual, cooperative, etc) depending on the type of interactional history that the units of the system have built up over time.

¹⁶ Keohane and Martin argue that the hard core of institutional theory consists of the following assumptions: i) states are the primary actors in world politics, ii) states behave as if they were rational, iii) states pursue their interests, iv) states operate in a world of anarchy, v) nevertheless despite iv, international institutions constructed by states are effective in improving the conditions of information transfer between actors, leading to an amelioration of the negative effects of anarchy. (Keohane and Martin, 2003) Furthermore, they assert that the first four assumptions are shared with neo-realism. In fact, this is not, strictly speaking, correct. The systems theory developed by Waltz does not require rationalist assumptions of state behavior, and as Waltz has written, neo-realism “requires no assumptions of rationality or of constancy of will on the part of all of the actors. The theory says simply that if some do relatively well, others will emulate them or fall by the wayside.” (Waltz, 1979: 118)

¹⁷ Moravcsik takes three assumptions to make up the hard core of such a liberal theory of IR: i) the fundamental actors in international politics are rational individuals and private groups, ii) the goals that states pursue via foreign policy represent the preferences of certain groups in domestic society, iii) the configuration of state preferences shape state behavior in the international system. (Moravcsik, 2003)

¹⁸ The hard core assumptions here for neoclassical realism being that “power is the fundamental feature of international politics; and that the essential nature of international politics is conflictual.” (Schweller, 2003: 325)

¹⁹ Waltz posed the question as that of whether a theory should be defined as a collection or set of laws pertaining to a particular phenomenon on the one hand, or whether a theory should be defined as statements that explain collections of laws, on the other. (Waltz, 1979) He argues that a theory should be defined according to the latter, (i.e. theories explain laws), and Waltz says that defining a theory as merely a collection of laws falls prey to the inductivist illusion, a form of flawed reasoning whereby it is believed “that truth is won and explanation achieved through the accumulation of more and more data and the examination of more and more cases.” (Waltz, 1979: 4) More data and more association of course do not automatically lead to explanations or theories – they lead merely to ever more facts which overwhelm without necessarily giving us the answer to the question of *why* the relation between a certain cause and effect obtains. And unlike laws, theories are neither true nor false. Instead, Waltz says that we should ask of a theory: “How great is their explanatory power?” (Waltz, 1979: 6)

²⁰ As Hollis and Smith have nicely summarized the debate, “in international affairs, and throughout the social world, there are two sorts of story to tell and a range of theories to go with each. One story is an outsider’s, told in the manner of a natural scientist seeking to explain the workings of nature and treating the human realm as part of nature. The other is an insider’s, told so as to make us understand what the events mean, in a sense distinct from any meaning found in unearthing the laws of nature...’explaining is the key term in one approach, ‘understanding’ in the other.” (Hollis and Smith, 1990: 1)

²¹ As Gideon Rose argues in a well-known paper, neoclassical realism “explicitly incorporates both external and internal variables, updating and systematizing certain insights drawn from classical realist thought. Its adherents argue that the scope and ambition of a country’s foreign policy is driven first and foremost by its place in the international system and specifically by its relative material power capabilities. This is why they are realist. They argue further, however, that the impact of such power capabilities on foreign policy is indirect and complex, because systemic pressures must be translated through intervening variables at the unit level. This is why they are neoclassical.” (Rose, 1998: 146)

²² In fact, Laudan persuasively argues that such knowledge loss is also a characteristic of disciplinary development in the natural sciences. In a word, “there are usually problem losses as well as problem gains associated with the replacement of any older theory by a newer one.” (Laudan, 1977: 148)

²³ As Milner (1998) has noted, there have been many ideas brought into rationalist research from traditions “that directly examine the sources of preferences and beliefs...recent work in the rationalist paradigm has tried to make use of these nonrationalist accounts of preferences and beliefs.” (p. 784)

²⁴ Hedley Bull, in his important 1966 article, encapsulated this debate as a contest between the classical approach and that of the scientific approach, and he forcefully argues that what strengths exist in the scientific approach can in fact be accommodated within the classical approach.

²⁵ As Schmidt as argued for the case of IR, “by continuing to either neglect or distort the history of the field, international relations scholars cut themselves off from a wealth of potentially useful insights about international politics and fail to recognize the degree to which the contemporary discourse of the field is deeply rooted in the past. In this respect, the history of political science is unlike the history of natural science. Many of the ideas that have been recovered from reconstructing the history of the field continue to have direct relevance for the present-day study of international politics.” (Schmidt, 1998: 229)

²⁶ Waltz decisively states the problem as follows: “We are faced with both an infinity of data and an infinity of possible explanations of the data. The problem is a double one. Facts do not determine theories; more than one theory may fit any set of facts. Theories do not explain facts conclusively; we can never be sure that a good theory will not be replaced by a better one.” (Waltz, 1979: 9)

²⁷ It is often assumed from Waltz’s treatment that anarchy and hierarchy are mutually exclusive, but a more nuanced view suggests that important aspects of hierarchy are present even within anarchical systems. As Jervis suggests, “to a significant extent, states may form an anarchical *society*, with many shared values, principles, norms, and practices...indeed, rather than conceiving of the states and the international system as separate levels and taking the existence of states for granted, theories should recognize that the latter are actually brought into being by the standing and legitimacy given them by other actors and the system they constitute.” (Jervis, 1997, 108) This line of argument is also pursued in greater depth in Bull, 1977, and Milner, 1991

²⁸ Wohlforth, 2001, 227

²⁹ Ibid, 220

³⁰ Ibid, 225

³¹ Buzan and Little, 2000, p.6

³² This view of theories as models of reality resembles certain aspects of Ludwig Wittgenstein’s early picture theory of language as a representational system, as set out in his *Tractatus Logico-Philosophicus*. There, Wittgenstein built a theory of language as a collection of propositions which can come together in various configurations to make up pictures of the world. And similarly to Waltz, he took it that these pictures are “a model of reality.” (Wittgenstein, 1922: 39) However, because all possible pictures are linked to reality through the logical substratum which is common to both of them, no picture is completely false – and “what is thinkable is also possible.” (Wittgenstein, 1922: 43)

Bibliography

- Almond, Gabriel, and Stephen Genco, (1977), "Clouds, Clocks, and the Study of Politics," *World Politics*, 29(4): 489-522
- Ball, Terence, (1976), "From Paradigms to Research Programs: Towards a Post-Kuhnian Political Science," *American Journal of Political Science*, 20(1): 151-177
- Ball, Terence, (1987), "Is There Progress in Political Science?," in Ball, Terence, (ed.), *Idioms of Inquiry: Critique and Renewal in Political Science*, Albany: State University of New York Press
- Bennett, Andrew, (2003), "A Lakatosian Reading of Lakatos: What Can We Salvage from the Hard Core?," in Elman, Colin, and Miriam Fendius Elman, (eds.), *Progress in International Relations Theory: Appraising the Field*, Cambridge MA: MIT Press
- Bernstein, Steven, Richard Ned Lebow, Janice Gross Stein, and Steven Weber, (2008), "God Gave Physics the Easy Problems: Adapting Social Science to an Unpredictable World," *European Journal of International Relations*, 6(1): 43-76
- Betts, Richard, (1978), "Analysis, War, and Decision: Why Intelligence Failures are Inevitable," *World Politics*, 31(1): 61-89
- Beyerchen, Alan, (1992), "Clausewitz, Nonlinearity, and the Unpredictability of War," *International Security*, 17(3): 59-90
- Blyth, Mark, (2003), "Structures Do Not Come with an Instruction Sheet: Interests, Ideas, and Progress in Political Science," *Perspectives on Politics*, 1(4): 695-706
- Blyth, Mark, (2006), "Great Punctuations: Predictions, Randomness, and the Evolution of Comparative Political Science," *American Political Science Review*, 100(4): 493-498
- Blyth, Mark, (2008), "ELEN Degenerates – or – Four (plus two) Reasons to Take Ideas Very Seriously Indeed," *mimeo*
- Brown, Chris, (2001), *Understanding International Relations*, Houndmills: Palgrave
- Bryce, James, (1901), 'Flexible and Rigid Constitutions,' in James Bryce, *Studies in History and Jurisprudence*, New York: Oxford University Press
- Bueno De Mesquita, Bruce, (1985), "Toward a Scientific Understanding of International Conflict: A Personal View," *International Studies Quarterly*, 29: 121-136
- Bull, Hedley, (1966), "International Theory: The Case for a Classical Approach," *World Politics*, 18(3): 361-377
- Bull, Hedley, (1995), *The Anarchical Society*, Second Edition, New York: Columbia University Press
- Bury, J. B., (1932), *The Idea of Progress: An Inquiry into its Origin and Growth*, New York: Macmillan
- Butterfield, Herbert, (1965), *The Whig Interpretation of History*, New York: W.W. Norton
- Buzan, Barry, and Richard Little, (2000), *International Systems in World History: Remaking the Study of International Relations*, Oxford: Oxford University Press
- Buzan, Barry, and Richard Little, (2001), "Why International Relations has Failed as an Intellectual Project and What to do About it," *Millennium: Journal of International Studies*, 30(1): 19-39
- Campbell, Donald, (1975), "Degrees of Freedom and the Case Study," *Comparative Political Studies*, 8(2): 178-193
- Carr, E. H., (1961), *What is History?*, Houndmills: Palgrave Macmillan
- Carr, E. H., (2001), *The Twenty Years' Crisis*, Houndmills: Palgrave Macmillan

- Chernoff, Fred, (2002), "Scientific Realism as a Meta-Theory of International Politics," *International Studies Quarterly*, 46: 189-207
- Christensen, Thomas, and Jack Snyder, (1997), "Progressive Research on Degenerative Alliances," *American Political Science Review*, 91(4): 919-922
- Cole, Stephen, (1994), "Why Sociology Doesn't Make Progress Like the Natural Sciences," *Sociological Forum*, 9(2): 133-154
- Cunningham, Andrew, (1988), "Getting the Game Right: Some Plain Words on the Identity and Invention of Science," *Studies in the History and Philosophy of Science*, 19(3): 365-389
- Dessler, David, (1989), "What's at Stake in the Agent-Structure Debate?," *International Organization*, 43(3): 441-473
- Dessler, David, (2003), "Explanation and Scientific Progress," in Elman, Colin, and Miriam Fendius Elman, (eds.), *Progress in International Relations Theory: Appraising the Field*, Cambridge MA: MIT Press
- Diesing, Paul, (1992), *How Does Social Science Work?*, Pittsburgh: University of Pittsburgh Press
- Doty, Roxanne Lynn, (1997), "Aporia: A Critical Exploration of the Agent-Structure Problematique in International Relations Theory," *European Journal of International Relations*, 3(3): 365-392
- Dryzek, John S., (1986), "The Progress of Political Science," *Journal of Politics*, 48: 301-320
- Dryzek, John S., and Stephen Leonard, (1988), "History and Discipline in Political Science," *American Political Science Review*, 82(4): 1245-1260
- Eckstein, Harry, (1975), "Case Study and Theory in Political Science," in Greenstein, Fred, and Nelson Polsby, (eds.), *Strategies of Inquiry*, Reading MA: Addison-Wesley
- Elman, Colin, and Miriam Fendius Elman, (1997), "Lakatos and Neorealism: A Reply to Vasquez," *American Political Science Review*, 91(4): 923-925
- Elman, Colin, and Miriam Fendius Elman, (2002), "How Not to Be Lakatos Intolerant: Appraising Progress in IR Research," *International Studies Quarterly*, 46: 231-262
- Elman, Colin, and Miriam Fendius Elman, (2003), "Lessons from Lakatos," in Elman, Colin, and Miriam Fendius Elman, (eds.), *Progress in International Relations Theory: Appraising the Field*, Cambridge MA: MIT Press
- Farr, James, John Gunnell, Raymond Seidelman, John Dryzek, and Stephen Leonard, (1990), "Can Political Science Be Neutral?," *American Political Science Review*, 84(2): 587-607
- Gerring, John, (2004), "What is a Case Study and What Is It Good for?," *American Political Science Review*, 98(2): 341-354
- Gross, Leo, (1948), "The Peace of Westphalia," *American Journal of International Law*, 42(1): 20-41
- Gourevitch, Peter, (1978), "The Second Image Reversed: The International Sources of Domestic Politics," *International Organization*, 32(4): 881-912
- Hacking, Ian, (1981), "Lakatos's Philosophy of Science," in Hacking, Ian (ed.), *Scientific Revolutions*, Oxford: Oxford University Press
- Hall, Rupert, (1983), "On Whiggism," *History of Science*, 21(March): 45-59
- Hempel, Carl, (1942), "The Function of General Laws in History," *The Journal of Philosophy*, 38: 35-48
- Hollis, Martin, and Steve Smith, (1990), *Explaining and Understanding International Relations*, Oxford: Clarendon Press

- Hollis, Martin, and Steve Smith, (1991), "Beware of Gurus: Structure and Action in International Relations," *Review of International Studies*, 17: 393-410
- Hollis, Martin, and Steve Smith, (1992), "Structure and Action: Further Comment," *Review of International Studies*, 18: 187-188
- Holsti, K. J., (1989), "Mirror, Mirror on the Wall, Which Are the Fairest Theories of All?," *International Studies Quarterly*, 33(3): 255-261
- Jackson, Patrick Thaddeus, and Daniel Nexon, (1999), "Relations Before States: Substance, Process and the Study of World Politics," *European Journal of International Relations*, 5(3): 291-332
- Jervis, Robert, (1976), *Perception and Misperception in International Politics*, Princeton NJ: Princeton University Press
- Jervis, Robert, (1985) "Pluralistic Rigor: A Comment on Bueno de Mesquita," *International Studies Quarterly*, 29: 145-149
- Jervis, Robert, (1994), "Hans Morgenthau, Realism, and the Scientific Study of International Politics," *Social Research*, 61(4): 853-876
- Jervis, Robert, (1997), *System Effects: Complexity in Political and Social Life*, Princeton NJ: Princeton University Press
- Jervis, Robert, (2003), "Realism, Neoliberalism, and Cooperation: Understanding the Debate," in Elman, Colin, and Miriam Fendius Elman, (eds.), *Progress in International Relations Theory: Appraising the Field*, Cambridge MA: MIT Press
- Keohane, Robert, and Lisa Martin, (2003), "Institutional Theory as a Research Program," in Elman, Colin, and Miriam Fendius Elman, (eds.), *Progress in International Relations Theory: Appraising the Field*, Cambridge MA: MIT Press
- King, Gary, Robert Keohane, and Sidney Verba, (1994), *Designing Social Inquiry*, Princeton: Princeton University Press
- Knorr, Klaus, and Rosenau, James, (1969), "Tradition and Science in the Study of International Politics," in Knorr, Klaus and James Rosenau, (eds.), *Contending Approaches to International Politics*, Princeton: Princeton University Press
- Krasner, Stephen, (1985), "Toward Understanding in International Relations," *International Studies Quarterly*, 29: 137-144
- Kratochwil, Fredrich, (2000), "Constructing a New Orthodoxy? Wendt's 'Social Theory of International Politics' and the Constructivist Challenge," *Millennium: Journal of International Studies*, 29(1): 73-101
- Kratochwil, Fredrich, (2006), "History, Action and Identity: Revisiting the 'Second' Great Debate and Assessing its Importance for Social Theory," *European Journal of International Relations*, 12(1): 5-29
- Kratochwil, Fredrich, (2007a), "Of False Promises and Good Bets: A Plea for a Pragmatic Approach to Theory Building," *Journal of International Relations and Development*, 10: 1-15
- Kratochwil, Fredrich, (2007b), "Evidence, Inference, and Truth as Problems of Theory Building in the Social Sciences," in Lebow, Richard Ned, and Mark Irving Lichbach, (eds.), *Theory and Evidence in Comparative Politics and International Relations*, New York: Palgrave Macmillan
- Kratochwil, Fredrich, (2007c), "Looking Back from Somewhere: Reflections on What Remains 'Critical' in Critical Theory," *Review of International Studies*, 33: 25-45

- Kuhn, Thomas, (1996), *The Structure of Scientific Revolutions*, Chicago: Chicago University Press
- Kuhn, Thomas, (1970), "Notes on Lakatos," *Boston Studies in the Philosophy of Science*, VIII: 137-146
- Laitin, David, (1995), "Disciplining Political Science," *The American Political Science Review*, 89(2): 454-456
- Lakatos, Imre, (1970), "Falsification and the Methodology of Scientific Research Programmes," in Lakatos, Imre, and Alan Musgrave, (eds.), *Criticism and the Growth of Knowledge*, Cambridge: Cambridge University Press
- Lakatos, Imre, (1981), "History of Science and its Rational Reconstructions," in Hacking, Ian (ed.), *Scientific Revolutions*, Oxford: Oxford University Press
- Lapid, Yosef, (1989), "The Third Debate: On the Prospects of International Theory in a Post-Positivist Era," *International Studies Quarterly*, 33: 235-254
- Laudan, Larry, (1977), *Progress and Its Problems: Toward a Theory of Scientific Growth*, Berkeley: University of California Press
- Laudan, Larry, (1981a), "A Problem-Solving Approach to Scientific Progress," in Hacking, Ian (ed.), *Scientific Revolutions*, Oxford: Oxford University Press
- Laudan, Larry, (1981b), "A Confutation of Convergent Realism," *Philosophy of Science*, 48(1): 19-49
- Leifer, Eric, (1992), "Denying the Data: Learning from the Accomplished Sciences," *Sociological Forum*, 7(2): 283-299
- Lieberson, Stanley, (1992), "Einstein, Renoir, and Greeley: Some Thoughts about Evidence in Sociology," *American Sociological Review*, 57(1): 1-15
- Lieberson, Stanley, and Freda Lynn, (2002), "Barking Up the Wrong Branch: Scientific Alternatives to the Current Model of Sociological Science," *Annual Review of Sociology*, 28: 1-19
- Lijphart, Arend, (1971), "Comparative Politics and the Comparative Method," *The American Political Science Review*, 65(3): 682-693
- Masters, Roger, (1964), "World Politics as a Primitive Political System," *World Politics*, 16(4): 595-619
- McKeown, Timothy, (1999), "Case Studies and the Statistical Worldview," *International Organization*, 53(1): 161-190
- Milner, Helen, (1991), "The Assumption of Anarchy in International Relations Theory: A critique," *Review of International Studies*, 17: 67-85
- Milner, Helen, (1998), "Rationalizing Politics: The Emerging Synthesis of International, American, and Comparative Politics," *International Organization*, 52(4): 759-786
- Moravcsik, Andrew, (2003), "Liberal International Relations Theory: A Scientific Assessment," in Elman, Colin, and Miriam Fendius Elman, (eds.), *Progress in International Relations Theory: Appraising the Field*, Cambridge MA: MIT Press
- Morgenthau, Hans, (1946), *Scientific Man vs. Power Politics*, Chicago: University of Chicago Press
- Onuf, Nicholas, (1995), "Levels," *European Journal of International Relations*, 1(1): 35-58
- Patomaki, Heikki, (1996), "How to Tell Better Stories about World Politics," *European Journal of International Relations*, 2(1): 105-133
- Platt, John, (1964), "Strong Inference," *Science, New Series*, 146(3642): 347-353

- Polsby, Nelson, (1998), "Social Science and Scientific Change: A Note on Thomas S. Kuhn's Contribution," *Annual Review of Political Science*, 1: 199-210
- Popper, Karl, (1963), "Truth, Rationality, and the Growth of Scientific Knowledge," in Karl Popper, *Conjectures and Refutations*, London: Routledge
- Popper, Karl, (1981), "The Rationality of Scientific Revolutions," in Hacking, Ian (ed.), *Scientific Revolutions*, Oxford: Oxford University Press
- Rose, Gideon, (1998), "Neoclassical Realism and Theories of Foreign Policy," *World Politics*, 51(1): 144-172
- Ruggie, John Gerald, (1986), "Continuity and Transformation in the World Polity," in Keohane, Robert (ed.), *Neorealism and its Critics*, New York: Columbia University Press
- Sartori, Giovanni, (1970), "Concept Misformation in Comparative Politics," *American Political Science Review*, 64(4): 1033-1053
- Schmidt, Brian, (1998), *The Political Discourse of Anarchy: A Disciplinary History of International Relations*, Albany: State University of New York Press
- Schweller, Randall, (1997), "New Realist Research on Alliances: Refining, Not Refuting, Waltz's Balancing Proposition," *American Political Science Review*, 91(4): 927-929
- Schweller, Randall, (2003), "The Progressiveness of Neoclassical Realism," in Elman, Colin, and Miriam Fendius Elman, (eds.), *Progress in International Relations Theory: Appraising the Field*, Cambridge MA: MIT Press
- Sewell, William, (1996), "Three Temporalities: Towards an Eventful Sociology," in T. J. McDonald (ed.), *The Historic Turn in the Human Sciences*, Ann Arbor: University of Michigan Press
- Simowitz, Roslyn, (2003), "Measuring Intra-programmatic Progress," in Elman, Colin, and Miriam Fendius Elman, (eds.), *Progress in International Relations Theory: Appraising the Field*, Cambridge MA: MIT Press
- Singer, J. David, (1961), "The Level-of-Analysis Problem in International Relations," *World Politics*, 14(1): 77-92
- Skinner, Quentin, (1969), "Meaning and Understanding in the History of Ideas," *History and Theory*, 8(1): 3-53
- Snyder, Jack, (2003), "'Is' and 'Ought': Evaluating Empirical Aspects of Normative Research," in Elman, Colin, and Miriam Fendius Elman, (eds.), *Progress in International Relations Theory: Appraising the Field*, Cambridge MA: MIT Press
- Toulmin, Stephen, (1961), *Foresight and Understanding: An enquiry into the aims of science*, New York: Harper and Row
- Toulmin, Stephen, (1967), "The Evolutionary Development of Natural Science," *American Scientist*, 55(4): 456-471
- Toulmin, Stephen, (1970), "Does the Distinction between Normal and Revolutionary Science Hold Water?," in Lakatos, Imre, and Alan Musgrave, (eds.), *Criticism and the Growth of Knowledge*, Cambridge: Cambridge University Press
- Tully, James, "The Pen is a Mighty Sword: Quentin Skinner's analysis of politics," in James Tully, (ed.), *Meaning and Context*, Princeton NJ: Princeton University Press
- Vasquez, John, (1997), "The Realist Paradigm and Degenerative versus Progressive Research Programs: An Appraisal of Neotraditional Research on Waltz's Balancing Proposition," *American Political Science Review*, 91(4): 899-912

- Vasquez, John, (2003), "Kuhn vs. Lakatos?: The Case for Multiple Frames in Appraising IR Theory," in Elman, Colin, and Miriam Fendius Elman, (eds.), *Progress in International Relations Theory: Appraising the Field*, Cambridge MA: MIT Press
- Waever, Ole, (1998), "The Sociology of a Not So International Discipline: American and European Developments in International Relations," *International Organization*, 51(4): 687-727
- Walt, Stephen, (1997), "The Progressive Power of Realism," *American Political Science Review*, 91(4): 931-935
- Waltz, Kenneth, (1959), *Man, the State and War*, New York: Columbia University Press
- Waltz, Kenneth, (1979), *Theory of International Politics*, Boston: McGraw Hill
- Waltz, Kenneth, (1990), "Realist Thought and Neorealist Theory," *Journal of International Affairs*, 44: 21-37
- Waltz, Kenneth, (1997), "Evaluating Theories," *American Political Science Review*, 91(4): 913-917
- Waltz, Kenneth, (2003), "Assaying Theories: Reflections on Imre Lakatos," in Elman Colin, and Miriam Fendius Elman, (eds.), *Progress in International Relations Theory: Appraising the Field*, Cambridge MA: MIT Press
- Watzlawick, Paul, Janet Beavin Bavelas, and Don Jackson, (1967), *Pragmatics of Human Communication: A Study of Interactional Patterns, Pathologies, and Paradoxes*, New York: W. W. Norton
- Weber, Max, (1949), "'Objectivity' in Social Science and Social Policy," in Shils, Edward and Henry Finch, (eds.), *The Methodology of the Social Sciences*, New York: The Free Press
- Wendt, Alexander, (1987), "The Agent-Structure Problem in International Relations Theory," *International Organization*, 41(3): 335-370
- Wendt, Alexander, (1991), "Bridging the Theory/Meta-Theory Gap in International Relations," *Review of International Studies*, 17: 383-392
- Wendt, Alexander, (1992), "Levels of Analysis vs. Agents and Structures: Part III," *Review of International Studies*, 18: 181-185
- Wendt, Alexander, (1992), "Anarchy is what states makes of it: the social construction of power politics," *International Organization*, 42(2): 391-425
- Wilson, Adrian, and T. G. Ashplant, (1988a), "Whig History and Present-Centred History," *The Historical Journal*, 31(1): 1-16
- Wilson, Adrian, and T.G. Ashplant, (1988b), "Present-Centred History and the Problem of Historical Knowledge," 31(2): 253-274
- Wittgenstein, Ludwig, (1922), *Tractatus Logico-Philosophicus*, London: Routledge & Kegan Paul
- Wohlforth, William, (2001), "The Russian-Soviet Empire: A Test of Neorealism," in Cox, Michael, Tim Dunne, and Ken Booth (eds.), *Empires, Systems, and States: Great Transformations in International Politics*, Cambridge: Cambridge University Press