

PATERSON REVIEW

VOLUME 8, 2007

The Norman Paterson School of
International Affairs

Carleton University
1401 Dunton Tower
1125 Colonel By Drive,
Ottawa, Ontario, Canada
K1S 5B6

www.diplomatonline.com/patersonreview/home.html

The *Paterson Review*, formally *E-merge*, is a scholarly journal exclusively showcasing the work of graduate students in the field of international affairs. Managed by students from the Norman Paterson School of International Affairs, the *Paterson Review* is dedicated to publishing articles on a wide range of emerging issues in the theory and practice of international affairs. Copy requests and submissions may be sent electronically to patersonreview@gmail.com or by mail to Paterson Review c/o Norman Paterson School of International Affairs 1401 Dunton Tower, Carleton University 1125 Colonel By Drive, Ottawa, Ontario, K1S 5B6 Canada.

Copyright © 2007, Paterson Review

1. International Relations – Periodicals. 2. World Politics – Periodicals. 3. Policy Sciences – Periodicals.

CONTENTS

vii *Letter from the Editors*

- 1 Quick Impact Projects: Towards a ‘Whole of Government’ Approach**
*Jon Baker, Norman Paterson School of International Affairs,
Carleton University*
- 23 The Changing Nature of “International Security”: The Need for an Integrated Definition**
*C. Elisabeth St. Jean, Norman Paterson School of International Affairs,
Carleton University*
- 35 Turning the Health Professional Carousel: Is Canada Undermining Human Rights in Developing Countries?**
*Karinne Lantz, Norman Paterson School of International Affairs,
Carleton University*
- 51 Human Rights, Cultural Tradition and International Development: Culturally Appropriate International Responses to Female Genital Cutting**
*Kristen Thomasen, Norman Paterson School of International Affairs,
Carleton University*
- 67 Maintenance of Sanctions Regimes: Lessons from Iraq**
*Nathan Lysons, Norman Paterson School of International Affairs,
Carleton University*
- 79 Russia’s Foreign Policy Ace**
*Jeremy D. Gordon, Maxwell School of Citizenship and Public Affairs,
Syracuse University*
- 93 Protection for Whom? Stabilization and Coercive Rule in Haiti**
*Jennifer Peirce, Norman Paterson School of International Affairs,
Carleton University*

Critical Commentary

- 109 Effects of Financial Sector Foreign Direct Investment on Access to Credit for Small and Medium-sized Enterprises in Emerging Market Economies: A Critical Commentary**
*Manuela W. Armenta, Norman Paterson School of International Affairs,
Carleton University*

Book Review

- 125 Colonial Terrorist, American Hero: Two Histories of the Original American Insurgency**
*Daniel Doktori, School of International and Public Affairs,
Columbia University*

EDITORIAL STAFF

Co-Editors-in-Chief

Patrick Cuenco Kristen Thomasen

Treasurer and Marketing Manager

Marion Sandilands

Editors

Laurie Gorman
Lindsey Higgs
Nathan Lysons
Ashlyn Milligan
Marion Sandilands
Ryan Stein

Publishing Editors

Ashlyn Milligan
Ryan Stein

Graphic Designer

Predrag Tomic

ACKNOWLEDGEMENTS

Expert Reviewers

Dr. Chantal Blouin, *Researcher, The North-South Institute*

Dr. Louise de la Gorgendière, *Associate Professor of Social Anthropology, Dept. of Sociology and Anthropology, Carleton University*

John Graham, *Chair of FOCAL, the Canadian Foundation for the Americas*

Dr. Tessa Hebb, *Director of Capital Strategies, Carleton Centre for Community Innovation, Carleton University*

Dr. Jez Littlewood, *Director of the Canadian Centre of Intelligence and Security Studies, Carleton University*

Dr. David Malone, *Adjunct Research Professor, NPSIA and High Commissioner of Canada to India*

Christopher Penny, L.L.M., *Assistant Professor of International Law, NPSIA, Carleton University*

Dr. Yiagadeesen Samy, *Assistant Professor, NPSIA, Carleton University*

Dr. Elinor Sloane, *Associate Professor and Associate Supervisor of Graduate Studies, Dept. of Political Science, Carleton University*

Additional Reviewers

Randall Blair, M.A., *The Woodrow Wilson School of Public and International Affairs, Princeton University*

Martin Fischer, *PhD. Candidate, NPSIA, Carleton University*

Darryl Whitehead, *PhD. Candidate, NPSIA, Carleton University*

Friends

Vivian Cummins

Cindy Halden

Donna Jacobs

Patricia Lacroix

Leslee McCabe

Brenda Sutherland

Charlene Youssef

Jijun (Jessie) Zhao

Blind Review Contributing Editors

Ashley Barnes
Stephen Buchta
Patrick Cuenco
Alia Dedhar
Laurie Gorman
Lindsey Higgs
Christopher Kang
Joshua Kilberg
Karinne Lantz
Nathan Lysons
Ashlyn Milligan
Aruna Rajulu
Marion Sandilands
Ryan Stein
Kristen Thomasen

Letter from the Editors

We are pleased to present the 2007 edition of the *Paterson Review*. Now entering its eighth volume, the *Paterson Review* annually publishes the work of graduate students in the field of international affairs.

The eighth volume of the *Paterson Review* brought together a large group of graduate students, primarily from NPSIA, who have been working hard to make this year's journal a success. For the first time since its inception, the *Paterson Review* has been published in a hard copy format, which will be circulated to international affairs schools around the world. The *Paterson Review* is also available online through *Diplomat & International Canada* magazine at www.diplomatonline.com.

This volume was highly competitive, with 37 submissions in total. A three-round blind review was held to determine which of these papers, submitted by students from across North America, would be selected for publication. In total, seven articles, one critical commentary and one book review were selected. The following is a list of the schools from which papers were received:

Graduate School of International Relations and Pacific Studies – University of California, San Diego

The Norman Paterson School of International Affairs – Carleton University

School of International and Public Affairs (SIPA) - Columbia University

Centre for Contemporary Arab Studies – Georgetown University

Paul H. Nitze School of Advanced International Studies – Johns Hopkins University

School of Public Policy – University of Maryland

Woodrow Wilson School of Public and International Affairs – Princeton University

Maxwell School of Citizenship and Public Affairs – University of Syracuse

Jackson School of International Affairs – University of Washington

The Whitney and Betty MacMillan Center – Yale University

Following the blind review, the selected articles went through a rigorous editing process. This process included an evaluation of each submission by one or more experts in the field of the article's subject matter, along with simultaneous editing and formatting by the *Paterson Review* editorial staff, in coordination with the authors.

Our nine contributors investigate a broad range of timely international affairs topics, sure to incite discussion and hopefully contribute to the field of policy development. Four articles and our book review address a broad range of security and defence issues. Jon Baker assesses the efficacy of Quick Impact Projects, implemented by the Canadian Forces in Afghanistan, and makes a case for the development of valid and reliable measures of effectiveness for these

projects. Jennifer Peirce examines the police reform initiatives which have taken place in Haiti since 1994, while C. Elisabeth St. Jean explores traditional and new definitions of 'international security,' and looks for an integrated definition which incorporates the benefits and rejects the drawbacks of both. Nathan Lysons provides a comprehensive assessment of lessons-learned from the United Nations Security Council's experience with sanctions on Iraq. Finally, Daniel Doktori reviews the books *Washington's Crossing* by David Hackett-Fischer and *The Glorious Cause* by Jeff Shaara, and draws a connection between America's first experience with insurgent-style warfare during the American Revolution and the current U.S. military experience in Iraq.

Looking at foreign policy issues in Eastern Europe, Jeremy D. Gordon's article investigates the ties between the Kremlin and Russia's largest oil and gas company, Gazprom, and assesses Russia's use of Gazprom as a foreign policy tool in numerous Former Soviet Republics. The final two articles and our critical commentary examine diverse topics in development. Karinne Lantz investigates the moral and legal obligations of developed countries, looking specifically at the United Kingdom and Canada, with respect to the recruitment of health professionals from developing countries, particularly those in Sub-Saharan Africa. Kristen Thomasen tackles the question of how to address, in an appropriate and effective way, an internationally defined violation of human rights that is also a cultural tradition, assessing development policies geared towards decreasing rates of female genital cutting. Finally, Manuela W. Armenta provides a critical assessment of the current literature relating to the effects of financial sector foreign direct investment on the ability of small- and medium-sized enterprises to access credit in developing countries.

The *Paterson Review* would like to express its sincere appreciation to everyone who was involved in the creation of the eighth volume of the journal. We would like to extend a special thank you to all editors and expert reviewers who dedicated valuable time to our publication; to the authors for their hard work and dedication; to the staff at the Norman Paterson School of International Affairs for their ongoing support; to *Diplomat & International Canada* magazine for their continued partnership; and to our sponsors, the Centre for Security and Defence Studies, the Centre for Trade Policy and Law and NPSIA, for making this all possible.

Patrick Cuenco & Kristen Thomasen

QUICK IMPACT PROJECTS: TOWARDS A ‘WHOLE OF GOVERNMENT’ APPROACH

Jon Baker

Norman Paterson School of International Affairs, Carleton University

Using the case study approach, qualitative evidence from the Provincial Reconstruction Team (PRT) in Kandahar, Afghanistan will be evaluated in an attempt to assess the impact that Quick Impact Projects have had. The research objectives are: (1) to determine the level of involvement civilian agencies *should* have in the QIP decision-making process; and (2) to determine whether the level of involvement that civilian agencies currently have meets this standard. This article will argue that the military and civilian agencies of the Canadian government have made progress in coordinating their efforts, but further improvement of the PRT structure will be severely hampered until valid and reliable measures of effectiveness for QIPs are developed and implemented.

INTRODUCTION

Quick Impact Projects (QIPs) are increasingly used by military actors during Peace Support Operations (PSOs)¹ around the globe. These are usually short-term, small-scale ‘hearts and minds’ initiatives that are designed to have an immediate impact contributing to post-conflict stabilization or recovery. QIPs can benefit military personnel by increasing force protection, intelligence gathering, acceptance of foreign troops, and, conversely, opposition to enemy forces. Some observers, however, argue that QIPs are neither effective nor efficient from a development perspective because the military is simply unsuited and ill-equipped to provide such aid; moreover, they argue, involvement by the military compromises the neutrality and political independence of those NGOs who are most appropriate to carry out such projects. The result, critics say, is an approach that is at best ‘stove-piped’ rather than genuinely ‘whole of government,’ and at worst operating at cross-purposes. The research question, therefore, is rooted in whether the use of QIPs by the military to win ‘hearts and minds’ helps or hinders the overall ‘whole of government’ (WOG) approach.

Using the case study approach, qualitative evidence from the Provincial Reconstruction Team² in Kandahar Province in Afghanistan will be evaluated in an attempt to assess the impact that QIPs have had. The research objectives are: (1) to determine the level of involvement civilian agencies *should* have in the QIP decision-making process; and (2) to determine whether the level of involvement that civilian agencies currently have meets this standard. All this is based on the assumption outlined in Canada’s International Policy Statement that “an integrated ‘3D’ approach, combining diplomacy, defence and development, is the best strategy for supporting states that suffer from a broad range of interconnected problems” (Government of Canada, IPS, 2005).³ This article will argue that the military and civilian agencies of the Canadian government have

made progress in coordinating their efforts, but further improvement of the PRT structure will be severely hampered until valid and reliable measures of effectiveness for QIPs are developed and implemented.

DEFINITIONS AND CONCEPTS

The rationale behind a formalized WOG⁴ approach is two-fold. The first motive is to avoid duplication of efforts, interfering with the plans of other departments, and the consequent waste of energy and resources. The resulting information flows between actors increases situational awareness and enhances the capacity for ‘strategic’ planning and intelligent decision-making, including more awareness of second- and third-order effects (Gizewski 2007). This is vital given the number of actors who must coordinate their activities in contemporary PSOs—military, diplomatic, development, police, and so on (Savage 2006, 19). The second motivation for a WOG approach stems from a recognition that the goals of military and civilian organizations are co-dependent: “without security, development cannot happen, and without development, lasting security cannot be sustained” (Peabody 2005, 4). Harmonizing these efforts requires compromises and understandings between military and non-military actors.

If the number of actors involved in contemporary PSOs has increased, so too has the number of belligerents. According to many scholars, developments since the Second World War indicate that warfare is increasingly returning to a decentralized form in which states have lost their monopoly on combat forces.⁵ Those of most concern are ‘spoilers’—non-state actors who profit from instability or lawlessness and therefore seek to prevent or ‘spoil’ the establishment of peace.⁶ Moreover, the strategy that spoilers have adopted is one of *asymmetric* warfare:

In the past, the asymmetric threat was normally just a component of the larger conventional conflict. Today however, there is only one superpower in the world and this ‘asymmetric’ gap between the U.S. military forces and those of her adversaries continues to grow. As a result, the only method of attack against this superpower and western coalitions will be through asymmetric means (Marcella 2006, 61).

Insurgents⁷ constitute a type of spoiler group that frequently makes use of asymmetric warfare.

In virtually any counter-insurgency (COIN) operation, the support of the local population constitutes the ‘centre of gravity’ for both insurgent and counter-insurgent forces (Lewis 2006, 19-20).⁸ Not only can they provide the former with money, shelter, and food, but more importantly they are the primary intelligence source for both the insurgents and counter-insurgents. According to John Lynn, local populations (based on perceptions of legitimacy, security, or a grievance) are divided into those sympathetic to the government, those sympathetic to the insurgents, or those who are neutral. If COIN operations are ineffective or indiscriminate, they may target not only insurgents and their leadership, but also the pro-insurgent population and the neutral population as a

whole—thereby breaking the population’s allegiance to and support for the government, and eroding counter-insurgency intelligence gathering capabilities. The problem is compounded further if little or no effort is made by the government to address the grievances of the insurgents and their supporters, allowing the insurgents to gain legitimacy in the eyes of the population (Breede 2006, 25-26). Therefore, in order for COIN operations to be successful, “government must make a real effort to address the grievances of the pro-insurgent population...By increasing the pro-government population base and building trust in the neutral population, intelligence will begin to flow which will allow for the more focused use of force against the insurgents (Breede 2006, 26). According to H. Christian Breede, one way governments (or counter-insurgent forces acting on their behalf) can address grievances is through Civil Military Cooperation (CIMIC) (Breede 2006, 27).

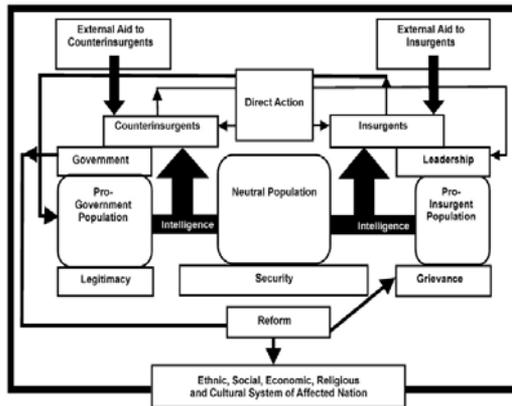


Figure 1: Lynn's Successful Counter-Insurgency Model

While there is always a civil dimension to consider whenever a military deploys,⁹ “in most complex peace operations as well as during successful counterinsurgency operations, civil–military cooperation [has] moved to centre stage, becoming the primary means to reach a combined political goal and establish peace (Zaalberg 2006, 414). Recognizing this, many western militaries have begun to institutionalize CIMIC¹⁰ within their force structures in order to facilitate effective communication and coordinated activity between military and civilian bodies towards a ‘whole of effort’ approach (let alone a ‘whole of government’ approach). At the tactical level, CIMIC’s core functions include: civil-military liaison and assessments of the local civil environment; maximizing civil support for the military force; and committing resources *where necessary* to facilitate the execution of operations and hasten a return to normalcy (Demers 2005, 2). This last function can and often does include the use of Quick Impact Projects, which are usually short-term, small-scale initiatives that are designed to have an immediate impact contributing to post-conflict stabilization or recovery

(Department for International Development 2005, 7). QIPs can be divided into at least two categories: (1) social projects, which include support to education, public health and sports programmes; and (2) infrastructure improvement projects, which include water, electricity and any minor repair or reconstruction of physical structures (Demers 2005, 3). Consistent with the WOG cooperative interdepartmental approach, Canadian Forces' (CF) CIMIC doctrine states that, "Within the Government of Canada (GoC), the planning for CIMIC activities and objectives must be coordinated with [other government departments], particularly Foreign Affairs and Canadian International Development Agency (CIDA)" (Department of National Defence 2006, 6).

In the context of COIN ops, QIPs can be one of the quickest and easiest ways to address the grievances of the local population (at least in the short-term) and thereby win 'hearts and minds.' Destroyed infrastructure, human suffering, economic collapse, and social divisiveness are all endemic in war-affected societies. The accrued benefits with respect to security from the QIPs that tackle these problems can then be used to undertake larger, long-term development projects¹¹ that go even farther in addressing grievances.¹² QIPs are not the exclusive domain of CIMIC; government agencies (or International Organizations [IOs] and Non-Governmental Organizations [NGOs] acting on their behalf) can and do implement QIPs as well—but their use by the military is what remains controversial, as discussed below.

THE DEBATE OVER QIPs

In 2004, retired Canadian Major General Lewis MacKenzie wrote that "soldiers are not social workers with guns. Both disciplines are important, but both will suffer if combined in the same individuals" (Adinall 2006, 50). Over the years, many have echoed this sentiment, arguing against the use of QIPs by the military for two primary reasons. The first is the 'blurring of the lines' effect, whereby militaries undertake many of the same reconstruction activities as humanitarian actors.¹³ In so doing, critics claim aid is used as a tool for implementing foreign policy and crisis management in order to serve the economic and geo-strategic interests of donor states, leading certain local actors to *perceive* that humanitarian actors, by extension, are no longer neutral and independent.¹⁴ The result, they say, is that these civilian actors face rejection or have to deal with high levels of insecurity that reduce their room for manoeuvre (Groupe Urgence Réhabilitation Développement 2006, 4). The second set of criticisms stem from the execution of QIPs. Since the military lacks expertise in aid delivery and development work, civilian actors claim that aid projects undertaken by the military have not proven to be cost-effective, high quality or sustainable, yet these military projects have displaced other aid and development efforts.¹⁵ Moreover, military interventions are subject to political shifts and thus tend to be more short-lived and less reliable than the presence of humanitarian actors, who generally stay as long as it takes to relieve suffering. Similarly, the high turnover of military personnel (e.g. six month rotations for Canadian soldiers) is also

frustrating for civilians, as they constantly have to adjust to working with inexperienced counterparts (Peabody 2005, 7).

Canadian Forces CIMIC doctrine acknowledges that, when it comes to QIPs, NGOs and IOs are often “better funded, more knowledgeable in this field, and better suited to take on this responsibility” (Department of National Defence 2006, 95). However, the doctrine goes on to state that “especially in the earlier and riskier stages of an operation there may be a role for CIMIC projects where a small amount of money, well spent, will do much to ameliorate hardship in the civil sector” (Department of National Defence 2006, 95). There are, after all, times when the military has specific skills, assets, and capabilities that are suitable to QIPs.¹⁶ Nevertheless, the military’s use of QIPs is generally kept to a minimum in both scope and time precisely because CIMIC activities are intended to support the mission goals of the force, not supplant or control the efforts of civilian agencies like CIDA. As much as the military might want to “ameliorate hardship in the civil sector,” its reasons for undertaking QIPs clearly go beyond altruism. As outlined in Lynn’s COIN model, addressing local grievances enables intelligence gathering and prevents the unfocused use of force that could otherwise have “blowback” consequences for force protection.¹⁷ This has tangible benefits as a ‘force multiplier.’ To put it another way, “without good intelligence, a counterinsurgent is like a blind boxer, wasting energy flailing at an unseen opponent and perhaps causing unintended harm. With good intelligence, a counterinsurgent is like a surgeon, cutting out cancerous tissue while keeping other vital organs intact” (Department of the Army 2006, 1-19). Indeed, evidence suggests insurgents themselves recognize that QIPs can have this effect.¹⁸ In response to the criticisms of poor execution, Captain Graham M. Longhurst, argues:

If CIMIC is conducted properly, as trained for by CIMIC operators, there should be little conflict between what the military and the civilian organizations are doing. . . . The only time conflict occurs is when organizations don’t take the time to coordinate the distribution of limited resources. It could be argued that the causes of these conflicts are more as a result of poor training and/or personality conflicts between the organizations (Longhurst 2005, 47).

In any case, a report by the U.S. Institute of Peace that evaluated QIPs with the American-led PRTs noted that, “although there have been reports of PRT assistance projects being poorly implemented, the incidence of this is on par with similar NGO-sponsored projects” (Dziedzic and Seidl 2005, 10). Furthermore, only a small number of respondents to a 2005 survey of NGOs in Afghanistan actually pointed to ‘blurring of the lines’ between humanitarian and military space as the cause of violence directed toward them.¹⁹ Instead, the threat to ‘humanitarian space’ seems to have more to do with the nature of asymmetric warfare. In a report for the Humanitarian Practice Network, Max P. Glaser observes:

Rather than being tactical targets, civilians have moved to the forefront of warfare as objective targets in a deliberate strategy of control. [The commanders of armed non-state actors] have concomitantly adapted their strategies both with regard to civilians and *with regard to [Non-Governmental Humanitarian Agencies] and other humanitarian agents coming to the assistance of civilians* (Glaser 2005, 6).²⁰

Glaser notes that this is particularly true of spoilers, spawned under the ‘War on Terror’ rubric, who seek to “exploit instability and oppose Western occupation. These groups demonstrate little concern for civilians or aid workers, as both are considered proxy targets in the achievement of their goals” (Glaser 2005, 5). Nor does it help that civilians and aid workers are ‘soft targets’ compared to military personnel, who generally have weaponry and training far superior to that of the insurgents. The result is that the provisions of international humanitarian law may be breached and humanitarian access challenged as a deliberate strategy of war by spoilers. In light of this operating environment, CIMIC could serve as a force multiplier not only for the military, but also for civilian organizations working toward common goals (Longhurst 2005, 47).²¹

THE WAY FORWARD: A METHODOLOGY FOR QIP DECISION-MAKING

With respect to the first research objective, it would seem that the compromise is for *all* government agencies (not just the military) to identify and implement only those projects that will have a net positive ‘3D’ impact. Not only must ‘negative’ projects be avoided, but also those projects whose net impact is simply ‘neutral,’ so as not to waste valuable resources. In fact, even projects that have a net positive impact need to be prioritized, as it may not be possible to implement all of them given the scarcity of resources that are usually available to a PSO. Clearly, all this is easier said than done, but at least this sort of WOG approach rejects the principle that the military mission has supremacy irrespective of the short- and long-term consequences for development and diplomacy.

Despite the emphasis on a WOG *impact*, a WOG *process* is implied in this methodology. It is based on the assumption (consistent with the basic premise of this article) that a single dimension, acting on its own, lacks the ‘perspective’ required to achieve a truly ‘3D’ impact in the absence of meaningful consultation with the other dimensions.²² The modalities of this process can take a variety of forms and change over time,²³ but ultimately they must be a function of and responsive to the impacts measured on the ground. After all, the greatest collaborative process is of little value if its virtues fail to translate into a positive impact.

In Figure 2, a matrix is presented for gauging the utility of a QIP with respect to the security and development dimensions. The value -1 signifies a project that will have a *negative impact* on a given mission; the value 0 signifies a project that will have *no substantial impact* on a given mission (positive or

negative); and the value 1 signifies a project that will have a *positive impact* on a given mission. Nine possibilities are given for projects, of which only three have a net positive impact (seen in bold).

		Development Mission		
		Negative	Neutral	Positive
Defence Mission	Negative	-1,-1	-1, 0	-1,1
	Neutral	0,-1	0,0	0,1
	Positive	1,-1	1,0	1,1

Figure 2: QIP Utility Matrix – 2-Dimensional

There are, however, (at least) three dimensions to the WOG approach. As such, the QIP impacts listed in Figure 2 need to be cross-referenced with the impact they will have on the overall political environment, as presented in Figure 3. This time twenty-seven possibilities are given for projects, of which ten have a positive impact (seen in bold).

		Diplomacy Mission								
		Negative			Neutral			Positive		
		Development Mission			Development Mission			Development Mission		
		Neg	Neut.	Pos.	Neg.	Neut.	Pos.	Neg	Neut.	Pos.
Defence Mission	Negative	-1,-1,-1	-1, 0,-1	-1,1,-1	-1,-1,0	-1,0,0	-1,1,0	-1,-1,1	-1,0,1	-1,1,1
	Neutral	0,-1,-1	0,0,-1	0,1,-1	0,-1,0	0,0,0	0,1,0	0,-1,1	0,0,1	0,1,1
	Positive	1,-1,-1	1,0,-1	1,1,-1	1,-1,0	1,0,0	1,1,0	1,-1,1	1,0,1	1,1,1

Figure 3: QIP Utility Matrix – 3-Dimensional

If ranked according to net positive impact, the priority scale for QIPs would appear as follows:

	Defence	Development	Diplomacy	Net Impact Value	Outcome
1	1	1	1	3	'win-win-win' (tridimensional impact)
2	1	1	0	2	'win-win' (bidimensional impact)
3	1	0	1	2	
4	0	1	1	2	'slow and steady' (unidimensional impact)
5	1	0	0	1	
6	0	1	0	1	
7	0	0	1	1	'2 steps forward, 1 step back' (ambivalent impact)
8	1	1	-1	1	
9	1	-1	1	1	
10	-1	1	1	1	

Figure 4: QIP Net Impact Rankings

A few observations must be made. First, each dimension can only be expected to pursue projects from which it will receive a positive impact—unless another dimension is for some reason (e.g. security) incapable of implementing its own projects. Thus, CIMIC would normally only be involved in implementing the QIPs in the ‘defence’ column that are shaded. Second, even when the net impact value is the same, there might still be a need to rank projects according to their ‘dimension.’ For example, in the early stages of a PSO when establishing security is paramount, those projects with a net impact value of 2 *and* those that include a value of 1 in the defence dimension might need to be prioritized over those that do not. Conversely, development might be the dominant dimension for projects of with a net impact value of 2 on more mature PSOs. Third, those projects that include a negative impact on any particular dimension should be ranked lower than those projects of an equal net impact value that do not include a negative impact; a feeling of resentment might develop if two dimensions gain from a particular project at the expense of another dimension, possibly undermining future support for the WOG approach. Fourth, all values listed in Figures 2, 3, and 4 need not be restricted to whole integers. A more sophisticated project assessment approach could make qualitative distinctions to one decimal point in order to further refine the rankings.²⁴ Finally, strides will need to be taken to establish objective criteria (as much as reasonably possible) that can be used to quantify a project as either +1, 0, or -1 (or a value in between) for each dimension.

FROM THEORY TO PRACTICE: KANDAHAR PROVINCE

The situation in Kandahar province very much reflects the model for COIN operations presented earlier. The central government is struggling to consolidate its power and exert influence in the Afghan countryside while a variety of illegal, armed spoiler groups disrupt the nation-building process.²⁵ Similarly, the population can be divided into three groups: the first comprises hardcore Taliban, Al-Qaeda, and criminal elements (pro-insurgent); the second group is the vast majority of the Afghan population who desire security and stability, regardless of who provides it (neutral); and the third group is the central government and its supporters (pro-government) (Eyre 2006, 82). Rather than employing its own COIN forces, the weak central government has had to invite foreign militaries into Afghanistan to fight the insurgency until such time as the Afghan National Army and Afghan National Police are strong enough to do so. This decision has not been without problems in Kandahar province, where rural Pashtuns have a deep-seated xenophobia and an innate hostility towards foreign occupation. According to Gordon Smith:

Reaction to the presence of armed forces has prompted a reflexive and violent resistance in the isolated rural south, especially as NATO troops conduct invasive anti-Taliban actions in the villages. The active combat mission, designed to seek out and destroy both dormant and operational Taliban forces, has resulted in large

civilian casualties... Because of the loss of civilian life, these actions have, on the whole, fostered increased support for anti-NATO and anti-government insurgency, including support for the Taliban....[v]ictimimized Kandahari tribes will also seek revenge for the death of their family members. [The now deceased] Mullah Dadullah precisely expressed this vicious circle with his statement that “for every Talib you kill, I can recruit twenty more” (Smith 2007, 16).

As outlined in Lynn’s model, indiscriminate attacks of this nature by foreign forces threaten to produce a self-defeating vicious cycle by breaking the population’s allegiance to and support for the central government, and by eroding counter-insurgency intelligence gathering capabilities of NATO forces. The other half of the vicious cycle equation arises from the government’s inability to address the grievances among the population:

[I]nsecurity obstructs reconstruction; blocked reconstruction fostered insecurity; and security and reconstruction difficulties prevented sufficient progress in the area of governance reform for the promotion of public order and significant economic progress. By the fall of 2006, such interdependent difficulties resulted in a highly deteriorated situation...” (Smith 2007, 20).

This insecurity has meant that few civilians are engaged in aid and development in the province, and NGOs are leaving because of the same concern for their safety.²⁶ Coupled with a severely flawed policy towards poppy production,²⁷ the situation in Kandahar province represents a textbook scenario for implementation of QIPs by CIMIC.

PRT KANDAHAR: A QIP CASE STUDY

Structure and Roles

Canada took over command of Kandahar’s Provincial Reconstruction Team (PRT) from the United States in August 2005. The PRT’s composition includes officials from the Canadian International Development Agency (CIDA), the Department of Foreign Affairs and International Trade (DFAIT), the Royal Canadian Mounted Police (RCMP),²⁸ Correctional Services Canada, and the Canadian Forces—consistent with the WOG approach. Within this context, a division of labour seems to have naturally developed. CIDA and DFAIT interact with Afghan government officials at the ministerial level (both provincial and national) and implement large projects over the medium- and long-term that have a significant impact on development. Since all major development projects are implemented centrally through the Afghan government (to build its capacity and enhance its legitimacy), very few staff from CIDA and DFAIT are needed in the PRT for oversight of these projects and minimal travel is required outside of Camp Nathan Smith (where the PRT is based). Meanwhile, the military (through CIMIC) interacts with locals at the district level (e.g. mayors, chiefs of police) and implements QIPs in order to meet the immediate needs of locals in the short-term and build their confidence in the Canadian effort—effectively buying time

until the larger projects can come online (Chase 2007). Implementation of the QIPs is completely decentralized, meaning that the number of projects that can be undertaken at any given time is limited by the number of CIMIC teams and their mobility outside of the camp.

The justification for this particular division of labour appears to be three-fold. First, the security situation makes it too dangerous for staff from CIDA and DFAIT to be ‘outside the wire’ carrying out decentralized projects for extended periods of time. Second, the number of CIDA and DFAIT (*vis-à-vis* CF) personnel available for deployment to the PRT is insufficient to meet the demand that exists for QIPs. Finally, the small-scale and limited scope of QIPs means that soldiers, rather than civil servants, can be trained to execute them without much difficulty. The same is not necessarily true of CIDA’s and DFAIT’s large-scale projects, which require greater expertise. In any case, all departments were subject to a learning curve, with numerous examples of CIMIC, CIDA, and DFAIT projects alike that had less than successful results (Howard 2007). Moreover, no department had any particular advantage in terms of knowledge of the local context; CIDA gained most of its information from locally engaged staff—a practice that the CF mirrored as well (Howard 2007).

Decision-Making Processes & Principles

Four budgets are available to fund CIMIC’s QIPs: the Commander’s Contingency Fund (CCF) and the departmental budgets of CIDA, DFAIT, and USAID,²⁹ respectively. A ‘Targeting Board’ composed of the PRT commander and the various leadership positions in the PRT (including at least one representative from each government department) is convened periodically in order to approve large projects of strategic value and determine the most appropriate budgetary source. This is a key step in the coordination process. Not only do all the stakeholders come away with increased situational awareness, but projects are exposed to various ‘lenses’ in order to consider second- and third-order effects. If projects are not feasible, they are either rejected or sent back for retooling. However, the Targeting Board only reviews projects above CDN\$5000; anything costing below that amount (which includes the vast majority of QIPs) can be authorized by the PRT commander alone and funded from his/her contingency fund. Indeed, this is the primary purpose of the CCF (Chase 2007). The advantage is that it saves both time and bureaucracy. After all, the very purpose of a ‘quick impact project’ is defeated if several months are spent filling out the paperwork that is required by Treasury Board regulations. QIPs can be successful precisely because they ‘buy time’ until the larger projects that are subject to stricter standards of accountability (sometimes valued in the millions of dollars) can be approved and implemented.³⁰ Yet this is not meant to imply that CIDA and DFAIT have no knowledge of or input into those QIPs costing less than CDN\$5000. CIMIC teams deployed to the field (sometimes for weeks at a time) send back daily reports to the PRT that are distributed to all relevant personnel by the operations officer. Moreover, the officer in command of CIMIC usually asks other departmental representatives informally for their

input on projects not going to the Targeting Board (Tondreau 2007). However, this system is not perfect; one of the problems with these daily reports is insufficient communications equipment,³¹ the result of which is that CIDA and DFAIT would frequently learn about QIPs after the fact and thereby have no opportunity to offer input.³²

Since Canada took over the PRT, principles such as capacity-building, sustainability, and ‘do no harm’ are always kept in mind during the planning process of a QIP,³³ which indirectly leads to consideration of the net impact. U.S. officials working in Kandahar, however, were less effective at this. Rather than going through mayors and *wakils* (village elders) to implement QIPs, which allows local leaders to share in the credit, reinforces their authority in the eyes of their people, and has a net positive impact for the diplomacy and governance dimension, American policy insists on ‘showing the flag’ and maximizing the credit for the U.S. government. Without this sort of local ‘buy in,’ there were cases of schools being built “whether people wanted it or not” (Chase 2007). In another extreme case, the Canadian PRT actually declined a U.S. offer of tractors after American officials had refused to let the farm equipment go through the *wakils* for distribution (Chase 2007).

Measuring Impact

CIMIC teams supervise the implementation of QIPs and perform follow-up assessments after a period of one or two weeks. In the case of infrastructure, for example, pictures are taken and a site report is filled out. Yet there are no explicit measures of effectiveness (MOEs) in use: “we learned about [MOEs] during our training, but we never used them. If we’re building an irrigation canal and water goes through it, the project is effective” (Tondreau 2007). After all, the Canadian approach to QIPs stresses consultation with the locals—usually through a *shura* (council)—during which complicated issues that might have unintended consequences are usually worked out. Thus, if the project is functional and the villagers do not complain about it afterwards, it is deemed ‘effective.’ Another sign of a successful CIMIC project (at least for the defence dimension) is the amount of intelligence received from the local people, such as the locations of hide outs, weapons caches, and so on.³⁴ The biggest hurdle to employing sophisticated MOEs is time:

There were days when we were working 24 hours straight. Other days were slower, but even those were 10 to 12 hours long. We could use measures of effectiveness, but they would have to be streamlined. If they took more than 15 minutes to complete, you risk not having them complied with (Tondreau 2007).

Moreover, there is effectively no incentive on the part of CIMIC teams to employ MOEs so long as projects adhere to the aforementioned basic principles, which are thought to avoid clearly negative impacts. Distinguishing between different degrees of impact and efficiency is unnecessary because:

There is no shortage of money [to build QIPs]. When it ran out, we simply got more from the Canadian government. There is no need to prioritize like we do for the big projects [implemented by CIDA or DFAIT]. We're just trying to get them out the door as fast as we can (Howard 2007).

This results-based or 'shot-gun' approach assumes that the aggregate positive impact of *lots of projects*—even if many are by themselves marginal—is higher than that of a more deliberate, effects-based approach with *fewer projects* that are of higher individual quality. Thus, even if some projects have a net *negative* impact under the results-based approach due to hasty assessments or planning, the sheer volume of projects with a net *positive* impact will more than compensate. This sentiment is echoed by Rochus Pronk, the former political advisor to the Dutch PRT in neighbouring Uruzgan province: "Let's not make rocket science out of [quick impact projects]. They are meant to be quick. ...The military is not just handing out money regardless of the consequences. It is responsible with its spending" (Pronck 2007).

POLICY IMPLICATIONS

With respect to the second research objective, the evidence presented in this article suggests that the military and civilian agencies of the Canadian government have made progress in coordinating their efforts. The division of labour between military and civilian actors seems sensible, and the military makes attempts to seek input from the other two dimensions in order to improve its own situational awareness and enhance the net impact of its projects. While every project would be collectively vetted in an ideal world, the informal cooperation mechanisms that currently exist below the Targeting Board are likely the best that can be hoped for given the operating environment and the nature of QIPs. With proper training ahead of time in project management and cultural awareness, the evidence from Roto 2³⁵ suggests that there is no reason why CIMIC teams cannot achieve a net positive '3D' impact with QIPs that is comparable to (if not better than) what CIDA or DFAIT staff could achieve under the same circumstances. However, the informal opportunities that do exist for cooperation need to be properly exploited—sufficient communications equipment must be provided to all elements of CIMIC both inside and outside Camp Nathan Smith, and a liaison officer must be appointed to brief and debrief the CIMIC teams whenever they come back in from the field for re-supply.

The practice of having a self-contained military budget for QIPs, separate from any cross-government arrangements, appears a necessary one. While a jointly managed budget for QIPs might foster further cooperation between the dimensions, it would only be of value if the streamlined procedures were kept intact:

One of the lessons learned...has been the need for expedited project implementation over the spectrum of deployment and not just at the end of a given rotation. The reality on the ground is that the period between project

identification and implementation can be considerably drawn out due to the lengthy administrative oversight required to utilize CIDA funds. Projects identified at the start of tour were implemented (at best) by the third and fourth month, whereas projects under the force protection umbrella were completed in a much shorter time span. While the government's 3D approach stresses coordination, it does not imply subservience of one branch to another (Skidd 2005, 7-8).

Although accountability to the Canadian taxpayer is important, the bureaucratic processes that characterize peacetime actions need to be modified when operating in a quasi-war zone where people's lives are at risk. One suggestion for improved effectiveness is to increase the threshold for Targeting Board review from \$5000 to between \$10,000 and \$15,000 (Howard 2007).³⁶

In any case, it seems one of the obstacles to *interdepartmental* synergy has been a lack of *intradepartmental* synergy within the CF. According to Gordon Smith, "the Canadian Forces have a fundamentally different agenda than does the Canadian International Development Agency (CIDA)" (Smith 2007, 6), but this is not entirely true. The CIMIC operators who are on the ground interacting with the locals recognize that alleviating poverty must become a primary strategy for achieving security in Afghanistan. As Warrant Officer Dean Henley remarked, "You can't win the insurgency by killing insurgents. You can't kill the insurgency, you have to take away its support base" (Day 2007). Yet CIMIC operators feel there is a divide between the tactical and strategic level because the advice given from the bottom up is not followed (Longhurst 2005, 44).³⁷ Solving this problem requires not only spending more aid money in Kandahar province where Canada's major military commitment is,³⁸ but also refocusing an overly aggressive military strategy that has led to far too many innocent civilian deaths.³⁹ The CF should attempt to expand the size of the PRT by creating "satellite PRTs" in the 15 districts of Kandahar province outside the capital city.⁴⁰ Indeed, the Senlis Council reports that in the neediest areas, such as the remote villages and refugee camps in Kandahar province, the inhabitants have neither seen a foreigner nor heard of CIDA, let alone received food or any other aid from the Canadians (Senlis Council 2006, 18). While some have rightly made the argument that the size of the foreign military presence in Afghanistan has been well below what might have been expected,⁴¹ increasing the troop level in and of itself will be of little value if indiscriminate attacks persist and the local population continues to shift its support to the insurgents. Any increase in troop levels needs to be accompanied by an even greater proportional increase in the size and capabilities of PRT Kandahar.⁴² After all, as Maslow famously remarked, "If the only tool you have is a hammer, you tend to see every problem as a nail."

FUTURE RESEARCH

The absence of measures of effectiveness remains the thorniest issue. Anecdotally, QIPs seem to have a net positive impact, even across dimensions,

but this cannot be concluded with any certainty. This is because the methodology outlined earlier in this article is premised on an effects-based approach. This is not present (nor even encouraged) in PRT Kandahar at the tactical level, where the CIMIC element has been given *carte blanche* by the Canadian government. Presumably, the current results-based approach is preferred by politicians since the output of QIPs is higher; even if these QIPs are less effective, the absence of MOEs means it is difficult to conclusively prove this—meanwhile the Canadian government can be *seen* conducting copious amounts of reconstruction and thereby bolster public opinion for the Afghanistan mission. While it is quite unlikely that the results-based approach is more efficient than the effects-based one,⁴³ it is entirely possible that the results-based approach could achieve a greater net positive 3D impact if provided with unlimited resources. While citizens in a liberal democracy are entitled to consciously support inefficient policies when other values are prioritized (e.g. Europeans and the Common Agricultural Policy), they must be able to do so with the benefit of full information. Yet at present, Canadians have no idea exactly how effective or efficient QIPs have been. Canadians cannot know, nor can this author's methodology be effectively tested, until the PRT employs MOEs for its QIPs. Nor, by extension, can informed policy recommendations be made for improving the WOG coordination process within the PRT (beyond the relatively straightforward ones made in the previous section) without a firm grasp of the impact that the current processes are having. Consequently, the most pressing research agenda must be to develop valid and reliable MOEs that are sufficiently streamlined for rapid use by CIMIC teams deployed in the field. After all, it is doubtful whether the unlimited financial support for QIPs can carry on indefinitely, especially if the CIMIC elements of the PRT are greatly expanded, as recommended earlier.⁴⁴

Once MOEs are developed, a relatively straight forward comparison between the results-based and effects-based approaches could divide up the districts of Kandahar province between the two methods. What's more, within the sample of effects-based approaches, further refinement of this author's methodology for evaluating 3D processes could be conducted. For example, does a dimension for local input need to be *explicitly* added to the methodology? Local input is only *implicit* in the current methodology—that is to say, if a particular QIP is not a priority for the local population, the net impact value for each dimension is unlikely to be very, if at all, positive. By the same token, do other dimensions also need explicit consideration with their own dimension (environment, human rights, gender, etc.), or are these simply sub-elements of development? How many dimensions can be considered before the methodology becomes too unwieldy? What sorts of timelines are required before the impact of each type of project can be effectively measured?

More broadly, a similar methodology needs to be developed to study how well QIPs are being coordinated with respect to the 'whole of effort' approach.⁴⁵ Initial evidence of the coordination between elements of PRT Kandahar and other civilian actors operating in the area suggests this has been limited.⁴⁶ The

additional (and much more problematic) variable in assessing the ‘whole of effort’ approach is the lack of a superordinate authority. Unlike the WOG approach, in which all government departments are subordinate to the will of the Canadian cabinet (which insists on unity of effort), is it feasible for this methodology to create dimensions for actors such as NGOs over whom there is no executive authority? Perhaps the best that can be expected is a multi-tiered, or ‘almost whole of effort’, approach in which true ‘coherence’ happens only with certain NGOs who volunteer to coordinate their efforts with the military, while nothing more than simple ‘awareness’ can take place with other NGOs like Médecins Sans Frontières who wish to maintain complete independence from anything associated with military operations.⁴⁷

CONCLUSION

As the local environment improves in some districts, it is likely that PRT Kandahar will be able to refocus its efforts away from QIPs and concentrate instead on longer-term projects that strengthen the capacity of the provincial government. Yet many districts in Kandahar province remain unvisited by the PRT, meaning that QIPs will most likely remain in use for the foreseeable future in order to fill the initial void in the flow of reconstruction assistance to these new areas.

QIPs (and the PRT structure that engenders them)⁴⁸ are valuable components of the WOG toolbox that will surely be used again in future Canadian PSOs to other failed and fragile states. Despite the natural aversion of certain organizations to cooperate, the imperatives of development, diplomacy, and defence are rarely mutually exclusive. As more practitioners recognize this and use it as a starting point for discussion, the task of harmonizing military activities with both the desires of the local population and the activities of the rest of the international community will become easier.

NOTES

¹ PSOs are defined as “*multifunctional operations in which impartial military activities are designed to create a secure environment and to facilitate the efforts of the civilian elements of the mission to create a self sustaining peace*.” PSOs may include peacekeeping and peace enforcement, as well as conflict prevention, peacemaking, peacebuilding and humanitarian operations” (emphasis in original, Peace Missions Programme 2000, 27).

² PRTs are small joint civil-military teams that first began deploying across Afghanistan in early 2003, in an attempt to expand the legitimacy of the central government, enhance security, and facilitate reconstruction. Although only lightly armed, their very presence is intended to serve (1) as a deterrent to insurgents and criminals who have been active in these areas and (2) as a catalyst by enabling local government authorities, the UN, and NGOs to operate (Olson 2006, 5). Each of the individual implementing countries has developed its own PRT designs in an attempt to customize the concept to meet local conditions—with varying degrees of success.

³ Although donor governments in the developed world can agree on neither the policy prescriptions to address failed states, nor how to prioritize failed states, nor a single definition for ‘failed state,’ there is consensus that policy responses should use the “whole of government” (WOG) approach.

The 2005 International Policy Statement has placed the Canadian government firmly within this consensus (Prest, Gazo, and Carment 2005).

⁴ Picciotti et al. identify three other levels of coordination in fragile state policy besides WOG: Intra-departmental rationalization, Inter-donor harmonization, and Donor-partner alignment (Prest, Gazo, and Carment 2005, 17-18). Together, these constitute a ‘whole of effort’ approach. While all levels are important (individually and collectively) and require further investigation in the context of the Canadian PRT in Kandahar, WOG coordination will be the focus of this article.

⁵ For, for example, Lind (1989).

⁶ Stedman classifies various types of spoilers, the most dangerous being ‘total spoilers,’ who are innately opposed to central authority and are unappeasable. The only method to constrain and ultimately uproot such implacable figures appears to be the application of sustained pressure, most likely involving the use of force (Sedra 2003, 2). Durch has argued that in Afghanistan, the Taliban and al Qaeda constitute ‘total spoilers’ (Durch 2003, 9).

⁷ According to Breede “insurgencies are characterized by a desire to achieve some form of societal and political change. There is a grievance amongst the population and insurgencies form as a reaction to it.” This is distinct from a guerilla movement, which is purely a reaction to an external invasion without any binding ideology (except patriotism) or greater political end above and beyond the ousting of foreign (or in some cases even domestic) aggressors (Breede 2006, 25). Both of these are distinct from paramilitary groups and criminal gangs, who may not even have a legitimate grievance. In many cases, they are merely taking advantage of the insecure environment that has been fostered by an insurgency for their own gain.

⁸ In the case of a failed or fragile state like Afghanistan, the counter-insurgent forces may be foreign militaries deployed on a PSO. Under these circumstances, insurgents have a huge tactical advantage because they are locals who speak the same language as the villagers, have the same history, and are often related by blood (Day, 2007).

⁹ It may be in the form of the local population, refugees, local officials, or workers from international organizations. Moreover, this element exists not only in PSOs and COIN operations, but also in humanitarian assistance and warfighting deployments [Longhurst 2005, 35].

¹⁰ According to the NATO definition, CIMIC is “the co-ordination and co-operation, *in support of the mission*, between the NATO Commander and civil actors, including national population and local authorities, as well as international, national and non-governmental organizations and agencies” (emphasis added, NATO Standardization Agency 2003, 1-1).

¹¹ There remains a debate—one that goes to the heart of the WOG approach in failed and fragile states—over the practicality of even undertaking long-term development work in a ‘war zone.’ As one aid worker put it, “PRTs doing school reconstruction doesn’t make a lot of sense. If the security situation is so bad that an NGO can’t do this work, then kids can’t go to school anyways” (quoted in Olson 2006, 24). Yet even if large-scale, long-term projects should not be pursued simultaneously with a COIN operation, there is still arguably a ‘force multiplier’ effect (discussed further in the following section) that small-scale, short-term QIPs can have for the military regardless of their consequences for development.

¹² Arguably, QIPs are one of the quickest and easiest ways to curry favour with *foreign* populations as well. In the case of Afghanistan, where foreign military forces are carrying out COIN operations on behalf of the threatened Karzai government, the support of the domestic population in each NATO member state is often conditional upon seeing evidence of tangible progress. In this regard, foreign populations can be divided along similar lines as local populations in a COIN environment—pro-intervention, neutral, and anti-intervention. QIPs can help shield foreign governments from anti-interventionist criticisms of the mission until long-term development projects show results.

¹³ Humanitarian actors have traditionally provided *humanitarian aid*, which consists of material or logistical assistance provided in response to a crisis in order to save lives, alleviate suffering, and maintain human dignity. This can be distinguished from *development aid*, which seeks to address the underlying socioeconomic factors (such as poverty or a lack of human rights) that may have led to a crisis or emergency. Moreover, development aid is typically not made available in situations of protracted crisis because the political framework for development is absent, nor is it necessarily bound by humanitarian principles like impartiality and neutrality. Virtually all of the QIPs

undertaken by PRT Kandahar province development, rather than humanitarian, assistance (Howard 2007).

¹⁴ Taliban and other anti-government elements in Afghanistan are reported to have told locals that Provincial Reconstruction Teams (PRTs) are really 'spies' and should not be cooperated with, which has clear implications for the humanitarian actors who are seen to work closely with the PRTs. In one region, for example, local partners of the international NGO CARE were reportedly approached by the Taliban and told that as long as they continued to operate in the same way, they would not be targeted; but, if they accepted funding from the military-led PRT, "their security might be threatened" (Olson 2006, 14-15).

¹⁵ In Afghanistan's Badghis province, for example, one of CARE's local partners had started up a micro-loan business that charged 10% interest as part of a long-term community project. The PRT came in and set up a short-term loans project with no fee, which attracted many people to what CARE considers a less sustainable option. Furthermore, sometimes differences between military-run aid projects and those run by NGOs have fueled distrust and cynicism towards NGOs, who often ask people to contribute labour and other local resources in order to build in local ownership of a project. When military-led projects in the same area refrain from requesting such inputs, rumours start that the NGOs have simply stolen money and materials meant for the project (Olson 2006, 17).

¹⁶ According to the British Department for International Development, these include: 1) when there is a gap which cannot be filled by another actor; 2) where value is added for the recipient community; 3) when it is acceptable to the local population and culture; 4) when it is planned in such a way to support other assistance efforts; and 5) where the military has a specific comparative advantage [Department for International Development 2006, 8]. There is also the argument that CIMIC could not effectively execute its other roles (such as liaison) *without* the use of QIPs: "To simply go to meetings and try to build friendships without being able to help solve the issues . . . generates distrust on the part of the locals and is no way to build a trusting relationship" (Longhurst 2005, 44).

¹⁷ The CF defines force protection as "comprising all measures taken to contribute to mission success by preserving freedom of action and operational effectiveness through managing risks and minimizing vulnerabilities to personnel, information, materiel, facilities and activities from all threats" (Skidd 2005, 7).

¹⁸ In September 2004, for example, Canadian CIMIC operators in Kabul were advised that opposing military forces intended to specifically target them due to the success their projects were having in winning 'hearts and minds' (Longhurst 2005, 45).

¹⁹ Indeed, M.H. Mayar, deputy director of ACBAR (the NGO coordination body in Afghanistan), points to the public perception that the work of humanitarian actors directly supports the Karzai government: "NGOs are seen to be helping the government because they are working to improve people's lives. If the lives of regular Afghans improve, more ordinary people would support the government and this is what the insurgents do not want" (Olson 2006, 15-16).

²⁰ Emphasis added.

²¹ It should also be noted that, regardless of the emerging asymmetric threat, the Government of Canada has issued guidelines for CIMIC that state: "To the greatest extent possible, CF operations should be conducted with a view to respecting the humanitarian operating environment" (Government of Canada, Guidelines, 2005).

²² This does not mean that *implementation* of a given project always has to be shared by multiple government agencies if the planning process (in which all departmental stakeholders are involved) deems it unnecessary. Yet if the plan needs to be amended during the implementation phase, all stakeholders should be consulted.

²³ Julie Boileau and Richard Garon identify four models for civil-military interactions—Independent, Collective, Top-down, and Bottom-up—the first three of which are suitable to military stabilization missions (*vis-à-vis* non-military ones) like the one taking place in Kandahar province. While the stated goals of Canada's WOG approach resemble the collective model, Boileau and Garon claim the present CF approach to CIMIC is more akin to the stove-piped aspects of the independent model (Boileau and Garon 2005).

²⁴ If this were to happen, an additional category of outcomes would have to be created in the table presented in Figure 4: "1 *large* step forward, 2 *small* steps back" (unidimensional ambivalent impact).

For example, a QIP could have a military impact of +0.8, but development and diplomacy impacts of -0.2 and -0.3, respectively. This project would still be acceptable because it results in a net positive impact of +0.3. Nevertheless, caution must be taken not to build a sense of added accuracy into the methodology that is not actually measurable in reality.

²⁵ Some of these spoilers are anti-government insurgents, while others are paramilitary groups or criminal gangs (the distinction between which was described earlier) (Smith 2007, 15).

²⁶ Numerous Afghan organizations (such as the Danish Committee for Aid to Afghan Refugees) have halted all operations in the south, as have large international agencies like Oxfam (Smith 2007, 18).

²⁷ The British, American, and Afghan governments have sought to eradicate what is frequently the population's principal livelihood and sole source of desperately needed income. While the Taliban and warlords benefit from the poppy industry, these policies of forcible crop destruction only play into the hands of the Taliban (Smith 2007, 14).

²⁸ CivPol actually includes officers from other provincial and municipal police forces—not just the RCMP. Administratively, they come under the authority of DFAIT (Howard 2007).

²⁹ The United States, which ran the PRT prior to Canada, still has a USAID representative with a discretionary fund for projects attached to Canada's contingent. A fifth budget, the Assistance To Afghanistan Fund, is operated solely by the CIMIC chain of command. It consists of charitable donations made by service clubs, businesses, and ordinary Canadians back home who wish to have specific projects undertaken in Afghanistan (Chase 2007).

³⁰ Even some QIPs over \$5000 do not pass through the Targeting Board if the CIMIC staff and the PRT commander feel they are relatively "straight forward," i.e. there would be no value-added by having additional input from the other agencies in the PRT (Howard 2007).

³¹ A shortage of Tactical Satellite Kit (TSK) radios meant that CIMIC teams had to use the Battle Group's TSKs if they were in the vicinity. Even when these could be borrowed, the CIMIC leadership in the PRT did not have their own TSK either, meaning that messages had to be relayed through the Battle Group or the higher chain of command in the PRT. When TSKs were not available, cell phones were used, but their signals were not encrypted (if a signal was available at all) (Tondreau 2007).

³² This communication deficit was a two-way street, with poor passage of information from the operational level to the CIMIC teams deployed outside the wire. When these teams did return to camp, they were not always debriefed on the activities of the other government departments (Tondreau 2007).

³³ For example, past mistakes from QIPs in the Balkans, where soldiers themselves completed all stages of a project while community members watched [Peabody 2005, 10], have not been repeated. Projects in Afghanistan are almost always implemented by locals themselves, with CIMIC simply providing the funding (Tondreau 2007).

³⁴ While intelligence was often gained as a result of a QIP, it must be noted that this was never a condition for its provision in the first place by Canadian CIMIC operators. The primary purpose of QIPs is to win the 'hearts and minds' of the locals by addressing their needs, thereby encouraging their acceptance of the international presence and their patience for larger development projects to come. Any intelligence that is provided is simply a welcome side-benefit: "CIMIC tries to be passive intelligence collectors. If it comes, great. But we can be targeted if we're seen otherwise" (Chase 2007).

³⁵ This refers to the rotation of Canadian soldiers to PRT Kandahar during the period August 2006 to February 2007.

³⁶ Indeed, the Dutch PRT in neighbouring Uruzgan province has a mechanism similar to the Targeting Board, but its threshold is €10,000 (Pronck 2007).

³⁷ Many of those interviewed by the author noted that this barrier decreased over time. By the end of Roto 2, the Battle Group commander not only recognized the value of 'non-kinetic' operations, but even insisted on using CIMIC as much as he could. Whether this realization carries on into Roto 3 remains to be seen.

³⁸ Only 20% of CIDA's \$100 million budget for Afghanistan will be spent in Kandahar province (Smith 2007, 18). A recent Senate Report went so far as to conclude that CIDA should channel more

of its budget directly through CIMIC for QIPs until NGOs are able to safely function in Kandahar (Standing Senate Committee on National Security and Defence 2007).

³⁹ The U.S. Counterinsurgency Manual identifies nine ‘paradoxes of counterinsurgency,’ the first two of which are extremely noteworthy: “The More You Protect Your Force, The Less Secure You Are,” and “The More Force Used, the Less Effective It Is” (Department of the Army 2006, 1-22). The Dutch appear to have adopted these principles in Uruzgan province (Chivers 2007).

⁴⁰ By the end of the Roto 2, CIMIC teams were operating primarily in the districts of Zhari and Shah Wali Kot, but there was also a presence in Spin Boldak, Argendat, Daman, Maywand and Dand, which includes Kandahar City (Tondreau 2007). However, this still leaves nine districts without any direct assistance from the PRT.

⁴¹ See, for example, Smith (2007). There is indeed a case to be made for additional troops to provide general security. As Zaalberg points out, QIPs in and of themselves cannot win ‘hearts and minds,’ at least not in the long-term; locals also need to believe that COIN forces will protect them against reprisals and intimidation before they can truly turn against the insurgents (Zaalberg 2006, 412). Currently, there are insufficient troops to provide general security in a province as large as Kandahar (roughly equivalent to the size of Nova Scotia). If Canada had the resources, it would be better to provide an entire brigade—just as the British have done in neighbouring Helmand province—instead of a mere battle group (Chase 2007).

⁴² Although the Dutch have integrated their Battle Group in Uruzgan province with the PRT in order to increase the mobility of its CIMIC teams (Pronck 2007), some have argued that it might be too risky for Canada to decentralize its Battle Group in the same way given the security environment in Kandahar province. If Canada is unwilling to re-task its Battle Group assets, then more reinforcements from Canada—or even freshly graduated ANA units—will be required in order to provide the additional force protection required for the mobility of an expanded PRT.

⁴³ It is conceivable that the PRT’s CIMIC teams are so skilful that they can achieve high impact QIPs intuitively—without expending many resources on extensive assessments. Alternatively, if the Canadian mission can end sooner or if fewer Canadian lives are lost with a results-based approach, the short-term trade-off of efficiency for effectiveness could lead to efficiency gains in the long-term.

⁴⁴ The development of MOEs could spark a further policy recommendation for PRT expansion—the addition of an ‘assessment cell’ dedicated strictly collecting and analyzing data on QIPs on an ongoing basis, as well as improving the MOEs themselves over time.

⁴⁵ The other level of coordination identified by Picciotti et al.—inter-donor harmonization—is less of an issue for the PRTs as each one is either primarily or entirely composed of military and civilian staff from the same country.

⁴⁶ Savage argues that communications between the PRT and these bodies occur on an ad hoc and random (rather than formalized) basis, and that meetings with UNAMA and the Disbandment of Illegal Armed Groups (DIAG) committee have resulted in only information sharing rather than true coordination (Savage 2006, 59).

⁴⁷ NATO classifies four levels of coordination on its ‘spectrum of interaction’—awareness, deconfliction, cooperation, and coherence. In order to be flexible and accommodate the diverse interests it encounters in an operational theatre, NATO has developed a number of different coordination mechanisms across this spectrum that cater to non-NATO actors who seek either higher or lower levels of coordination (Icayan 2007).

⁴⁸ This implies that the military should avoid implementing QIPs in the absence of a WOG structure like the PRT—a principle which is consistent with the premise that WOG approaches achieve better results when addressing the problems of failed and fragile states.

ACKNOWLEDGEMENTS

I would like to acknowledge all those who read drafts of my paper and provided feedback, including Dr. David Carment, David Vogt, Bonnie Butlin, John Cadham, and Ashlyn Milligan. I would also like to give special thanks to the Canadian soldiers who shared with me their first-hand experiences from the field, particularly MWO Denis Tondreau, WO Sean Chase, and Sgt Ted Howard.

REFERENCES

- Adinall, Robert. 2006. Transformations of War and Public Perception: Implications for 21st Century Warfare. *Canadian Army Journal* 9(1): 50-67.
- Boileau, Julie, and Richard Garon. 2005. Who is in charge? The need for a comprehensive model of Civil-military cooperation. *Conference of Defence Associations Institute*.
<http://www.cda-cdai.ca/symposia/2005/Boileau.pdf> (accessed February 13, 2007).
- Breede, Captain H. Christian. 2006. Intelligence Lessons and the Emerging Canadian Counterinsurgency Doctrine. *Canadian Army Journal* 9(3): 24-40.
- Chase, Warrant Officer Sean. 2007. Interview by author. Telephone. Ottawa, March 27.
- Chivers, C.J. 2007. Dutch Soldiers Stress Restraint in Afghanistan. *New York Times*, April 6.
- Day, Adam. 2007. Task Force Afghanistan: The Battle For The People. *Legion Magazine*. <http://www.legionmagazine.com/features/militarymatters/07-01.asp#1> (accessed January 29, 2007).
- Demers, Major A. 2005. To Help or Not to Help—CIMIC and Project Management. *The Bulletin*, 11(1): 1-5.
- Department for International Development. 2005. Quick Impact Projects: A handbook for the military. *DFID*. <http://www.dfid.gov.uk/pubs/files/qip/booklet.pdf> (accessed February 15, 2007).
- Department of National Defence. 2006. *Civil-Military Cooperation Tactics, Techniques and Procedures*. Fort Frontenac, Kingston, Ontario: Army Publishing Office.
- Department of the Army. 2006. Counterinsurgency. *Federation of American Scientists*. <http://www.fas.org/irp/doddir/army/fm3-24fd.pdf> (accessed March 20, 2007).
- Durch, William J. 2003. Peace and Stability Operations in Afghanistan: Requirements and Force Options. *The Henry L. Stimson Center*. <http://www.stimson.org/fopo/pdf/afghansecurityoptions070103.pdf> (accessed November 26, 2006).
- Dziedzic, Michael J., and Colonel Michael K. Seidl. 2005. Provincial Reconstruction Teams and Military Relations with International and Nongovernmental Organizations in Afghanistan. *United States Institute of Peace*.
<http://www.usip.org/pubs/specialreports/sr147.pdf> (accessed March 18, 2007).
- Eyre, Lieutenant-Colonel Wayne D. 2006. The Role of the Infantry in the War of the Snakes. *Canadian Army Journal* 9(3): 80-90.
- Gizewski, Peter. 2007. Collaborative Responses to Conflict: Toward a 'JIMP-Capable' CF. Address made at the CIIA National Foreign Policy Conference, Montreal, March 22.
- Glaser, Max P. 2005. Humanitarian Policy Group Report – Humanitarian engagement with non-state armed forces: The parameters of negotiated access. *Overseas Development Institute*. No. 51
<http://www.reliefweb.int/library/documents/2005/odi-gen-30jun.pdf> (accessed November 26, 2006).
- Government of Canada. 2005. Government of Canada Guidelines on Humanitarian Action and Civil-Military Coordination. Unpublished Paper. 2005. Overview – Canada's International Policy Statement. *DFAIT*. <http://geo.international.gc.ca/cip-pic/ips/overview-en.asp> (accessed January 13, 2007).
- Groupe Urgence Réhabilitation Développement. 2006. Humanitarian Space in Jeopardy? *Urd.org*.
http://www.urd.org/fichiers_urd/pole_dactivites/evenements/UAH/UAH2006/Issues_discussion_UAH2006.pdf (accessed November 22, 2006).

- Howard, Sergeant Ted. 2007. Interview by author. Telephone. Ottawa, March 28.
- Icayan, Tony. 2007. Enabling the 'Future Comprehensive Civil-Military Interaction (FCCMI) Concept'. Address made at the Pearson Peacekeeping Centre, Ottawa, February 26.
- Lewis, Colonel Fred A. 2006. The Ability to Do Old Things in New Ways—Counter-Insurgency and Operational Art. *Canadian Army Journal*, 9(3): 6-23.
- Lind, William S., et al. 1989. The Changing Face of War: Into the Fourth Generation. *Marine Corps Gazette*: 22-26.
- Longhurst, Captain Graham M. 2005. Evolution of Civil-military Cooperation (CIMIC). *Canadian Army Journal*, 8(4): 33-49.
- Marcella, Lieutenant-Colonel Tim D. 2006. Force Protection for Logistics: Lessons and Recommendations for the Asymmetric Environment. *Canadian Army Journal*, 9(2): 60-69.
- NATO Standardization Agency. 2003. AJP-9: NATO Civil-Military Cooperation (CIMIC) Doctrine. *NATO*. www.nato.int/ims/docu/AJP-9.pdf (accessed March 20, 2007).
- Olson, Lara. 2006. Fighting for Humanitarian Space: NGOs in Afghanistan. *Journal of Military and Strategic Studies*, 9(1): http://www.jmss.org/2006/2006fall/articles/olson_ngo-afghanistan.pdf (accessed November 15, 2006).
- Peabody, David. 2005. The Challenges of Doing Good Work: The Development of Canadian Forces CIMIC Capability and NGOs. *Conference of Defence Associations Institute*. October 2005. <http://www.cdacda.ca/symposia/2005/Peabody.pdf> (accessed February 13, 2007).
- Peace Missions Programme. 2000. Peace Support Operations: A Working Draft Manual for African Military Practitioners. *Institute for Security Studies*. <http://www.issafrica.org/Pubs/Other/PeaceSupportManualMM/Ch2.pdf> (accessed May 22, 2007).
- Prest, Stewart, John Gazo, and David Carment. 2005. Working out strategies for strengthening fragile states—the British, American and German experience. *Failedstates.org*. <http://www.failedstates.org/documents/UK-US-German.pdf> (accessed January 13, 2007).
- Pronck, Rochus. 2007. Address made to the Peace Operations Working Group, Ottawa, March 29.
- Savage, Owen Alfred James McMenemy. 2006. *Peacebuilding in Afghanistan: The Case of the Canadian Provincial Reconstruction Team 2005/06*. Royal Roads University, master's dissertation.
- Sedra, Mark. 2003. Afghanistan: Between War and Reconstruction – Where Do We Go From Here? *Foreign Policy In Focus*. <http://www.fpif.org/pdf/papers/SRafghan.pdf> (accessed November 26, 2006).
- Senlis Council. 2006. Hearts and Minds in Southern Afghanistan. *SenlisCouncil*. http://www.senliscouncil.net/documents/HM_c1 (accessed January 29, 2007).
- Skidd, Lieutenant Alden. 2005. The Evolving Role of Civil-Military Cooperation (CIMIC) Operations in International Settings. *The Bulletin*, 11(1): 6-8.
- Smith, Gordon. 2007. Canada in Afghanistan: Is it Working? *Canadian Defence & Foreign Affairs Institute*. <http://www.cdfai.org/PDF/Canada%20in%20Afghanistan%20Is%20it%20Working.pdf> (accessed March 12, 2007).
- Standing Senate Committee on National Security and Defence. 2007. Canadian Troops in Afghanistan: Taking a Hard Look at a Hard Mission. *Parliament of Canada*. <http://www.parl.gc.ca/39/1/parlbus/commbus/senate/com-e/defe-e/rep-e/repfeb07-e.pdf> (accessed February 27, 2007).

Tondreau, Master Warrant Officer Denis. 2007. Interview by author.

Telephone. Ottawa, March 28.

Zaalberg, Thijs W. Brocades. 2006. Countering Insurgent-Terrorism:

Why NATO Chose the Wrong Historical Foundation for CIMIC. *Small Wars and Insurgencies*, 17(4): 399-420.

THE CHANGING NATURE OF “INTERNATIONAL SECURITY”: THE NEED FOR AN INTEGRATED DEFINITION

C. Elisabeth St. Jean

Norman Paterson School of International Affairs, Carleton University

Historically, the concept of international security has been equated with the use of force between nations, with a particular focus on great power warfare. However, during and since the 1980s, this description became increasingly questioned in terms of who should be secured, the nature of international threats, and the type of responses that were subsequently warranted to manage such threats. The new formulations of international security seemingly rectified the problems raised by the narrow conception of international security. However, as this article will demonstrate, these new security definitions are not without their own difficulties. This article will contrast the benefits and drawbacks of both the traditional and new definitions in order to illustrate that neither provides a sufficient conception of international security on its own. Instead, the article will put forward an integrated definition that incorporates the benefits, and rejects the drawbacks, of both. The article maintains that expanded notions are important but not at the expense of conventional understandings of international security. The article will also establish that the term “international security” should include security challenges that could threaten the international system, including non-military threats that may destabilize a state to the point of implosion.

INTRODUCTION

In international relations, words play an important role in determining how academics and policy makers view the world and make decisions on issues of great concern. The term “international security” is perhaps the most salient phrase in demonstrating the importance of semantics, as it is imbued with a sense of urgency and significance. As Adrian Hyde-Price points out, securitizing an issue means removing it from the regular political discourse and “signal[ling] a need for it to be addressed urgently and with exceptional means” (Hyde-Price 2001, 38). For much of contemporary history, and certainly since World War II, the concept and study of international security has been equated with the use of force between nations, with a particular focus on the role of great powers. This reflected the view that international security involved the territorial integrity of nations and the greatest threat to such territorial integrity was posed by wars between states, and particularly between great powers (Nye and Lynn-Jones 1998).

During and since the 1980s, this description became increasingly questioned in terms of who or what should be secured, the nature of international threats, and the type of responses that were subsequently warranted to control these threats. New conceptions of international security arose to incorporate, *inter alia*, different actors (such as human security), different forms of threats (such as environmental security), and different responses (such as non-military collective action). Analysts, activists, and policy makers promoted these new definitions because of the perceived shortcomings of traditional notions of

international security. The new formulations of international security seemingly rectified the problems raised by the narrow conception of the traditional definition. However, as this article will demonstrate, these new security definitions are still plagued by their own difficulties and challenges. As a result, there has been somewhat of a vindication of the traditional notion that international security should be primarily concerned with violence towards states, as it demonstrates merit when contrasted against the problems of the new forms of international security.

This article will contrast the benefits and drawbacks of both the traditional and modern definitions in order to illustrate that neither provides a sufficient conception of international security. Rather, the article will put forward an integrated definition that incorporates the benefits and rejects the drawbacks of both traditional and modern conceptions of international security. The analysis will begin with an examination of the fundamentals of the phrase “international security.” The article will then describe how the new definitions arose in response to the problems posed by the old conception. Following this, there will be a discussion of the subsequent problems with the new definitions and the corresponding benefits that traditional notions of international security bring to bear. The article will conclude by arguing that although these new understandings of international security present advantages, they raise issues that are not always equivalent to international security. It is necessary to conceptualize “international security” in a manner that neither subsumes all environmental or human problems under an international security rubric, nor limits international security to warfare alone. Instead, as the article will argue, an integrated definition focuses on the *impact* of threats rather than the nature or the source of the threat itself.

FUNDAMENTALS OF THE PHRASE “INTERNATIONAL SECURITY”

Security is an elusive subject for study. Adrian Hyde-Price (2001) describes how some academics argue it cannot be defined in any “objective” way, and that any problem can become a security issue once it has been securitized by policymakers. Security, then, manifests itself tautologically: any problem that is labelled security is in fact a security concern. However, as Hyde-Price (2001) then points out, this makes the security field entirely reactive to what policy makers deem a security threat, removing any independent analytical value. Such definitions of international security cannot, therefore, help to guide or inform policy, and although it may be of theoretical interest, this article will instead focus on the more objective definitions of security that can be used for academic and policy analysis.

In his article “Human Security: Paradigm Shift or Hot Air?” Roland Paris provides a basic but nevertheless useful definition: “a ‘security threat’ connotes some type of menace to survival” (2001, 98). The dilemma lies in interpretation, as there are three major aspects to the definition. First, there is a “menace to survival;” security is about threats and even threat perception. Second, security involves a referent object or unit of analysis, in that the “menace” poses a threat *to* someone (or something) and is also posed *by* someone (or something). For instance, an attack by one state against another is a classic international security

threat. Finally, discussions about security often include the means to secure the referent object from the threat, so a third area of debate is over the best response to a security threat.

The expansion of “international security” is consequently characterized by a shift in thinking with respect to the referent object, the threat to security, or the means to provide security. This shift is often achieved by adding adjectives to the term “security.” For instance, and as will be explained below, environmental security shifts focus from military to environmental threats; human security shifts focus from the state to individuals as the referent object; collective security shifts focus from unilateral to cooperative responses. Traditional definitions can thus be viewed as the foundation from which modernists expanded the concept of international security. The first two areas of debate (threat and referent object) form the crux of much of the international security debate, although the third aspect (response) also has a role.

The first debate focuses on the “threat” itself. Proponents of new conceptions of security maintain that the security definition must be broadened to incorporate new threats – environmental degradation, for example – that were previously relegated to other fields for analysis. Richard H. Ullman provides one broad definition, stating that a threat is an “action or sequence of events that... threatens drastically and over a relatively brief span of time to degrade the quality of life for the inhabitants of a state” (1983, 133). Mohammed Ayoob relates security to vulnerability and threats, maintaining that there exists a continuum of “vulnerability,” and problems become vulnerabilities when they “threaten to have political outcomes that affect the survivability of states” (1991, 259). Human security likewise moves the discourse beyond traditional threats facing the state towards human or individual-centric threats. Although the exact definition of human security is contested, it was born out of a 1994 United Nations Development Program report and contained seven security elements: economic, food, health, environmental, physical harm, community, and political (Paris 2001, 89-90). Traditionalists disagree with these broad notions of threats, arguing that military force is the primary threat and other issues, such as the environment and poverty, should only be considered as potential secondary causes of insecurity but not an international security issue per se.

In addition to disagreement over what constitutes a proper threat for inclusion in the definition of international security, referent objects are often vaguely described and are thus the focus of the second debate. While individuals, societal groups, and states all appear as focal referent objects in contemporary writings on international security, traditional notions of security, though they come in various guises, can be generally understood as the “military defence of state interests and territory” (Paris 2001, 87). It became the norm to view the state as the primary unit of analysis, and as a result, the notion of protecting the territorial integrity of the state became the end in and of itself. Thus, as Nicholas Thomas and William T. Tow (2002) point out, the state is the primary focus of analysis and action; a state faces a threat from another state, and it is the state that primarily responds. Yet the purpose of state security is, at its basic level, intended to protect the *people* within that state. Alternatively, new conceptions of security – human security in particular – have considered the

individual to be the unit of analysis. The consequence is that there is no agreement over what constitutes the proper referent object for international security.

Beyond the referent object and the threat, there is also a third disagreement, which exists over the proper response to any given threat. According to Hyde-Price, in terms of responses, “security has two dimensions: avoiding war (its negative dimension) and building peace (its positive dimension)” (2001, 28). In essence, when the referent object can reduce its vulnerability to a threat, its security is thereby increased. This can be achieved in two ways. First, the object can concentrate on the negative dimension by eliminating the threat directly, through political, economic, military, or other means. The second method focuses on building the positive dimension, where the object reduces its vulnerability to a threat by increasing its capacity to deter or prevent a threat from posing a direct risk. Deciding which to pursue is in some ways tied to the threat under consideration. In his review article, “The Security Problematic of the Third World” Mohammed Ayoob (1991) describes how traditionalists have placed the emphasis largely on using military capacity to reduce vulnerability, whereas many advocates of new formulations of security instead focus on non-military responses. In 2003, for instance, there was a debate among academics and policymakers whether invasion or diplomatic and other pressures was the best response to the potential threat of weapons of mass destruction in Iraq. Furthermore, new security definitions often promote a cooperative response to threats. And as J. Ann Tickner (1994) argues, there are some threats to the global system that cannot be solved by territorial protection.

THE SHIFT FROM TRADITIONAL TO MODERN FORMS OF “INTERNATIONAL SECURITY”

Because traditional notions of security focused on the use of force between great powers, the focus of international security studied during the Cold War was naturally on superpower conflict and nuclear war. With the end of the Cold War, analysts began to argue that the subject of international security “had to be recast to reflect the changing nature of conflict” (Freedman 1998, 48). As Mats Berdal (2000) notes, internal conflicts came to dominate the security agenda for two reasons. The main reason is simply because of the increase in incidence of internal conflict; few conflicts today are international in the sense that all the combatants are state actors. However, he also points out that the international community is becoming more involved in intrastate conflict. Indeed, as evidenced by interventions in Somalia and *ex post facto* acceptance of the NATO intervention in Kosovo, the Security Council itself has adopted a broader understanding of the international peace and security agenda.

Simultaneously, a view arose that “the new security agenda is increasingly composed of more intangible and diffuse risks and challenges” (Hyde-Price 2001, 27). With this shift in focus, there has been a concurrent shift towards analysing the social conditions that cause these new conflicts. Lawrence Freedman (1998) expands on this idea, pointing out that this shift towards the analysis of root causes of conflict is in fact sensible even to traditionalists, because there will always be a wider context to the use of force. Joseph S. Nye

Jr. and Sean M. Lynn-Jones (1988) describe how interdisciplinary approaches are a key aspect of international security studies. However, traditionalists limit psychological, economic, sociological, and other fields to analysing *aspects* of the traditional threat: war. Thus, for traditionalists, economics is only important insofar as it affects the likelihood of war, and typically that between great powers.

Analysts and advocates further appealed for the expansion of international security to consider the large number of individuals trapped in suffering. This occurred because “the end of the Cold War generated a major re-evaluation of normative and policy assumptions... [of] what made *people* ‘secure’”¹ (Thomas and Tow 2002, 177). This appeal occurred largely because with the end of the Cold War, there seemed to be space in the academic as well as policy arena to consider non-military problems facing the world. The result was that issues such as access to food, a clean environment, and economic welfare increasingly became issues of concern for international security studies. These matters first became important as state security threats in their own right, as analysts pointed out that these threats could exacerbate existing tensions and thus spark conflict. Further, there was a growing sense of a global consciousness that the international community – and the West in particular – was morally compelled to assist those individuals suffering in other nations.

WHEREFORE TRANSFORMATION? OUT WITH THE OLD, IN WITH THE NEW

The recent shift towards new ways of thinking about security arose because of the problems associated with the traditional concept of security. There are four problems with the old notion of security and five merits to the new definition’s response to these problems.

The first problem with the old definition is that in focusing on the state as the unit of analysis, it does not allow for an analysis of threats posed by the state itself to the people within that state. The problem with this is that citizens are often directly threatened by the state in which they reside. Thus, the first benefit of the broader definition is that human security allows for an analysis of the harm that a state can do to its own citizens. Moreover, Barry Blechman (1998) highlights the fact that although protection of the state is important for human welfare, it is not a sufficient guarantor of individual well-being. Indeed, as Tickner (1994) discusses in the context of the Cold War, traditional notions of security were at odds with the insecurity of those citizens who, firstly, lived with the threat of nuclear destruction, and secondly felt adverse impact from the amount of resources expended on nuclear weaponry. By focusing on the individual level, human security thus incorporates territorial protection while likewise paying heed to other threats facing people. Human security thus points out that even though states may be “secure,” the individuals that reside within them are not always similarly secure.

The second problem with conventional understandings of international security is the focus on the external nature of a threat. Because they use states as their unit of analysis, traditionalists look at the threat constituted by other states. They thus fail to consider insecurity that emanates from within a state, such as

that caused by political repression. Moreover, although traditionalists point to external threats from other states, most of the world is not immediately threatened by such concerns. The rest of the world is instead more concerned with internal conflicts and threats to society such as “indiscriminate violence, illegal immigration, drug dealing, and organized crime” (Gasteyer 2000, 90). These have been, however, largely considered domestic issues by the traditionalists. As a second merit, then, the new definitions of security endorse considerations of all threats, whether they originate from within the threatened state’s borders, such as displaced persons or terrorism, or transcend borders themselves, such as environmental concerns. Indeed, much of the human security analysis has been focused on the detrimental role that human insecurity can have in states neighbouring the host of the problem (Thomas and Tow 2002). Similarly, the third merit to new definitions is that they are not limited to interstate violence. Paris (2001) points out that the notion of human security emerged from the criticism that the traditional notion of security was too narrow for contemporary considerations. As Tickner (1994) explains, in the developing world, many threats originate from within the state, and not externally from another state. Furthermore, proponents of environmental security argue that not only can its decline at times lead to conflict, but the more general impact is a “downward pull on economic performance and, therefore, on political stability” (Tuchman Mathews 1989, 166).

A third drawback to traditional definitions is its lack of focus on longer term or potential threats, such as HIV/AIDs and health security. New definitions, particularly human security, incorporate such threats, thus providing a fourth merit in that they allow for an appreciation of threats that do not immediately pose acute distress. Paris (2001) explains that although some threats, such as environmental ones, are only projected, advocates maintain that they require immediate attention in order to prevent them from becoming actual threats. Further, analysts such as Jessica Tuchman-Mathews (1989) and Laurie Garrett (2005) have demonstrated that non-military threats, such as environmental and health issues, can pose threats and produce enormous costs. Likewise important in terms of understanding the origins of future threats, particularly for Western states, is the fact that threats “are now more likely to emanate from some distant turbulence” (Freedman 1998, 57). The subsequent argument, then, is that it is important for Western states to pay more attention to sub-state problems in other countries as they could give rise to serious threats in the future. Directing attention to these non-military issues via a security label is thus important, given that often such considerations are ignored unless they present an immediate and apparent danger.

Traditional definitions retain a fourth problem in their considerations of proper responses, where they have tended to focus on enhancing unilateral military capabilities. This is most well known in terms of the policy of mutually assured destruction during the Cold War, whereby the United States and the Union of Soviet Socialist Republics strove to attain enough nuclear armaments to ensure that in the event of an attack, each would retain sufficient nuclear weapons to destroy the other. The result of an attack by one would therefore be total annihilation of both. Another merit to focusing on individuals and non-

military threats, then, is the fact that the analysis of a proper response usually includes alternatives. Indeed, as Hyde-Price (2001) points out, many have argued that given globalization and the “new” threats, states alone are no longer best able to deal with threats.

PROBLEMS WITH THE NEW DEFINITION

Despite the variety of benefits that these new definitions of security bring, they have also presented a novel set of challenges and problems. Indeed, despite the motivations behind an expansion of the phrase, analysts soon began to question where limitations would be drawn between those problems that belong under domestic policy, and those threats that require attention in terms of international security policy. This has resulted in the re-emergence of advocates for traditional notions of security who provide five critiques of modernist definitions, and thereby support a return to traditional thought.

The first accusation levelled at proponents of the new interpretations of international security focuses on human security, and the fact that the term is often vaguely defined. The lack of a clear definition is partly because human security proponents cannot agree as to whether all problems facing individuals should be included or whether there should be some cut off between “development” and “security,” neither of which is well-defined. The vague definition is due to the fact that the phrase has brought together activists of varying issues, and a narrowing of the definition would likely make it difficult for such a diverse coalition to function as a whole (Paris 2001, 87). However, the resulting definition can conceptually encompass virtually anything that can be construed as discomfiting to an individual, and “if human security means almost anything, then it effectively means nothing” (Paris 2001, 93). It is difficult for policymakers to assess the relative importance of each aspect of human security, and as a field of study, the concept loses analytical value once it has broadened to such a point of inclusiveness. Without “clear criteria for specifying what is, and what is not, a security problem... an expanded definition of security will lose its intellectual coherence” (Gärtner and Hyde-Price 2001, 5). Some proponents have thus taken on traditional notions that use a narrower concept, where violence is the key threat. A common consequence has thus been the narrowing of the human security field to encompass only violent threats to individuals (Mack 2005).

The narrowing of human security has in fact been taken one step further, whereby human security is defined to incorporate only those threats to individuals that result in threats that transcend borders (Thomas and Tow 2002, 179). For instance, a refugee flow that destabilizes a region, such as the Great Lakes region in the mid 1990s, would constitute a human security threat. However, this gave rise to the second critique of the new understanding of security, which originated as a complaint by some human security proponents, who claim that such a definition has returned to the focus on the state as the referent object. A third problem with the new understanding of security is that the majority of the new threats are difficult to measure in terms of their actual impact. Hyde-Price (2001) emphasizes how such problems are largely potential threats rather than actual threats. Both environment and health threats constitute

prime examples, as it is difficult to prove that they can be the exclusive, or even primary, cause of a conflict.

A fourth critique arose because the new versions of international security were often borne out of a belief that the world had become a relatively safer place, which allowed for the focus to shift towards non-military threats. Yet it is difficult to argue that military threats facing states have disappeared from the international scene. Freedman further points out that if analysts have too much of a focus on non-military threats facing non-state actors then there is the potential to develop complacency towards analysing the aggressive nature of states under anarchy. As aggressiveness and anarchy are still features of the current international order, conventional military threats are still relevant to security studies. This is particularly the case considering that although the external threat of territorial integrity may not be an imminent threat to the West, it continues to be a very real threat in many other parts of the world. Moreover, a relatively short amount of time has passed since the end of the Cold War, and given the recent war in Iraq and regional tensions, particularly in the Middle East, South Asia, and East Asia, it seems rather short-sighted to assume that external threats of international aggression have dissipated entirely.

Finally, as Thomas and Tow (2002) discuss, another problem with contemporary threats is that they have no originating enemy. In other words, the value to traditional conceptions of security was that the threat was pre-meditated and originating from a specific source. How can non-traditional threats, such as the environment, be neutralized without an “enemy?” This speaks to the larger difficulty that once such problems are characterized as threats, military solutions are immediately considered, given that for much of modern history security threats had a military aspect that required a military response. Consequently, when problems are “securitized,” policymakers will tend to reach for a military solution. Freedman (1998) further argues that this is the result of forcing non-military problems into an analytical framework that was constructed to deal with military threats.

CONCLUSION: AN INTEGRATED DEFINITION

The debate between the two concepts raises serious questions regarding how we conceptualize and define international security. One extreme offers the view that international security is threatened when any human being suffers; the other extreme is that international security is only affected by the inter-relationships between states. With the numerous benefits and drawbacks displayed by both traditional and modern definitions, there is a need for an integrated conception of international security. This article will now establish that the term “international security” should be focused on the impact on state stability rather than attempt to classify which threats and referent objects should be included.

Concentrating on the impact of threats recognizes that not all human, environmental, and other security problems are necessarily of concern to international security. To illustrate, few would argue that gang warfare in the United States or the low human development indicators in some Canadian Aboriginal communities constitute international security concerns. They do, however, represent human security concerns, as defined by proponents of the

narrow and broad schools respectively. We thus require a definition that encompasses some of the aspects of the new threats without losing sight of the importance of the security of nations.

Indeed, military concerns will clearly occupy an important space in the international security discourse so long as anarchy prevails, even though one might argue, as Hyde-Price (2001) does, that we are witnessing limited anarchy through self-imposed measures, such as international laws and norms. Heinz Gärtner and Adrian Hyde-Price (2001) go further in their introduction to *Europe's New Security Challenges*, maintaining that there is conceptual space for considering non-military threats to the international order. Yet as Hyde-Price states elsewhere, international security studies “cannot and should not encompass all human distress; as an analytical approach to conflict, it must only deal with such issues when they threaten to provoke conflict and insecurity” (2001, 28). That is, although problems such as poverty should be considered as a potential cause of conflict, poverty in and of itself should not be considered a security threat.² Yet despite the traditional view of security, threats to international security are brought about not only by military threats. While traditional conceptions have focused on inter-relationships between states, the international system is likewise affected by unstable states, and particularly state implosion.

State instability and implosions are brought about by internal issues, which can be traced back to non-military causes, even if the destruction is through military means. This is best understood by considering the concept of structural violence: the indirect violence facing individuals within a state. Structural violence reduces life expectancy because of “deeply embedded socioeconomic inequalities... [resulting in a] lack of access to basic material needs” (Hyde-Price 2001, 33; Tickner 1994, 187). Thus, expanded notions of security, such as human security, are important inclusions in international security definitions because they are directly relevant to structural violence, which in turn can affect a state’s stability. Yet, international security is only affected by such problems when the threat to human security is of such magnitude that it endangers the state in question. To return to the example mentioned above, current human security problems within the United States and Canada are not considered international security issues because they do not threaten the stability of either state. Warring factions and low human development indicators in Iraq, however, constitute a direct threat to the Iraqi state and are thus threats to international security. The critical aspect which must be included in an integrated definition is that the *impact* of a threat to state stability defines whether it is an international security issue.³ Thus, threats facing international security include direct violence between states, but also involve other threats, such as famine or disease, that bring about state instability.

A second characteristic of international security, when defined by impact, is that it can involve any referent object. Indeed, this article maintains that the concept of international security has moved beyond interstate events; even contemporary traditionalists have acknowledged that international security involves the protection of states from threats originating within their own or other states. Thus, the notion of human security certainly resonates with this

understanding of international security. Therefore, it may be useful to turn to a point made by Diego Rojas Coronel: “the idea of an *international order* as a specific, protected object” (2004, 9). Taking this one step further, we can use Barry Buzan’s notion of “units of security,” whereby states, individuals and societal groups can each be a referent object, so long as their actions – regardless of whether they are caused or affected by the threat – have an impact on the international system (Coronel 2004, 3). This concept complements the above definition, as it also focuses on consequences, rather than attempting to pre-identify the referent objects in international security.

Finally, turning to the nature of response, the traditional definition tends to focus on unilateral and often military reactions to security threats. This, however, is a function of interpretation and not the definition itself. According to the United Nations Charter, the Security Council is afforded the right to interfere in any issue it deems a threat to international peace and security. The international community decides to impose its will in order to mitigate a rising problem when it views that problem as an international security issue, regardless of whether it originated internally or externally. Moreover, individuals are frequently vulnerable to events and processes at the international level. This is compounded by the fact that states are often unable or unwilling to address vulnerabilities at the sub-state and individual level, and therefore require assistance from the international community. Focusing specifically on the impact of a threat, therefore, allows for both international and unilateral responses, depending on the nature of the threat’s actual or potential impact.

Human security and other⁴ expanded notions of security thus seem to be better considered, as Paris (2001) argued, as labels for categories of research, not frameworks of analysis. In order to broaden the nature of international security studies without expanding to the point of inadequacy, it is important to recognize that this field is concerned first and foremost with organized violence that affects the international system. There are times when international security will be primarily concerned with conventional military threats; however, there will also be occasions when the conditions for peace, insofar as they affect state stability, will be of primary salience. Hyde-Price asserts, “security involves preventing war through military preparations to deter armed aggression from within and without and, more positively, fostering conditions conducive to building a legitimate and enduring peace order” (2001, 40). This can be demonstrated by juxtaposing two questions. First, does it matter if immediate inter- or intra-state violence is prevented but we fail to avert a global environmental catastrophe that destroys the planet? Conversely, what does soil erosion matter if the earth has been destroyed in a nuclear holocaust? While extreme, these two examples demonstrate that neither military nor non-military threats are unimportant, and one cannot be forgotten at the expense of the other.

An integrated definition should therefore include those challenges that could threaten the international system, either through direct violence between states or by means of state instability, particularly state implosion. This definition recognizes that international security can be affected by non-military threats as well as state and/or non-state groups. Further, an integrated definition acknowledges that although unilateral responses may be in order at times, many

issues affecting current international security involve – and indeed require – an international response. Thus, in this integrated definition, threats are defined by their *impact* (international violence or state instability) rather than their type (such as environmental) or origin (such as refugee flow). This provides a starting point for an integrated analysis of international security that allows for the inclusion of any threat, referent object, or response, so long as it affects the international system or involves the international community.

NOTES

¹ Emphasis added.

² Here it is also important to note that principles will often come into conflict with security; for instance, as Freedman explains, a state will often find itself choosing between security and civil liberties. This was evident in the United States and other Western states following the September 11, 2001 terrorist attacks, as measures were implemented to better discover and contain terrorist groups.

³ Richard Ullman’s (1983) definition provided a starting point for looking at the impact of security related concerns, arguing that we must look at the impact on the people within the state, rather than simply the physical integrity of the state’s borders. Yet his definition still focused on national, or internal, security, and he does not explain why such concerns should be considered *international* security problems. For instance, he makes a compelling case for including “decimating epidemics, catastrophic floods, or massive and pervasive droughts” as security problems (Ullman 1983, 133), but he stops short of making a case that these problems should likewise be included in a definition of international security.

REFERENCES

- Ayoob, Mohammed. 1991. The Security Problematic of the Third World. *World Politics* 43: 257-83.
- Berdal, Mats. 2000. International Security After the Cold War. In *Towards the 21st Century*, eds Kurt R. Spillmann and Andreas Wenger, 21-68. Bern: Peter Lang.
- Blechman, Barry M. 1998. International Peace and Security in the Twenty-first Century. In *Statecraft and Security: the Cold War and Beyond*, ed. Ken Booth, 289-307. Cambridge: Cambridge University Press.
- Coronel, Diego Rojas. 2004. International order as a referent object. *Standing Group on International Relations Fifth Pan-European International Relations conference*. <http://www.sgir.org/conference2004/papers/> (accessed 15 March 2006).
- Freedman, Lawrence. 1998. International Security: Changing Targets. *Foreign Policy* 110: 48-63.
- Garrett, Laurie. 2005. The Lessons of HIV/AIDS. *Foreign Affairs* 84: 51-64.
- . 2005. The Next Pandemic? *Foreign Affairs* 84: 3-23.
- Gärtner, Heinz and Hyde-Price, Adrian. 2001. Introduction. In *Europe's New Security Challenges*, eds. Heinz Gärtner, Adrian Hyde-Price, and Erich Reiter, 1-26. Boulder: Lynne Rienner Pub.
- Gasteyer, Curt. 2000. Old and New Dimensions of International Security. In *Towards the 21st Century*, eds Kurt R. Spillmann and Andreas Wenger, 69-108. Bern, Peter Lang.
- Hyde-Price, Adrian. 2001. ‘Beware the Jabberwock!’ Security Studies in the Twenty-First Century. In *Europe's New Security Challenges*, eds. Heinz Gärtner, Adrian Hyde-Price, and Erich Reiter, 27-54. Boulder: Lynne Rienner Pub.
- Levy, Mark. 1995. Is the Environment a Security Issue? *International Security* 20: 35-62.
- Mack, Andrew. 2005. *Human Security Report 2005: War and Peace in the 21st Century*. Vancouver: The Human Security Centre.
- Messari, Nizar. 2002. The State and Dilemmas of Security: The Middle East and the Balkans. *Security Dialogue* 33: 415-19.
- Newman, Edward. 2004. A Normatively Attractive but Analytically Weak Concept. *Security Dialogue* 35: 358-9.
- Nye, Joseph S. Jr. and Lynn-Jones, Sean M. 1988. International Security Studies: A Report of a Conference on the State of the Field. *International Security* 12: 5-27.

- Owen, Taylor. 2004. Human Security – Conflict, Critique and Consensus: Colloquium Remarks and a Proposal for a Threshold-Based Definition. *Security Dialogue* 35: 373-87.
- Paris, Roland. 2001. Human Security: Paradigm Shift or Hot Air? *International Security* 26: 87-102.
- Thomas, Nicholas and Tow, William T. 2002. The Utility of Human Security: Sovereignty and Humanitarian Intervention. *Security Dialogue* 33: 177-92.
- Tickner, J. Ann. 1994. Re-visioning Security. In *International Relations Theory Today*, eds. Ken Booth and Steve Smith, 175-197. University Park, Pa: Pennsylvania State University Press.
- Tuchman Mathews, Jessica. 1989. Redefining Security. *Foreign Affairs* 163-77.
- Ullman, Richard H. 1983. Redefining Security. *International Security* 8: 129-53.

TURNING THE HEALTH PROFESSIONAL CAROUSEL: IS CANADA UNDERMINING HUMAN RIGHTS IN DEVELOPING COUNTRIES?

Karinne Lantz

*Norman Paterson School of International Affairs, Carleton University
Faculty of Law (Common Law Section), University of Ottawa*

This article will address the issue of health professional migration, with a specific focus on how this migration affects health systems in developing countries. The central question being examined is whether or not states have an obligation to ensure that their policies – or actions by private actors based in their states – do not undermine the delivery of healthcare in other states. After exploring this obligation, this article will analyze how the issue may be effectively addressed by drawing upon the experience of the United Kingdom; how successful has the U.K. been in meeting its obligation? What are the most effective policy responses for developed states to implement? By framing this problem as a human rights issue, it will be argued that developed countries have a moral and legal responsibility to mitigate the negative effects of active recruitment of health professionals from developing countries. In light the U.K.'s experience, Canada's potential role in undermining human rights in developing countries will be examined and policy recommendations will be made. Although the issue is complex, this policy analysis will centre on the unmet demand for health professionals as a primary driver of international migration, as it is argued that this factor may be most effectively addressed by developed states.

INTRODUCTION: INTERNATIONAL MIGRATION OF HEALTH PROFESSIONALS

Individuals have moved from one country to another for centuries; however, globalization has made migration easier, resulting in “a new dynamic in population movements in terms of size and velocity” (Bach 2003, 2). Even in the short term, the surge in migration is evident: it has been estimated that in 2000, 2.9 percent of the world's population – almost 175 million people – were living a country other than that of their birth, compared to one hundred million in 1995 (Stilwell et al 2003, 5). This increased mobility has facilitated the movement of health professionals, who constitute the largest group of skilled migrants (Martineau, Decker and Bundred 2002, 1). An estimated thirty-five million people work in health services worldwide; however, international migration has resulted in an inequitable sharing of “this essential global resource” (Martineau, Decker and Bundred 2004, 1).

The loss of health professionals presents a challenge to maintaining effective health systems in both developing and developed countries because the provision of adequate healthcare services depends on trained professionals. Although long recognized, the challenges arising from the migration of health professionals have not been comprehensively addressed. At the 1965 Edinburgh

Commonwealth Medical Conference, concerns about the effects of health professional migration were expressed, prompting a thorough study of the issue by the World Health Organization (WHO) (Ibid.). Released in 1979, the Mejia study reported that in 1972, 140,000 (6 percent) of the world's physicians were practicing in a country other than that of their nationality, while "86 percent of all migrant physicians were working in five countries: Australia, Canada, the Federal Republic of Germany, the United Kingdom...and the United States" (Bach 2003, 3). Despite the attention paid to the issue, the migration of health professionals increased and the Mejia study remained the only one of its kind until the WHO's *World Health Report 2006*, which focused on the global health workforce (see WHO 2006). It is widely recognized that the character of health professional migration has changed, as the number of migrants has increased significantly and the relocation is now "often permanent" (Eastwood et al. 2005, 1893). The nationalities of migrants have also changed, with new major source countries located in the Caribbean, Egypt, sub-Saharan Africa (SSA), Cuba and the former Soviet Union (Martineau, Decker and Bundred 2004, 2). The increased rate and shifting pattern of migration endanger the sustainability of health systems in many developing countries. Given the fragile nature of many of their health systems, the countries in SSA are most threatened by migration (Bach 2003, 5).

This article will address the issue of health professional migration, with a specific focus on how this migration affects health systems in developing countries. The central question being examined is whether or not states have an obligation to ensure that their policies (and actions by private actors based in their states) do not undermine the delivery of healthcare in other states. After exploring this obligation, this article will analyze how the issue may be effectively addressed by drawing upon the experience of the United Kingdom; how successful has the United Kingdom been in meeting its obligation? What are the most effective policy responses for developed states to implement? By framing this problem as a human rights issue, it is argued that developed countries have a moral and legal responsibility to mitigate the negative effects of active recruitment of health professionals from developing countries. In light the U.K.'s experience, Canada's potential role in undermining human rights in developing countries will be examined and policy recommendations will be made. Although focusing on the pull factors in developed countries "distorts a more complex picture" (Bach 2003, 7), this policy analysis will centre on the unmet demand for health professionals as a primary driver of international migration, as this factor may be most effectively addressed by developed states.

THE INTERNATIONAL HEALTH PROFESSIONAL CAROUSEL

Myriad factors influence the migration of health professionals, including individual choices, government policies and recruitment activities by governmental and private agencies. Complex "push" and "pull" factors influence individual choices to migrate, as conditions in the source state can encourage (push) individuals to leave, while conditions in receiving states attract (pull) individuals to those states. Globalization has facilitated this process, by making individuals increasingly aware of more favourable conditions in other countries,

which encourages migration: “the internet, easier international communication generally, and the rise of commercial agencies organizing health care staff migration have all greatly increased access to information about means, costs and consequences of migration” (Mensah 2005, 205-6).

The WHO estimates that the worldwide shortage of health professionals is 4.3 million workers (WHO 2006, 12). This unmet demand for health professionals is a major contributing factor to migration, as many developed countries including Canada, the United States, the United Kingdom and Australia have a shortage of health professionals and rely on immigration to maintain their healthcare systems. For example, the health plan for the United Kingdom released in 2000 required 10,000 more doctors and 20,000 more nurses to meet its needs and looked internationally for some of these professionals (Martineau, Decker and Bundred 2002, 3). The migration of health professionals has been analogized as an international “carousel” that moves health professionals from one country to another: a doctor from Tanzania, where there are two to three doctors per 100,000 people, may migrate to South Africa¹ to fill a vacancy left by a doctor who emigrated to the United Kingdom, which in turn was a result of the emigration of a British doctor to fill a vacancy in Canada that resulted from the loss of a Canadian doctor to the United States (Eastwood et al 2005, 1893). In recent years, there has been “a greater polarization [in migration] towards flows from poorer to richer countries,” (Martineau, Decker and Bundred 2002, 7) and due to the more permanent nature of migration from developing countries, “the medical carousel unfortunately does not turn full circle, as it has in the past, so the poorest nations experience all drain but no gain” (Eastwood et al 2005, 1894).

HUMAN RIGHTS: THE RIGHT TO HEALTH

Reliance on the migration of health professionals to support domestic health systems results in “a serious ethical issue to be confronted,” because it is clearly “inappropriate for nations as wealthy as Canada to solve [their] own domestic health human resources problems of undersupply and maldistribution by relying on the immigration of health professionals from developing countries” (McIntosh, Torgerson and Klassan 2007, 4). Although this normative argument is compelling, states relying on health professional migration may also incur a legal obligation to minimize the adverse effects of migration on developing countries in light of the human right to health that is recognized in the WHO Constitution. This right to health is further entrenched in international conventions, as “every country in the world is now party to at least one human rights treaty that addresses health-related rights, including the right to health” (WHO, 2007). In recognizing the right, “governments have an obligation to ensure that functioning public health and health-care facilities, goods and services, as well as programmes, are available in sufficient quantity to the population” (WHO 2003, 11). As a party to such conventions, Canada is bound by the obligations arising from these treaty regimes and by its membership in the WHO.

First reflected in the WHO Constitution of 1946, the right to health is implicit within the 1948 UN Universal Declaration of Human Rights in Article

25, which states that, “everyone has the right to a standard of living adequate for the health and wellbeing of himself and his family” (UN 1948, Art. 25). The right was given further definition within the International Convention on Economic, Social and Cultural Rights (ICESCR), which states that all parties to the convention “recognize the right of everyone to the highest attainable standard of physical and mental health” (UN 1976, Art. 12(1)), and obliges states to take steps toward “the creation of conditions which would assure to all medical service and medical attention in the event of sickness” (Ibid., Art. 12(2)). The Committee on Economic, Social and Cultural Rights has adopted numerous General Comments clarifying the nature and content of the treaty and the obligations of states parties to the ICESCR (WHO 2002, 10). Adopted in May 2000, General Comment No. 14 identified the four essential elements of the right to health as availability, accessibility, acceptability and quality (Committee on Economic, Social and Cultural Rights 2000, Art. 12), while reiterating that, like all human rights, the right to health imposes the obligations to “respect”, “protect” and “fulfil” [*sic*] on states parties (Ibid., Art. 33). Of particular importance, “the obligation to respect requires States to refrain from interfering directly or indirectly with the enjoyment of the right to health” (Ibid.). International co-operation, originally noted in General Comment No. 3, is highlighted once again within General Comment No. 14, as “States parties should recognize the essential role of international cooperation and comply with their commitment to take joint and separate action to achieve the full realization of the right to health” (Ibid., Art. 38). Furthermore, states that have the means to do so must “respect the enjoyment of the right to health in other countries, and to prevent third parties from violating the right in other countries” in order to comply with their obligations (Ibid., Art. 39). While exploring the customary status of the right to health is beyond the scope of this article, it is evident that states – including Canada – as parties to the WHO and the ICESCR, have the treaty-based obligation to recognize the right to health.

Given the widespread recognition of the right to health, this framework is useful in determining the legal obligation that may arise from the active recruitment of health professionals from developing countries experiencing a shortage of trained medical staff. Clearly, a normative dimension surrounds the recruitment of health professionals from other countries, as one is forced to ask if it is moral for a state to pursue policies or permit behaviours that undermine the health systems of other states. Framing the activity as a legal issue underscores the importance of the issue and provides an impetus for states to address the practice. The result of the legal obligation to respect other states’ right to health is that state sovereignty cannot be used to shield domestic policies from scrutiny when they adversely affect other states, and “it is increasingly recognized that ‘recruiting’ countries should assess the impact of their policies on the fulfillment of human rights in other countries” (WHO 2003, 12).

The migration of health professionals clearly has the potential to undermine a state’s ability to provide health services. Poignant examples are Ethiopia, Nigeria and Uganda, as they have all “cited shortage of health professionals as the main constraint to mobilizing responses to health challenges” (Eastwood et al 2005, 1894). The costs of losing health

professionals are substantial and present a potentially insurmountable challenge to developing countries. Lacking resources to invest in health care systems, the loss of trained professionals may have dire consequences in developing states. The United Nations Conference on Trade and Development has “estimated that each migrating African professional represents a loss of US\$184,000 to Africa” (Ibid.), while the overall effect has been conceptualized as an annual subsidy of \$500 million by Africa to North America, Western Europe and Asia (Mensah 2005, 210). Beyond directly affecting the quality of health services, emigration of health professionals may also result in inequitable access when individuals move from rural areas to fill urban vacancies. In-country migration can concentrate the lack of health professionals in rural areas, and since the poorest citizens are often located in these areas, they feel the loss most acutely. Ghana provides a clear example of unequal access: in 2000, not a single pediatrician was practicing in the northern two-thirds of the country due to in-country migration (Eastwood et al 2005, 1895).

Foreign trained health professionals represent a potentially enormous savings to developed countries, while acting as a drain on the resources of developing states that lack the financial ability to provide equitable health services. New Zealand, for example, is estimated to have saved \$37 million by having six hundred South African trained doctors practicing in the country (Mensah 2005, 210), giving strength to the claim that, “by actively enticing medical graduates from the less-developed world...health planners are looking for an inexpensive solution to inadequate human resource planning” (Bundred and Levitt 2000, 246). The obligation that the right to health imposes is progressive, such that a state must provide its population with the highest level of health services for which it has the means to provide. As such, one may argue that states that have the means to meet their own health needs but rely, as a matter of acknowledged or unacknowledged policy, on the migration of health professionals to support their health systems, should be held accountable for the adverse effects of their policies on other populations.

HEALTH, DEVELOPMENT AND INEQUALITY

Beyond the right to health, the ability to provide health services has wide-reaching effects. Good health is central to development, as healthy individuals are more likely to lead productive and fulfilling lives. Successful development depends on functioning health systems because poor health leads to lower life expectancy and increased suffering among the most vulnerable segments of society worldwide. As the World Bank has noted, improving health outcomes may be justified on economic grounds, because it fuels economic growth: “It reduces production losses caused by worker illness; it permits the use of natural resources that had been totally or nearly inaccessible because of disease; it increases the enrollment of children in school and makes them better able to learn; and it frees for alternative uses resources that would otherwise be spent on treating illness” (World Bank 1993, 17). Investing in health services may also be a central means of combating poverty, while “the adverse effects of ill health are greatest for poor people, mainly because they are ill more often, but partly because their income depends exclusively on physical labour” (Ibid., 20). Good

health is a “fundamental goal of development as well as a means of accelerating it;” therefore, “[t]argeting health as a part of development efforts is an effective way to improve welfare in low-income countries” (Ibid., 21). Sen, for example, views freedoms and capabilities as central to the enjoyment of some human rights (Sen 2005, 153-55). Through this framework, it is clear that health is necessary to achieve many human rights goals, as individuals suffering from the effects of poor health may be unable to meet or enjoy their other fundamental human rights due to a lack of capability.

Health professional migration has a disproportionate effect on the world’s poor. Since developed countries are often the destination of health professionals, the benefits of migration are concentrated there, while the negative effects of migration are largely felt by developing countries. The result is that “losses are compounded in developing countries that contend with a disproportionate burden of disease with far fewer resources than developed countries” (McIntosh, Torgerson and Klassan 2007, 1). Given the link between health and development, the migration of health professionals may clearly undermine the development prospects for many states. The WHO confirms this linkage, as it has reported that, “the correlation between the availability of health workers and coverage of health interventions suggests that the public’s health suffers when health workers are scarce” (WHO 2006, 10).

Just as human and financial resources are unequally divided worldwide, so too are health resources, as “countries with the lowest relative [health] need have the highest numbers of health workers, while those with the greatest burden of disease must make do with a much smaller health workforce” (Ibid., 8). Africa experiences this inequality acutely, as it “suffers more than 24 percent of the global burden of disease but has access to only 3 percent of health workers and less than 1 percent of the world’s financial resources – even with loans and grants from abroad” (Ibid.). The WHO has identified fifty-seven countries with a critical shortage of health workers and estimates that 2.4 million additional health professionals would be required for these countries to reach the target level of health professionals (Ibid., 12). Given that thirty-six countries in SSA have a critical shortage and “an increase of almost 140 percent is necessary to meet the threshold [number of health workers],” the vulnerability of the region is difficult to overstate (Ibid.).

Struggling with being the centre of the HIV/AIDS pandemic, the loss of health professionals has dire consequences for the health systems in SSA. Beyond its drain on health resources, HIV/AIDS compounds poor working conditions for health workers in affected countries, through the “loss of staff due to illness, absenteeism, low staff morale...and through the increased burden of [HIV/AIDS] patient load” (Eastwood 2005, 1895). The Director-General of the WHO has noted that, “the brain drain from Africa is severely limiting the ability of health workers to combat the HIV/AIDS epidemic and achieve any substantial progress towards the millennium development goals [MDGs]” (Mensah 2005, 203). Furthermore, “the dire shortage of health workers in many places is among the most significant constraints to achieving the three health-related [MDGs]: to reduce child mortality, improve maternal health, and combat HIV/AIDS and other diseases” (WHO 2006, 19). A 2005 study “estimated that an additional 1

million extra workers [would] be needed in [SSA] ...to meet the [MDGs] by 2015” (Eastwood 2005, 1895). If the brain drain continues, very few of these extra workers will be realized, while development and the right to health will remain severely challenged.

The ability of developing countries to respond to the migration of health professionals is questionable, as migration often surpasses their capacity to train replacements. In SSA, twenty-four out of forty-seven countries have only one medical school, while eleven have none (Ibid., 1893). Ghana is a prime example of a country experiencing the negative effects of health professional migration: it requires 520 nurses annually, but is only able to enroll 375 students each year due to fiscal constraints (Mensah 2005, 207). Compounding the health crisis, it is estimated that 60 percent of doctors trained in Ghana in the 1980s have left (Eastwood 2005, 1893), while 90.3 percent of doctors trained between 1985-94 migrated to the United States or the United Kingdom and 90.8 percent of those trained between 1998-2002 have emigrated (Mensah 2005, 205). A similar situation is occurring in Zambia. The WHO has estimated that Zambia requires 1,500 doctors to meet its medical needs, but there are only eight hundred doctors registered in the country and only fifty of the six hundred Zambian graduates in the history of the medical school in Lusaka remained working in the Zambian public-health sector in 2000 (Bundred and Levitt, 245). With five thousand doctors trained in SSA working in the United States in 2002 and 3,199 SSA-trained nurses registered in the United Kingdom in 2002-3 (Mensah 2005, 207), it is highly unlikely that countries in the region will be able to provide adequate and equitable health services to their populations if the current migration trends continue.

THE EFFECTS OF HEALTH PROFESSIONAL MIGRATION

The issue of health professional migration is extremely complex, as no two source nor receiving countries are the same, and the issue is further complicated by the fact that it is ultimately an individual decision to emigrate. The interplay of influences on individuals contemplating migration is a complex mix of push-and pull-factors. Diallo has identified the main pull-factors for migration to be salaries and benefits, better working environments and improved quality of life for the worker and family (Diallo 2004, 602). For health professionals, the opportunity for professional development is also a major draw: “for highly skilled workers, continuous professional development is an integral component of individual career planning and progression,” while the lack of opportunity for professional development at home “in contrast to the perceived opportunities abroad, reinforces the attractiveness of overseas employment” (Bach 2003, 11). This factor is confirmed by a 2002 survey of 1,119 foreign nurses working in the United Kingdom, in which pay and professional development were cited as the most preferred aspects of their work (Mensah 2005, 204). Major push-factors can include the working environment, violence inside and outside of the workplace, and the toll of HIV/AIDS. The structural adjustment programs promoted by international lending institutions during the 1970s and 1980s have also been identified as a potential push-factor in states where the reform programs led to a deterioration of living and working conditions (Ibid.).

Additionally, social networks can play a significant role in migration patterns, as migrants who communicate positive experiences to their colleagues and relatives at home encourage further migration and establish a potentially self-sustaining “migration pathway” between states (Bach 2003, 10).

The effect of migration is highly dependent on its nature – who is migrating, where they are coming from, where they are going and how long they will remain abroad. As mentioned above, the migration of health professionals represents a financial benefit to the receiving state. Although difficult to quantify, the benefit may be illustrated by the savings to the state by not having to train as many health professionals. It is estimated that every doctor trained in the United Kingdom costs £200,000 to £250,000 and takes five to six years to train; therefore, employing foreign-trained doctors is like importing “human capital at zero cost, with immediate effect” (Eastwood et al 2005, 1895). This loss of human capital and resources through migration is commonly referred to as ‘brain drain.’ The negative effects of such brain drain are further compounded when the individual migrant was involved – directly or indirectly – in training. Specific to the health sector, senior doctors are important because they act as a useful role model and source of information for new doctors (Ibid., 1894). When assessing the effects of migration, particular attention must be paid to *who* is migrating because the loss of an individual is difficult to quantify with merely the cost of training.

In contrast to brain drain, migration in highly skilled fields can also result in ‘brain circulation’ or ‘brain exchange’ through which temporary migration provides a long-term benefit for the source state because the migrant develops new skills through employment or educational experiences abroad that were not available in the source state (Bach 2003, 14). Given the unique circumstances within the health sector, brain drain is the more likely result of migration than brain exchange (Ibid., 12). Doctors emigrating from SSA also tend to migrate on a permanent basis therefore, the potential benefit of brain exchange to the region is limited (Eastwood et al 2005, 1894). Brain drain threatens the stability of the source state’s health system, as the remaining workers “experience added stress and greater workloads,” resulting in the staff being “ill-motivated, not only because of their workload, but also because they are poorly paid, poorly equipped, inadequately supervised and informed and have limited career opportunities” (Stilwell et al 2003, 5).

Even within developing countries, however, the effects of health professional migration may be quite different. One of the widely recognized benefits of migration is the potential for remittances, as money is sent back by those working abroad to support relatives remaining in the source country. Remittances are an important source of income for many developing countries, as it is “the second-largest source, behind FDI, of external funding for developing countries,” amounting to \$72.3 billion in 2001 (Ibid., 6). The Philippines is a clear example of one of the few states that is known to benefit from the migration of its health professionals. Ten percent of the population in the Philippines lives or works abroad. As early as 1970, more Filipino nurses were registered in the United States and Canada alone than the Philippines (Martineau, Decker and Bundred 2002, 2), while the policy of sending workers

overseas was recognized as being central to the country's economic growth in the *2001-2004 Medium Term Philippines Development Plan* (Bach 2003, 4). As the Secretary of Labour and Employment has stated, overseas employment is "an industry. It's not politically correct to say you are exporting people, but it's part of globalization, and...countries like [the Philippines], rich in human resources, have that to contribute to the rest of the world" (Ibid.) By training more health professionals than the country can employ, "remittances from physicians working abroad [are] estimated to compensate for the costs of training and emigration" (Diallo 2004, 602).

Although the Philippine experience shows that the international health carousel may be used to a state's benefit, the situation is much different in countries experiencing a health care crisis. Due to the fact that the funds "sent by health workers...are not directly reinvested in human capital for the health systems," there is a net loss of human capital within the healthcare system (Stilwell et al 2003, 6). As Eastwood *et al* have stated, "the idea that repatriated overseas earnings could make up for the deficiency [caused by the brain drain] is unrealistic, as there is no way of ensuring that repatriated income will find its way into investments in health care," a problem that is exacerbated when the state lacks professionals to champion the cause of investment in the health system (Eastwood 2005, 1894). Rather than hope for the potential indirect benefits of migration, "developing capable, motivated and supported health workers is essential for overcoming bottlenecks to achieve national and global health goals" (WHO 2006, xv).

RESPONDING TO THE CRISIS: INCOMPLETE INFORMATION AND CONFLICTING RIGHTS

A major challenge to addressing the health professional migration issue is incomplete information, as "the lack of reliable, up-to-date information greatly restricts the ability of policy-makers at national and international levels to develop evidence-based strategies to resolve the health workforce crisis, or to develop health systems to serve the needs of disadvantaged people" (Ibid., 15). Rather than being based on reliable evidence, most reports about the effects of migration are based on anecdotal information, or limited and inconsistent data (Stilwell et al 2003, 2). A major obstacle is the lack of international standards for measuring migration and the inability of many developing states to gather the information necessary to quantify the problem. Lacking concrete data on emigration from developing countries, studies largely rely on the health professional registers in developed states, which cannot provide all of the necessary information. This "reliance on incomplete data or incompatible data from different sources often means that it is not possible even to have an accurate picture of the trend in outflow of health workers, let alone any assessment of the impact of this outflow on the health services" (Buchan and Dovlo 2004, 38). To overcome the problem of incomplete information, the WHO, the International Labour Organization and other international organizations are working toward the harmonization of methods for collecting and analyzing data specific to health personnel migration (Diallo 2004, 607). Although the lack of data "makes it difficult to empirically assess the impact of

health worker migration on health systems” (Stilwell et al 2003, 8) and the effect of policies, it remains important for states to attempt to address the issue.

In order to minimize the negative effects of health professional migration, it is essential to develop ethical recruitment policies with the goal of successfully managing migration, rather than preventing its occurrence (Ibid., 2). Not only is prevention practically difficult, it raises ethical concerns, because there is an effective clash of individual rights at the heart of the process: the right to health and the right to mobility. Since “freedom of movement is a human right,” it is unrealistic to view an outright ban on all health professional migration from countries with a health worker crisis as an effective response to the issue (Ibid., 12). However, an ethical compromise may be found in limiting or preventing active recruitment of healthcare workers from states with fragile health systems. South Africa adopted this approach in 1995, when it banned the recruitment of doctors from other Organization of African Unity countries in order to reduce the number of doctors migrating from poor countries to fill vacancies in its rural regions (Bundred and Levitt 2000, 245). As *The Lancet* has stressed, “what must be remembered [by policy-makers] is that the objective is not limitation of mobility but equity of health care as soon as possible” (The Lancet 2000, 177).

THE UNITED KINGDOM EXPERIENCE

Like many other developed countries, the United Kingdom faces its own shortage of health professionals and relies on migration as a partial solution to its problem. The government “views overseas recruitment as an integral component of government policy to increase [the National Health Service] workforce and combat staff shortages” (Bach 2003, 12). In 2004, thirty-one percent of doctors and thirteen percent of nurses practicing in the United Kingdom were born in another country, clearly illustrating the state’s reliance on migration (Eastwood 2005, 1895). The health professional carousel also has a palpable effect in the United Kingdom, as many of its professionals are drawn to other countries like Ireland, Canada, Australia and the United States (Mensah 2005, 204). According to Mensah, the United Kingdom loses 15 percent of its doctors to foreign countries within three years of graduation, and 18 to 20 percent after ten to twenty years (Ibid.) Relying on migration to fill its vacancies, the United Kingdom acts as a drain on developing countries because “professional workers from developing countries are relatively unlikely to leave [the United Kingdom] after a few years of residence” (Bach 2003, 15). In response to this drain, Nelson Mandela criticized the United Kingdom in 1997 for actively recruiting South African nurses (Ibid., 21), providing a high-profile impetus for the government to act.

In 2001, the United Kingdom became the first country in the world to address its obligation to not undermine the health systems of other countries by introducing a code of practice for the recruitment of health professionals. Revised in 2004, the key aim of the Code of Practice for the International Recruitment of Healthcare Professionals is “to promote high standards of practice in the international recruitment of healthcare professionals,” which is “underpinned by the principle that any international recruitment...should not

prejudice the healthcare systems of developing countries” (U.K. Department of Health 2004, 4). The revised Code places a greater emphasis on ensuring that recruitment activities do not adversely affect developing countries’ healthcare systems by “preclud[ing] the active recruitment of healthcare professionals from developing countries, unless there exists a government-to-government agreement to support recruitment activities” (Ibid.). As a result of the revision, the Code applies to temporary as well as permanent staff, enables and encourages all healthcare organizations from both the public and “independent” sectors to sign on to its principles, and provides “best practice benchmarks” for recruitment activities (Ibid., 5). A central feature of the Code is that the Department of Health provides a list of developing countries “that should not be targeted [for recruitment] under any circumstances” (Ibid., 6). In recognition of the individual right to migrate, the Code only applies to active recruitment, so the National Health Service is not barred from employing individuals who applied for employment as a result of individual initiative without active solicitation.

A report for the U.K. Department for International Development “suggested that the [Code of 2001 had] very significant weaknesses and loopholes,” some of which were addressed in the 2004 revision (Eastwood 2005, 1894). The principle problem cited with the Code was its failure to cover private employers and recruitment agencies and the ease with which indirect means could circumvent the active recruitment prohibitions (Ibid., 1895). Not being directly covered, the private sector has continued active recruitment directly from countries on the proscribed list (Buchan and Dovlo 2004, 17). In 2003, one quarter of nurse registrants in the United Kingdom were from the states on the proscribed list, while work permits for many health professionals from developing countries on the list were approved: 5,880 from South Africa, 2,825 from Zimbabwe, 1,510 from Nigeria and 850 from Ghana (Eastwood 2005, 1893). Despite the changes, concerns remain, largely due to the fact that adherence to the Code by private institutions and recruitment agencies is voluntary. For many analysts, the United Kingdom’s demand for foreign health professionals is the greatest threat to developing countries’ healthcare systems, which is not addressed by the Code; therefore, they find it “difficult to believe that strengthening the code on its own will overcome the demand from U.K. employers for more staff to run their hospitals,” while “it is this demand that appears to be a principle cause of the drain of health professionals from English-speaking sub-Saharan Africa” (Ibid., 1895). Given that the Code is relatively new and there is a lack of reliable data available, it is difficult to make a definitive statement about the Code’s effect on developing states; however, it has been recognized to provide a “relatively weak regulatory mechanism” due to its lack of legal enforceability (Bach 2003, 22).

LESSONS FOR CANADA

The U.K. example can act as a useful comparison when determining what approach Canada should adopt to meet its obligation to ensure that Canadian recruitment of foreign health professionals does not adversely affect the healthcare systems of other countries. Like the United Kingdom, it has been recognized that foreign health professionals have “always played an important

role in and formed a substantial portion of Canada's health system" (McIntosh, Torgerson and Klassan 2007, 1). Canada struggles with a brain drain of its own; an estimated eight thousand Canadian-trained physicians practice in the United States (Martineau, Decker and Bundred 2004, 3), representing a substantial loss of human capital, as the emigration of Canadian doctors is estimated to cost \$100 million annually (McArthur 1999). Asia and Africa are the largest sources of physicians migrating to Canada; therefore, Canadian reliance on foreign healthcare professionals has the potential to adversely affect developing countries. Also similar to the U.K. experience, Canada was publicly rebuked in 2001 by the South African High Commissioner to Canada for recruiting South African doctors (McIntosh, Torgerson and Klassan 2007, 1). Although Canada cannot seek to comprehensively address the complex factors that influence international migration, it can most effectively meet its moral and legal obligations by ensuring that its domestic healthcare worker shortage does not have adverse effects on the sustainability of developing country health systems.

Canada has not effectively met its domestic demand for healthcare workers, as evidenced by the lack of physicians in many rural communities. Although the twenty-seven Organization for Economic Cooperation and Development countries increased total output from medical schools by an average of 26 percent between 1985 to 1994, Canada's output increased by only 18 percent (Bundred and Levitt 2000, 245). A major factor cited for Canada's lack of doctors is the 1991 Barer-Stoddard Report, which "predicted a future oversupply of physicians and recommended a 10 percent cut in medical residents...as a way of controlling costs" and led to training-restrictions (Ibid., 246). As a partial result of the insufficient training of doctors, foreign graduates accounted for 24 percent of all doctors practicing in Canada in 1999 (Martineau, Decker and Bundred 2002, 4). In an attempt to overcome a shortage in doctors, Ontario licenses many international medical graduates (IMGs). According to the president of the Ontario College of Physicians and Surgeons, the province is "quite dependent upon [IMGs] and they provide a great source of expertise that we need" (Mahoney 2007, A9). Of significance, the number of licences issued to IMGs in Ontario has tripled since 1995 to a record number of 1,247 in 2006, accounting for 42 percent of all licences issued in Ontario (College of Physicians and Surgeons of Ontario 2006, 4).

Examples of Canada's potential over-reliance on foreign professionals are readily available: Saskatchewan is highly dependent on South African doctors, as 17 percent of its 1,530 doctors in 2001 earned their first medical degree in South Africa (McIntosh, Torgerson and Klassan 2007, 1); in 1997, a professional recruiter hired by the Alberta Health Ministry traveled to South Africa and recruited forty South African doctors to fill rural community vacancies (Bundred and Levitt 2000, 246); while the newly-opened Centre for Spinal Injuries in Boxburg, South Africa, which was intended to operate as the region's referral centre, was closed down indefinitely following the recruitment of two of its anesthetists by a Canadian medical institution in 2000 (Mensah 2005, 209). Compounding Canada's lack of health care workers is its ageing population, as 25 percent of Canadians will be over sixty years of age by 2020 (Martineau, Decker and Bundred 2002, 7). Without determined action to

increase the output of Canadian medical schools in the face of an ageing population, the effect of Canada's lack of health professionals will be felt even more strongly in developing countries. Developed country healthcare systems need not suffer such sustained staffing shortages: Italy and Germany have an oversupply of doctors, which results in minimal unemployment, while France and Germany both rely on less than 5 percent foreign-trained doctors (Eastwood 2005, 1895). In light of these numbers, it is logical for Canada to recognize that the best response to the issue is to work toward self-sufficiency in terms of health professionals, while managed recruitment from abroad is the second best policy (McIntosh, Torgerson and Klassan 2007, 6). To successfully address this issue, Canada must identify and deal with its own push-pull factors by determining what policies would best serve to meet Canadian staffing needs and prevent losses to the United States.

When recruiting foreign health professionals, a "managed migration" approach is preferred, as "it attempts to link international migration to the health policy goals of individual nation States, and tries to regulate the flows of health workers in a way that is beneficial to source and destination countries" (Bach 2003, 21). Concluding bilateral agreements like those envisioned within the U.K.'s Code of Practice may be an effective way of ensuring that migration is well managed, as it reduces the role for commercial recruiters and can provide a more transparent process that is more easily controlled by regulatory mechanisms (Ibid., 23). Monitoring of any recruitment policy is central to its efficacy; without monitoring recruitment activities and the maintenance of a "living list" of proscribed countries that responds to changes in developing countries, the goal of managed migration will be difficult to attain (McIntosh, Torgerson and Klassan 2007, 14). Legally binding guidelines also promise a much higher degree of compliance than merely voluntary codes; however, the political will and resources required for such an approach would undoubtedly be difficult to attain (Ibid., 17).

Any Canadian response to meet its international obligation will face challenges due to the division of power between the federal and provincial governments. Unlike the United Kingdom, Canada does not have a centralized system of healthcare provision, while in states like Canada where the central government "has less control over the actions of health-care providers...the impact of ethical codes of practice are likely to prove much less effective" (Bach 2003, 23). The Federal Government has the authority to enter into international agreements, but the provision of healthcare services largely resides in provincial jurisdiction; therefore, future international agreements the Federal Government could enter into in order to address the issue of international recruitment may not be automatically binding upon the provinces. Since the right to mobility is enshrined in the Constitution through the Charter of Rights and Freedoms, this could undermine efforts taken by individual provinces because health professionals cannot be barred from migrating between provinces. Given these challenges, an effective solution requires a dialogue involving all interested stakeholders – professional associations, hospitals, regional health authorities and recruitment agencies – as well as the provincial and federal governments (McIntosh, Torgerson and Klassan 2007, 20).

Beyond addressing its pull-factors, Canada could help minimize push-factors, by focusing development assistance funds on developing and sustaining healthcare systems in developing countries, and help build capacity in states with weak healthcare systems by lending its expertise in the area (Eastwood 2005, 1897-99). A good example of an effective capacity-building effort is the Tropical Health and Education Trust, which provides teaching modules for medical undergraduate and masters programs in SSA (Ibid., 1898). There is also the possibility of compensation to foreign governments to offset the losses of human capital that arise from active healthcare professional recruitment. The Commonwealth Code of Practice for International Health Workers of 2002 places “a strong emphasis on mutuality of benefits for both countries, including compensation” (Bach 2003, 27). Compensation, however, may be politically unfeasible – it proved to be a primary obstacle to Canada, the United Kingdom and Australia signing the Commonwealth Code – while Canada’s federal government would likely be unwilling to compensate for provincial-level recruitment of health professionals, and both the provincial and federal governments would undoubtedly be unwilling to pay compensation for the activities of private recruiting agencies (McIntosh, Torgerson and Klassan 2007, 10). Compensation may also be “poor pittance” because it “might temporarily assuage the guilty conscience of the receiving country, but does little to replace a doctor who has taken five years to train and is a teacher and a role model to students and junior doctors” (Eastwood 2005, 1894).

CONCLUSION: INTERNATIONAL MANAGEMENT OF THE CAROUSEL

Given the complex nature of the issue of healthcare professional migration, it would be unrealistic to believe that one state can solve the overall problem – or even its own problems – independently. There may always be the drain of professionals to other states because “modern medicine in the more-developed world creates an insatiable demand for doctors, and globalization offers easy solutions to the trade in physicians” (Bundred and Levitt 2000, 246). The lack of supply of healthcare professionals in developed countries presents a major pull-factor for migration; therefore, without international co-operation that achieves the identification and realization of “internationally agreed minimum training targets for developed countries, the most vulnerable countries will continue to lose a large proportion of health workers” (Eastwood 2005, 1899). The pull from developed countries is expected to increase in the next ten to twenty years, as it is estimated that the US will require one million more nurses during this period (Martineau, Decker and Bundred 2004, 4), while the “graying” of the workforce presents its own challenges. Countries like Canada and the United States must address their practice of relying on foreign medical graduates, while turning away thousands of potentially qualified medical school applicants, or the negative effects of their healthcare policies will become more concentrated on developing countries like those in SSA (Martineau, Decker and Bundred 2002, 9).

In the end, codes of conduct will likely remain “only as strong as the will to conform to them” (McIntosh, Torgerson and Klassan 2007, 22). If Canada and

other developed countries engage in active international recruitment, it “needs to be fair, transparent, consider mutuality of benefits and reciprocity and adhere to the principles of global justice and personal autonomy” (Ibid., 6). In developing a policy, attention should be paid to the push-factors in developing states, as the “rebuilding of source country health systems in order to provide a decent wage and working conditions represents a response that is ethically acceptable and also has a chance of working” (Mensah 2005, 214). In this sense, a moral obligation exists to “assist [developing countries] in achieving workforce self-sufficiency, maintaining educational institutions, and providing quality of care” (McIntosh, Torgerson and Klassan 2007, 7). Developed countries must recognize that they owe more than just a moral obligation to developing states. A legal obligation exists for states to ensure that their policies do not adversely affect the right to health in developing countries; therefore, states that rely too heavily on foreign health professionals must act to resolve their shortages of healthcare professionals and slow the international health professional carousel that is undermining the healthcare systems of many developing states.

NOTES

¹ With the introduction of South African policies, to be discussed below, the recruitment of health workers from poor states in the region is restricted.

REFERENCES

- Bach, Stephen. 2003. International migration of health workers: Labour and social issues. *Working Paper: Sectoral Activities Programme*, WP.209. Geneva: International Labour Office.
- Buchan, James, and Delanyo Dovlo. 2004. *International Recruitment of Health Workers to the U.K.: A Report for DFID*. London: DFID Health Systems Resource Centre.
- Bundred, Peter E., and Cheryl Levitt. 2000. Medical Migration: Who are the Real Losers? *The Lancet* 356, no. 9225: 245-6.
- College of Physicians and Surgeons of Ontario. 2007. *Annual Physician Survey Results: 2006 Physician Resources in Ontario; Small Triumphs, Big Challenges*. March 2007.
- Committee on Economic, Social and Cultural Rights. 2000. *General Comment 14: The Right to the Highest Attainable Standard of Health*. Geneva: Committee on Economic, Social and Cultural Rights.
- Department of Health (U.K.). 2004. *Code of Practice for the International Recruitment of Healthcare Professionals*. London: Department of Health.
- Department of Health (U.K.). 2001. *Code of Practice for NHS Employers Involved in the International Recruitment of Healthcare Professionals*. London: Department of Health.
- Diallo, Khassoum. 2004. Data on the migration of health-care workers: sources, uses and challenges. *Bulletin of the World Health Organization* 82: 601-607.
- Eastwood, J.B. et al. 2005. Loss of health professionals from sub-Saharan Africa: The pivotal role of the U.K. *The Lancet* 365, no. 9474: 1893-1900.
- Lancet. 2000. “Medical migration and inequity of health care. *The Lancet* 356, no. 9225: 177.
- Mahoney, Jill. 2007. Licences for foreign trained MDs on increase; Ontario body issues overall total of 2,961. *The Globe and Mail*, March 30.
- Martineau, Tim, Karola Decker, and Peter Bundred. 2004. Brain Drain of Health Professionals: From Rhetoric to Responsible Action. *Health Policy*, 70: 1-10.
- Martineau, Tim, Karola Decker, and Peter Bundred. 2002. *Briefing note on international migration of health professionals: Leveling the playing field for developing country health systems*. Liverpool: Liverpool School of Tropical Medicine.

- McArthur, William. 1999. The Doctor Shortage: Part One. *The Fraser Forum*. The Fraser Institute, 1999. http://oldfraser.lexi.net/publications/forum/1999/06/06_doctor_shortage_part1.html.
- McIntosh, Tom, Rnee Torgerson, and Nathan Klassan. 2007. *The Ethical Recruitment of Internationally Educated Health Professionals: Lessons from Abroad and Options for Canada*. Ottawa: Canadian Policy Research Networks.
- Mensah, Kwadwo. 2005. "International Migration of Health Care Staff: Extent and Policy Responses, with Illustrations from Ghana." In *Commercialization of Health Care: Global and Local Dynamics and Policy Responses*, edited by Maureen Mackintosh and Meri Koivusalo, 201-215. London: Palgrave Macmillan.
- Osmani, Siddiqur Rahman. 2000. Human Rights to Food, Health, and Education. *Journal of Human Development* 1, no. 2: 273-298.
- Sen, Amartya. 2005. Human Rights as Capabilities. *Journal of Human Development* 6, no. 2: 151-166.
- Stilwell, Barbara, et al. 2003. Developing Evidence-Based Ethical Policies on the Migration of Health Workers: Conceptual and Practical Challenges. *Human Resources for Health* 1, no. 8 (2003), <http://www.human-resources-health.com/content/1/1/8>.
- Susser, Mervyn. 1993. Health as a Human Right: An Epidemiologist's Perspective on the Public Health. *Public Health Then and Now* 83, no. 3: 418 – 426.
- UN Commission on Human Rights. 2003. Statement by the World Health Organization. *Agenda Item 10: Economic, Social and Cultural Rights*. Geneva: UN Commission on Human Rights, http://www.who.int/hhr/information/en/item10_final.pdf.
- United Nations. 1966. *International Convention on Economic, Social and Cultural Rights*. New York, UN, 1966 (entered into force 3 January 1976).
- , 1948. *Universal Declaration of Human Rights*. New York: UN.
- Willettts, Annie, and Tim Martineau. 2004. *International recruitment of health professionals: Will codes of practice protect developing country health systems?* Liverpool: Liverpool School of Tropical Medicine, <http://www.liv.ac.uk/lstm/research/documents/codesofpracticereport.pdf>.
- World Bank. 1993 *World Development Report 1993: Investing in Health*. Washington: World Bank Group.
- World Health Organization (WHO). 2007. "Health and Human Rights" WHO, <http://www.who.int/hhr/en/>. (accessed 19 August 2007).
- , 2006. *World Health Report 2006: Working Together for Health*. Geneva: WHO.
- , 2003. International Migration, Health & Human Rights. *Health & Human Rights Publication Series* 4. Geneva: WHO.
- , 2002. 25 Questions and Answers on Health and Human Rights. *Health and Human Rights Publication Series* 1. Geneva: WHO.

HUMAN RIGHTS, CULTURAL TRADITION AND INTERNATIONAL DEVELOPMENT: CULTURALLY APPROPRIATE INTERNATIONAL RESPONSES TO FEMALE GENITAL CUTTING

Kristen Thomasen

Norman Paterson School of International Affairs, Carleton University

This article investigates the question of how to address, in an appropriate and effective way, an internationally defined violation of human rights that is also a cultural tradition. The practice of female genital cutting (FGC), an accepted and valued tradition, which has existed for centuries in some parts of the world, is investigated, and Amartya Sen's framework of human rights and development is used to critically assess development initiatives aimed at reducing rates of FGC. Sen's concepts of human rights and freedoms provide the foundation for the central argument of this article that: rates of FGC will only be reduced if families and individuals feel that their choice to reject FGC will not bring with it long term negative socio-cultural consequences for their daughters or themselves. The United States Agency for International Development (USAID), and its support for the Tostan program, is used as a case study to further consider the applicability of Sen's framework to development initiatives. It is argued that the Tostan approach to FGC was successful at reducing rates of the practice because it encouraged community-wide education, attitude change and abandonment of FGC, thereby reducing the potential negative socio-cultural consequences that might otherwise result from a rejection of the practice. The article concludes that this approach can be applied to other human rights issues, including early marriage and abortion. The wellbeing of the individuals affected by development policy should always be a top priority when policies are implemented.

INTRODUCTION

“Human rights, to justify that name, demand universality.”
(Sen 1999, 228)

The practice of female genital cutting¹ (FGC) is a tradition that has existed for centuries, but which has also been internationally accepted as a violation of basic human rights over the past two decades. International opposition to the practice of FGC inevitably raises the question of how a widespread practice, globally defined as a violation of human rights, can be handled appropriately and effectively by the international community when it is also a deeply embedded cultural tradition. Using the framework of development and human rights, as defined by Amartya Sen, this article argues that rates of FGC will only be reduced if families and individuals feel that their choice to reject FGC will not bring with it long term negative socio-cultural consequences for their daughters or themselves. By relaxing the social and cultural forces which surround FGC, thus allowing families to more freely choose to reject or pursue it, the cultural significance of the practice and the significance of protecting human rights can be respected simultaneously. A case study of the United States Agency for International Development (USAID) and its support for the non-governmental

organization (NGO) Tostan, a community education program based in Senegal, will be utilized to examine the applicability of Sen's theories to international development initiatives.

This article is divided into four sections. First, it provides a brief background to FGC as a global health and human rights concern. The second section provides the theoretical framework, which outlines how FGC can be discussed as a human rights and development issue. This section also assesses the international commitment to reduce rates of FGC. Third, this article evaluates the USAID policy on FGC through the theoretical framework, looking specifically at one successful case of policy implementation: financial support for the Tostan program. Finally, the article briefly explores how the theoretical framework applies or does not apply to other important development topics, specifically child marriage and abortion policies.

BACKGROUND TO FEMALE GENITAL CUTTING

The practice of female genital cutting (FGC) has existed for hundreds of years, primarily in Africa, although it has been also been carried out by communities in Asia and the Middle East. The practice is becoming more common in North America, Europe and Australia as a result of international migration of peoples from practicing cultures. FGC is defined by the World Health Organization (WHO) as comprising all procedures involving partial or total removal of the external female genitalia for cultural, religious or non-therapeutic reasons (WHO 2000). The age at which a girl may undergo the procedure will vary from region to region and can range between infant and mature woman, however it is typically performed between the ages of 4-12 years. FGC is usually carried out at the request of the girl's parents or relatives, and is often performed by traditional, medically untrained elderly female birth attendants, particularly in rural areas. The form and severity of FGC vary widely; the most common form is excision² and the most severe is infibulation³ (Ibid.).

A number of negative health consequences can arise as a result of FGC, which will vary according to the type and severity of the procedure performed. Short-term consequences include pain, shock and hemorrhage. Infection and hemorrhage are primary causes of short-term mortality (Ibid.). Because FGC is not consistently performed in medical settings, such as hospitals or clinics, the procedure is often carried out without any anesthetic, and the potential for transmission of infectious disease, including HIV/AIDS, can be high due to repeated use of the same, un-sterilized surgical instruments. Long-term consequences can be severe and include cyst and scar tissue formation, abscesses, frequent pain and obstetric complications, which increase the likelihood of infant and maternal mortality. Psychological consequences include feelings of incompleteness, anxiety and depression (Ibid.).

Despite the negative health consequences that can, and regularly do, arise as a result of FGC, prevalence rates are almost universal in a number of practicing regions. Rates vary widely, for example, in Cameroon prevalence was estimated at 1.4% in 2004, whereas in Guinea prevalence was estimated at 95.6% in 2005 (UNICEF 2006).⁴ In 2000, the WHO estimated that the number of girls and women who had undergone the procedure was between 100 and 140

million. According to these estimates, each year approximately 2 million girls are at risk of undergoing FGC (WHO 2000). The WHO recognizes five primary factors driving the practice, with the significance of each varying from region to region. These factors are: (1) sociological in nature (initiation of girls into womanhood, social cohesion and identification with cultural heritage, significance for marriage); (2) based on beliefs about hygiene (belief that FGC promotes hygiene and improves aesthetic appeal); (3) based on myths (enhancement of child survival, increased fertility); (4) based on religious belief (some communities practice because they believe it is demanded by Islam; while it predates the faith, it has been incorporated into Islam in many cases); and, (5) can be psychosexual in nature (the most controversial and gender-discriminating of the forces which act to maintain FGC, including attenuation of sexual desire, maintenance of chastity until, and fidelity during, marriage and enhancement of male sexual pleasure) (Ibid.). Underlying each of the above factors are nuanced cultural relationships of gender and power. While an understanding of these relationships is important to broader discussions of FGC, an in-depth investigation is beyond the scope of this article due to the complexity of these issues.

Motives for performing FGC vary widely between regions; however, one overriding factor influencing the continuation of the practice is community-wide embeddedness and acceptance of FGC. For example, if FGC is considered essential within the marriage market, whereby a non-circumcised woman may not be marriageable, then, for those who are dependent on marriage as a source of economic and social well-being, abandonment of FGC will be impossible (Oboler 2001, 316). If the majority of the community abandons the necessity for FGC in the marriage market, parents will have the freedom to choose to have their child circumcised or not, without the threat of significant social, economic, religious and/or cultural harm to their daughters.

Attempts to eradicate FGC have been ongoing since the seventeenth century when Catholic missionaries in Africa first took note of the practice (Boyle et al. 2002, 38). Despite long-term international attention to the issue, prevalence rates remain high, as identified above (UNICEF 2006). In the past, the power of the cultural forces which act to maintain FGC were oftentimes misunderstood, and, as such, numerous interventions have been misdirected. This article uses Amartya Sen's framework to examine a culturally appropriate approach to international development and the elimination of FGC.

THEORETICAL FRAMEWORK: HUMAN RIGHTS, DEVELOPMENT AND INTERNATIONAL OBLIGATIONS

The international debate over the elimination of FGC exemplifies the cultural relativism (the premise that practices specific to a culture cannot be properly judged by those outside that culture because they can only be interpreted in terms of one's own beliefs and cultural understandings) versus universalism (the premise that, in this case, fundamental human rights apply equally to all human beings regardless of culture, race, religion, etc) debate.⁵ An extensive literature spanning numerous disciplines, including international relations, international law, anthropology, and women's studies, has engaged in this debate (see Perry

(1997) for an excellent review of the theoretical aspects of the relativism/universalism debate; see Boyle (2002: 41-59) for a summary of the discourse from the mid-20th century to date). There is now a general consensus that FGC is a violation of human rights, however, many scholars and development agencies also recognize that cultural adaptation to eliminate FGC and promote the integration of new practices and norms must be done in a way that does not compromise the cultural integrity of the peoples for whom this practice is a tradition (Ibhawoh 2000, 839). In other words, though the cultural relativist argument has been insufficient in protecting the practice from international efforts to eliminate FGC, the embeddedness of the practice in some cultural groups is generally acknowledged. This article accepts the premise that FGC is a violation of human rights, and so moves beyond the universalism/relativism debate to focus more directly on culturally appropriate strategies for reducing rates of FGC, using Amartya Sen's framework of human rights to answer the question of whether and how to address a human rights violation that is also a cultural tradition.

Amartya Sen argues that while human rights and human capabilities have similar motivations, they are two distinct concepts (Sen 2005, 152); both, however, are relevant to a discussion of FGC. Human rights are roughly defined by Sen as one's rights to certain identifiable *freedoms*. This idea of freedoms incorporates both 'opportunity' and 'process.' The 'opportunity' aspect of freedom is similar to Sen's concept of capabilities, namely, the ability of an individual to take advantage of his or her personal means and prospects (Sen 2005, 154). If only the 'opportunity' aspect of freedom is considered, freedom and capability would appear to be essentially the same. Therefore, the 'process' aspect is crucial to Sen's framework of human rights, and to the application of this framework to FGC. The 'freedom of process' is one's ability to make a choice with respect to his or her actions. In short, 'freedom of process' measures an individual's agency, or power, to choose. Freedom in 'opportunity' and 'process' implies that the actions taken by an individual are the actions that the individual wishes to take *and* chooses to take, from among multiple options (Sen 2005, 152). Any action that is forced upon an individual, even if it is the action that individual might have taken anyways, is a violation of the 'process' aspect of that person's freedom. The 'opportunity' aspect, in contrast, is only violated if the individual is forced to take an action they might not have otherwise taken. An example, below, further elucidates this theoretical approach to human rights.

When applied to the case of FGC, Sen's framework asks whether one has the freedom to choose the actions that are carried out on one's body. The 'opportunity' and 'process' aspects of freedom have different, but equally significant applications. The 'process' of freedom is violated in communities in which the practice is an absolute requirement for females to be economically and socially supported. Parents have very little agency when the alternative to FGC has significant negative ramifications for their daughters. As the practice is traditionally carried out on young children, the girls who are circumcised also have little or no ability to opt into or out of the procedure; the 'choice' is made by parents or relatives. With this imposition, the opportunity aspect of freedom is violated if the parents of children on whom the practice is performed might

not have taken this action if another opportunity or course of action were available. Additionally, a woman's capabilities are limited if negative mental and/or physical health consequences result from the procedure; she will not be able to take full advantage of her set of personal abilities if she suffers from infection and pain. Therefore, according to Sen's concepts of 'process' and 'opportunity,' FGC is a violation of human rights if there is no freedom of choice associated with its application.

This article focuses specifically on the concept of choice as defined in Sen's framework. An additional argument, worthy of mention with respect to FGC, is Sen's discussion on cultural relativism. In his 2005 article, "Human Rights and Capabilities," Sen addresses the relativism/universalism argument with respect to human rights. He re-interprets the Rawlsian theory of public reason to state,

The role of public reasoning in the formulation and vindication of human rights is important because any general plausibility an ethical claim (or its denial) will have in terms of human rights will be dependant upon its ability to survive and flourish when it encounters unobstructed and well informed public discussion and scrutiny (Sen 2005, 160).

In other words, the universality of a human right is justified when it is publicly accepted as legitimate. While an extensive discussion of the universalism versus relativism of human rights debate is beyond the scope of this article, the concept of universal acceptance is important to the elimination of FGC. Community-wide recognition that FGC is a violation of rights, or even just that FGC is not an absolute necessity for women, allows the social forces that serve to maintain the practice to be relaxed, and thus enhances parental freedom of choice with respect to the decision to have a daughter circumcised. The cultural relativism framework acts to reinforce the applicability of Sen's framework of human rights by emphasizing the importance of community-wide recognition of FGC as a violation of human rights. This concept of community-level attitude change is significant to the success of the USAID policy and the Tostan program, as discussed below.

FGC, HUMAN DEVELOPMENT, AND INTERNATIONAL OBLIGATIONS

"Development can be seen ... as a process of expanding the real freedoms that people enjoy."
(Sen 1999, 3)

Sen's framework of human rights answers the question of *how* to approach FGC, but it does not explore the equally important question of *why*. This issue can be answered using a human development framework. As put forth by Sen in his book, *Development as Freedom*, development requires the removal of major sources of unfreedom (Sen 1999, 2). Unfreedom in the case of FGC is the inability of parents and women to make a choice, free from cultural persuasion, about whether or not to pursue circumcision. This is significant to human

development, which focuses on improving overall well-being, because human achievement is influenced by many factors including economics, political liberty, *and* good health (Sen 1999, 5). FGC is perpetuated by unfreedoms and inhibits a woman's ability to function to her fullest potential, because of its numerous negative mental and physical health implications.

This development framework is expanded by considering international commitments to development goals. At the United Nations (UN) Millennium Summit in 2000, world leaders agreed to the eight Millennium Development Goals (MDGs), a set of time-bound targets at the heart of the global development agenda. The MDGs provide a framework through which the UN system can work to improve the livelihood of all individuals, but particularly the world's poorest. Eradication of FGC fits into this framework through three goals: Goal 3, Promote Gender Equality and Empower Women; Goal 4, Reduce Child Mortality; and, Goal 5, Improve Maternal Health (UN MDGs 2005). International commitments and obligations again serve to reinforce the importance of developing culturally appropriate strategies for reducing the prevalence of FGC. This is especially important in countries where the practice is prevalent, as there can be significant local pressure against eradication efforts. Decreasing the prevalence of FGC must be both incorporated *into* development programs, for example, through education efforts, community development and human rights, and must be a *goal* of development programs in order to attain these international targets.

THE CASE OF USAID AND TOSTAN: ADDRESSING A CULTURALLY EMBEDDED VIOLATION OF HUMAN RIGHTS

This section examines the USAID policy on FGC, and specifically the Agency's support for Tostan, in order to elucidate the applicability of Sen's theory of human rights to international development initiatives. A brief history of the incorporation of FGC into American foreign policy is outlined below to explicate USAID's motivation for incorporating its FGC policy into its development spending. Subsequently, an analysis of different strategies for eliminating FGC provides the foundation upon which the argument is made in favour of choice-based development strategies, based on the framework provided by Sen. The lessons-learned from past strategies and the theoretical discussion above are used to explain why the Tostan program is a successful strategy for eliminating FGC.

CURRENT U.S. AND USAID POLICIES ON FGC

American foreign policy interest in FGC in developing countries arose because of domestic concerns over the practice, which surfaced in the United States in the 1990's as a result of migration of peoples from practicing countries. On September 30, 1996, the Department of Defense Omnibus Appropriations Bill was passed in the United States with anti-FGC legislation as part of this Bill (Dept. of Defense 1996). The legislation had both domestic and international concerns. It criminalized the practice of FGC on females under the age of 18 in

the United States, as an “infringement on the guarantees of rights secured by Federal and State law, both statutory and constitutional” (Dept. of Defense 1996, Sec. 645). The law also made education about the physical, medical and legal implications of the practice mandatory for all immigrants coming from cultures where FGC is practiced (Dept. of Defense 1996, Sec. 644). Finally, it instructed the U.S. Executive Director of each international financial institution to “use the voice and vote of the United States to oppose any loan or other utilization of the funds of their respective institution, other than to address basic human needs,” (Sec. 579) for the government of any country which the Secretary of the Treasury determines has a history of FGC and has not taken steps to implement programs designed to prevent the practice (Dept. of Defense 1996, Sec. 579). The first two aspects of the law are domestic in nature. National criminalization of FGC is largely symbolic due to the low rates of the practice in the United States. The third component of this public law links American foreign aid to the actions taken by other countries to eradicate the practice. The outward-looking nature of this component of the legislation laid the foundation for different government departments to implement anti-FGC policies of their own.

The United States, as one of the largest donors of bilateral aid, is a significant player in the international development community.⁶ Although USAID’s policies are sometimes controversial with other donors (for example, the overly strategic use of aid and USAID’s policies on HIV/AIDS prevention have been widely critiqued), the Agency is a leader within the development community. This article examines the USAID policy on FGC for this reason; the significance of American development aid to many of the developing countries where FGC is prevalent means that USAID (and consequently the United States) controls a powerful policy tool through which it can impact the domestic policies of other countries.⁷

United States Agency for International Development

The official USAID policy on FGC states that:

USAID recognizes FGC as a harmful, traditional practice that *violates the health and human rights* of women and *hinders development*. USAID opposes any practice of or support for Female Genital Cutting (FGC) and works toward the goal of total elimination of FGC (USAID 2000).⁸

The policy was designed to support existing U.S. Government and international policies and efforts to reduce rates of FGC around the world. In effect since September 1, 2000, the Policy Statement identifies two actions, which specifically direct USAID policy towards community-based, locally-engaged strategies; it is for this reason that USAID’s policy is argued to be appropriate. The two actions are: to support local-level actors (e.g., NGOs, community leaders, religious organizations) in order to ensure that attempts to eliminate FGC are culturally appropriate; and to target stakeholders, including men, and boys and to focus on reducing and eventually eliminating demand for the practice (USAID 2000).⁹ As will be argued below, development agencies, exemplified here by USAID, can appropriately seek to reduce rates of FGC by

adopting policies that direct funding and other actions towards locally oriented initiatives, such as the Tostan program.

Alternative Strategies

To answer the question of how to appropriately address a practice like FGC, an exploration of the effectiveness and lessons-learned from previously employed strategies is fitting. Past strategies for reducing rates of FGC have included: working with domestic governments to enact formal legislation; finding new employment for the women who perform circumcisions; debasing cultural justifications; educating mothers and communities about the negative health consequences associated with the practice; encouraging alternative rites of passage; and, the de-medicalization of the practice (Shabaan and Harbison 2005; USAID 2006a).

While not advocating that any of these strategies cease, it is crucial to recognize that the implementation of any one of these approaches individually will not decrease the demand for the practice significantly. Formal legislation supports efforts to reduce rates of FGC, however without enforcement and without public acceptance of the validity of such a law, legislation will have little effect on the rate at which FGC is carried out (Oboler 2001, 317). Employment for the traditional birth attendants who perform the practice is crucial, particularly through a development framework; however, this should follow a reduction in demand for such services. If families want their daughters to be circumcised, they will find alternative outlets through which to do so. Education about the negative health consequences of FGC is crucial to ending the practice. However, if a girl lives in a community where FGC is a requirement for marriage, for example, education alone is not enough, as a change in the community perception of the practice is necessary to relax or reconsider the social constraints on marriageability. The same problem applies to alternative rites of passage. Finally, de-medicalization, a more heavily debated issue where doctors/nurses are encouraged not to provide the service of performing circumcision,¹⁰ is important but will not alone change prevalence rates, as demand for the practice can move to alternative, potentially less safe providers (UN News Service 2007). In order to effectively address FGC, the broader cultural, social and religious implications of the practice must also be addressed and families must feel comfortable that their choice not to circumcise their daughter will not have negative long term implications for her or the family as a whole.

What Makes a Strategy Effective?

Interventions need to make it possible for people to *choose* whether or not to have FGC performed on them or their daughters. In order to successfully accomplish this, demand for the practice must be reduced and the economic and socio-cultural implications for a woman who has not been circumcised must be relaxed. For example, “if FG[C] is considered essential within the relevant marriage market, nobody can afford to drop it unilaterally without planning to move out of that market” (Oboler 2001, 316). Thus, if a woman has moved away from that marriage market, for example through emigration to North America or

Europe, she can afford not to be circumcised or have her children circumcised. In order to deal with the practice where it occurs most frequently, however, a community-based approach is required. Oboler (2001, 317) argues that elimination strategies must move beyond criminalization of the practice, and instead must focus efforts on community-based programs, which promote attitude and behavioural changes. “Effective eradication depends on attitude change, which can be best accomplished by gradual persuasion from community members committed to change” (Oboler 2001, 313).

Effective implementation, therefore, will recognize that traditional cultural beliefs are neither monolithic nor unchanging. Cultural change can result from individuals being exposed to new ideas, and being provided with the opportunity and freedom to adopt these ideas (Ibhawoh 2000). Amartya Sen’s framework of human rights supports this argument. Working within Sen’s framework, effective programs will be ones which recognize that those who choose to adopt these new ideas initiate a process of change, which may influence dominant cultural traditions (Ibhawoh 2000, 841). In the case of FGC, change can only come about as a result of broader cultural adoption of the norm or acceptance of non-circumcision. This article builds on this concept and argues that the USAID policy is effective, because it recognizes the importance of community-level action, which focuses on improving the individual’s ability to make an informed and free choice and thus encourages support for programs such as Tostan. The Tostan program is equally appropriate because it is designed to encourage voluntary, community-driven behaviour and attitude change. USAID has primarily provided financial support to Tostan for the implementation of the program in Guinea and for the evaluation of the program in Senegal (USAID 2006c). The Tostan program is not solely supported by USAID; numerous other international organizations and NGO’s are involved with Tostan, such as the Population Council and UNICEF. Additionally, *Deutsche Gesellschaft für Technische Zusammenarbeit* (GTZ), the German development agency, and the Swedish International Development Agency (SIDA), along with other development agencies and several private donor organizations, have provided substantial financial support to Tostan. This article focuses on USAID as *one* example of how development agency policy should be geared towards supporting appropriate means of addressing FGC. As argued by this author, Tostan is the means through which USAID effectively and appropriately seeks to reduce rates of FGC. The arguments put forth in this article apply to other international development agencies as well.

Tostan

Tostan is a non-profit NGO based in Thiès, Senegal, whose mission is to “provide human dignity to African people through the development and implementation of a participatory ... education program in national languages” (Tostan 2006). The village-based education program teaches residents about hygiene, human rights, women’s health and problem solving. Through education seminars, women learn about health issues and their rights with a central and significant focus on female genital cutting (Shabaan and Harbison 2005, 384). Tostan gives the responsibility to the participants, particularly women, to make

their own informed decisions as to whether FGC is right or wrong, or in other words, whether it is a violation of human rights or a tradition which must be carried on.

A key feature of the Tostan program is the “public declaration” against FGC, an activity which includes men, women, religious leaders and other stakeholders (Shabaan and Harbison 2005, 348). The declaration involves members of the community being given the opportunity to decide if they feel that FGC is harmful and should no longer be pursued. They may then make this opinion public at the declaration. The voluntary aspect of the declaration is crucial, and community members are invited, but not coerced, to join and publicly choose to oppose FGC. The community approach exemplified by the declaration reflects African socio-cultural realities, where decisions are generally taken by groups rather than by individuals (Diop et al. 2004, 35). It also builds into the framework provided by Sen, that the freedom to choose not to perform the practice on one’s children, or have it performed on oneself, can only be met by relaxing the social forces which maintain the practice, such as marital expectations and religious beliefs. Tostan and the declaration are run by indigenous actors with an emic, or local, understanding of the cultural variables that work to maintain FGC.

As of November 2000, FGC had been abandoned in a total of 174 villages where community members have participated in the Tostan program, affecting approximately 140,000 people (U.S. Department of State 2000). The program was implemented, for example, in ninety villages in the Kolda region of Senegal and an evaluation of this program by Diop et al. (2004)¹¹ in twenty of the ninety Senegalese villages showed that significant progress had been made in eliminating FGC. The proportion of women who thought it necessary to practice FGC decreased significantly in the intervention group immediately after the program (Diop et al 2004, 19). For example, in the baseline survey,¹² seven out of ten women said they wanted to have their daughters, who had not yet been circumcised, cut in the future. This proportion fell to about one in ten women among participants and two in ten amongst women indirectly exposed to the program¹³ (Diop et al 2004, 22). This finding, along with the public declaration held on June 5, 2002, is evidence of attitude change as argued by Diop et al. At the declaration, women and men denounced FGC as well as all forms of gender discrimination and violence. There was a dramatic decrease in the approval of FGC from both men and women living in the villages subject to interventions (Diop et al 2004, 32). Evidence of behavioural change includes the finding that the prevalence of FGC among daughters of women participating in the program decreased overall, and the proportion of non-cut girls aged 0-10 years increased significantly (Diop et al 2004, 33). The program was replicated in Burkina Faso, with similarly successful results (Shabaan and Harbison 2005, 348). The U.S Department of State account of USAID’s support for Tostan states,

These programs suggest that it is the women, men and leaders of the countries themselves who must bring about this cultural and social change. Governments and citizens of societies where FGM is practiced must take the initiative. It is clear, however, that outside support is often desired and helpful. The United

States can best support eradication of FG[C] by providing financial aid and technical assistance (U.S. Department of State 2000, 17).

Support for Tostan directly meets the two actions identified in the USAID Policy Statement (support for culturally appropriate local-level organizations and working towards eliminating demand for the practice), which direct Agency policy towards supporting locally-based initiatives.

AN EFFECTIVE APPROACH - THE HUMAN RIGHTS AND DEVELOPMENT FRAMEWORK

In response to the original research question posed by this article, the international community can effectively, from a results-based perspective, and appropriately, from a cultural perspective, work towards the elimination of FGC through a development and human rights framework. Tostan, for example, took an approach that allowed individuals to make informed decisions regarding FGC and gave these individuals the freedom to choose whether or not to pursue the practice by supporting community-wide attitude and behaviour change through the public declaration. Within Sen's framework, Tostan's success can be explained by assessing how the program addressed the 'opportunity' and 'process' aspects of freedom (Sen 2005). The 'opportunity' aspect of freedom may not be affected as a result of the Tostan program, such that any individuals who would have chosen to pursue circumcision for themselves or their daughters, regardless of whether another choice was available, can continue to pursue circumcision. What is crucial about the Tostan program is that it changes the 'process' aspect of freedom. By relaxing the socio-cultural forces acting to maintain FGC in a community (through such activities as the public declaration and the community-wide attitude change that came with it [Diop 2004, ii]), the negative consequences of not pursuing circumcision are also relaxed. Thus, for those individuals who would choose not to pursue circumcision, if given other options, their 'process' aspect of freedom is no longer violated. When addressing a violation of human rights that is also a significant cultural tradition, such as the act of FGC, this author argues that the educated choice of whether to continue the practice or not must rest primarily in the hands of those for whom the practice is tradition. FGC is not a cultural practice that can be quickly eliminated through the enactment of national or international laws, or even by educating communities about FGC's negative health consequences. If rates of the practice are to be reduced, the individuals for whom FGC is a tradition must feel comfortable that the choice to cease FGC will not have significant negative social and personal consequences.

From a development perspective, gender equality and maternal and infant health have arguably been improved as a result of the decrease in prevalence rates of FGC, as the health issues associated with the practice of FGC, discussed above (WHO 2000), will affect fewer women. Due to the reduction in prevalence rates, it may be more likely that in future generations fewer women will suffer from the physical and/or mental health complications associated with FGC and will thus conceivably have a greater opportunity to be economically productive.

Finally, support for the Tostan program exemplifies Sen's discussion of cultural relativism and the universality of human rights as public scrutiny of the ethical claim that FGC violates human rights led to rejection of the practice of FGC, thus further establishing the legitimacy of viewing FGC as a violation of rights (Sen 2005, 160-161). Tostan is a prime example of a program that achieved the human rights standards outlined by Sen. Any development program, which increases freedom of choice, however, should be considered successful by Sen's standards.

CONSIDERATIONS, IMPLICATIONS AND POLICY RECOMMENDATIONS

Numerous considerations, implications and recommendations emerge from this analysis of appropriate FGC strategies. The first and most obvious recommendation to emerge from this examination is that USAID and other international development agencies should increase support for locally-engaged organizations, like Tostan, as an appropriate and effective way to decrease prevalence rates. A successful approach such as the one outlined in this article should be at the forefront of international efforts to address FGC. Criminalization, de-medicalization, the adoption of alternative rites-of-passage and other methods can provide broader support for approaches that are community-based and focus on reducing the demand for FGC.

A second recommendation is that the approach employed by USAID and Tostan should be expanded to other similar practices, which restrict choice, and inhibit an individual's capabilities and freedoms. Child marriage is one example of a practice that, like FGC, inhibits or eliminates an individual's ability to *choose*. Although early marriage does not have the same immediate health consequences as FGC, it does have significant indirect consequences such as increasing a girl's risk of experiencing gender-based violence, creating a greater likelihood of contracting HIV/AIDS, and subjecting the girl to the risks associated with adolescent pregnancy, which include higher rates of maternal and child mortality (UNICEF 2005; USAID 2006b). The young age of girls at marriage, community and family pressures and economic and gender-based factors all suggest that the girl's choice, with respect to when and whom to marry, is limited (UNICEF 2005, 1). Additionally, early marriage may limit educational prospects, thus limiting future freedoms and opportunities (Ibid.). Early marriage therefore, like FGC, restricts the individual's freedoms and agency. Development programs, which seek to increase the age of girls at marriage (and effectively eliminate child marriage), could work to expand the ability of families and girls to *choose* when and whom to marry. Factors such as poverty, social expectations and gender inequality (UNICEF, 2005) may all play a role in limiting one's ability to choose in this case. Further investigation is required to determine the main barriers to such choices, which may vary between regions and cultures.

The first concern to emerge from this discussion relates to the public declaration associated with the Tostan program. The declaration raises the question of whether there is public pressure for community members to join the event, and whether there may be negative social consequences for not publicly

opposing FGC. If stakeholders within the community participate in the declaration as a result of social pressure and not sincere, voluntary acceptance that FGC is a violation of human rights, then the reduction in rates may not be sustainable. Additionally, if individuals who support FGC are coerced into public opposition of the practice, it could be argued, according to Sen's framework, that this is also a violation of the process aspect of freedom. Further examination of the social forces driving an individual or family to join the public declaration in opposition of FGC is required.

A second consideration to emerge from the discussion in this article is that of the longer term effects of cultural change. Elimination of FGC may be viewed in the short term as a success from a development and human rights perspective, however longer term investigation into the effects of an elimination or reduction of the practice on gender and power relationships should be pursued. FGC is a deeply embedded cultural tradition, and arguably, elimination of such a tradition could have significant social consequences. An understanding of long term implications of cultural change can help guide initiatives, which seek to eliminate the practice, in more appropriate directions.

A third consideration is the risk of overlooking the significance of individual and community-level choice in development programs and the potential for negative consequences associated with this. Bilateral aid allocation is one of the foreign policy tools available to donor countries and is determined by both the self interest of the donor country and the recipient country's need for aid (Neumayer 2003, 650). However, state interests should not supersede the interests of those individuals affected by a development initiative, or else the goal of improving the lives of others might be negated. USAID and other international actors must take caution to consider freedom and choice when implementing development policies. For example, from 1984 to 1993, and reinstated in 2001, the American Global Gag Rule required NGOs to agree that any funds received from the USAID would not be used to perform or promote abortion (Cohen 2000, 137). This policy was criticized in both popular media and academic literature as a case of domestic policy and state self-interest taking priority over the well being of individuals in the aid recipient state. While an in-depth investigation of these arguments is beyond the scope of this article, the criticism of the Global Gag Rule provides an important reminder of the relevance of the argument made in this article; that the goal of development strategies should be to encourage and expand the range of choices and enhance the ability of aid recipients to make free and informed choices. If the Global Gag Rule in fact limits a woman's ability to *choose* with respect to abortion, then further analysis of the appropriateness of the policy is necessary. The USAID approach towards FGC differs from the abortion policy in that an acknowledgement of local needs and concerns is written explicitly into the policy. Making local concerns a central focus of development policy enhances the sustainability, appropriateness and effectiveness of the initiatives which are implemented as a result of such a policy.

CONCLUSION

This article has sought to answer the question of how a widespread violation of human rights can be appropriately and effectively handled by the international community when it is also a deeply embedded cultural tradition. By analyzing the USAID policy and the Agency's support for Tostan, this article concludes that such traditional practices must be addressed at the local-level, and must work to allow individuals to make their own informed decisions. This article has argued that the USAID approach was effective because it recognized the importance of community education and voluntary, locally-driven behaviour and attitude change. The discussion in this article contributes to the literature on human rights and development not only by explaining that FGC *is* a human rights issue, but also *why* it is appropriate that donors address it in their development policies, and *how* they can implement those policies effectively.

NOTES

¹ Strong opponents of the practice tend to use the term 'genital mutilation' to emphasize its negative implications for the women on whom it is performed, however this term connotes more vicious motives on the part of parents who have the procedure performed on their daughters than this author believes is fair. Proponents tend to use 'female circumcision' to compare it more closely to male circumcision, where in fact FGC is more similar to male castration. Throughout this paper the practice will be referred to as 'female genital cutting,' as this is a more neutral term, similar in meaning to 'female genital modification.' For lack of appropriate terminology however, the act of performing the procedure will occasionally be referred to as 'circumcision.' This author does not equate at least the most severe and invasive types of FGC with male circumcision.

² Involves removal of the clitoris and labia minora - Type II as defined by the WHO (2000)

³ Involves removal of part or all of external genitalia and stitching/narrowing of the vaginal opening - Type III as defined by the WHO (2000)

⁴ Statistics of the prevalence of FGC among women age 15-49 (including both rural and urban statistics).

⁵ These definitions are provided by the author.

⁶ For example, in 2005, the Development Assistance Committee of the Organization for Economic Cooperation and Development (OECD) showed the United States to be the largest donor of official development assistance (ODA) in the world, at USD\$27.62 billion. While not all American ODA is bilateral, the role and impact of the United States as an aid donor is significant (OECD 2005)

⁷ Boyle et al. (2001), for example, provide a case study of the effect of American aid in Tanzania. Soon after the adoption of the Dept. of Defense Omnibus Appropriations Bill, the Tanzanian Government began to act on international pressures to eliminate FGC by enacting formal legislation against the practice. The Tanzanian law, which soon followed, was written in English and phrased in much the same way as domestic American state laws against FGC (Boyle et al 2001, 534).

⁸ Emphasis added.

⁹ The Policy Statement outlines two additional actions, not addressed in this article. These are: to update the Agency strategy to incorporate FGC; and to develop a framework for research and advocacy to enhance eradication efforts (USAID, 2000).

¹⁰ Reliance on biomedically trained doctors became an alternative to the traditional birth attendants when parents started to learn about the negative health consequences.

¹¹ Funded by USAID.

¹² Survey of the twenty villages before the program was implemented.

¹³ Who learned about the information conveyed at the educational seminars through word of mouth (Diop et al 2004, 22).

ACKNOWLEDGEMENTS

The author would like to express her sincere appreciation to all those who read through her article, particularly Predrag Tomic; her co-editor Marion Sandilands; Laurie Gorman; Dr. Chantal Blouin; and her expert reviewer; for their incredibly insightful feedback.

REFERENCES

- Boyle, Elizabeth Heger, Fortuna Songora and Gail Foss. 2001. International Discourse and Local Politics: Anti-Female-Genital-Cutting laws in Egypt, Tanzania and the United States. *Social Problems*. 48(4): 524-544.
- Boyle, Elizabeth Hedger. 2002. *Female Genital Cutting*. Baltimore: John Hopkins University Press.
- Bureau for Global Health. 2004 *Strategy for Female Genital Cutting (FY 2004-FY 2006)*. Washington. http://www.usaid.gov/our_work/global_health/pop/techareas/fgc/fgc_strategy.pdf. (accessed March, 2007).
- Cohen, Susan A. 2000. Abortion Politics and U.S. Population Aid: Coping with a Complex New Law. *International Family Planning Perspectives*. 26(3): 137-139+145.
- Department of Defense Omnibus Appropriations Bill.1996. Public Law No. 104-208, 110 Stat. 3009-3708.
- Diop, Nafissatou, Modou Mbacke Faye, Amadou Moreau, Jacqueline Cabral, Hélène Benga, Fatou Cissé Babacar Mané, Inge Baumgarten, and Molly Melching. 2004. *The TOSTAN Program: Evaluation of a Community Based Education Program in Senegal*. The Population Council. August. http://www.popcouncil.org/pdfs/frontiers/FR_FinalReports/Senegal_Tostan%20FGC.pdf. (accessed March, 2007).
- Ibhawoh, Bonny. 2000. Between Culture and Constitution: Evaluating the Cultural Legitimacy of Human Rights in the African State. *Human Rights Quarterly*. 22: 838-860.
- Neumayer, Eric. 2003. Do Human Rights Matter in Bilateral Aid Allocation? A Quantitative Analysis of 21 Donor Countries. *Social Science Quarterly*. 84(3): 650-666.
- Oboler, Regina Smith. 2001. Law and persuasion in the Elimination of Female Genital Modification. *Human Organization*, 60(4): 311-318.
- OECD. 2005. *Final Official Development Assistance (ODA) Data for 2005*. Development Co-operation Directorate (DCD-DAC). <http://www.oecd.org/dataoecd/52/18/37790990.pdf> (accessed August 16, 2007).
- Perry, Michael. 1997. Are Human Rights Universal? The Relativist Challenge and Related Matters. *Human Rights Quarterly*. 19: 461-481.
- Sen, Amartya. 1999. *Development as Freedom*. New York: Knopf Press.
- . 2005. Human Rights and Capabilities. *Journal of Human Development*, 6(2): 151-166.
- Shabaan, Layla and Sarah Harbison. 2005. Reaching the Tipping Point Against Female Genital Mutilation. *The Lancet*. 366: 347-349. www.thelancet.com. (accessed March, 2007).
- TOSTAN. 2006. *About TOSTAN*. <http://www.tostan.org/about.htm>. (accessed March 27, 2007).
- UN Millennium Development Goals (MDG's). 2005. *The UN Millennium Development Goals*. <http://www.un.org/millenniumgoals/>. (accessed March 27, 2007).
- UN News Service. 2007. *'Medical Trend' with Female Genital Mutilation Disturbs UN Agency*. 5 February, 2007. <http://www.un.org/apps/news/story.asp?NewsID=21456&Cr=women&Cr1>. (accessed March 2007).
- UNICEF. 2006. *Prevalence of FGM/C among Women 15-49, by Place of Residence*. UNICEF Statistics. <http://www.childinfo.org/areas/fgmc/tables.php> (accessed August 16, 2007).
- . 2005. *Early Marriage – A Harmful Traditional Practice: A Statistical Exploration*. New York, NY: UNICEF.
- USAID. 2007. *About USAID*. http://www.usaid.gov/about_usaid/. (accessed March 25, 2007).
- . 2006a. Accelerating the Abandonment of Female Genital Cutting: Community Change to Support Human Rights. *Global Health Technical Briefs*. <http://www.maqweb.org/techbriefs/tb31fgc.shtml>. (accessed March 2007).
- . 2006b. *Issue Brief: Preventing Child Marriage: Protecting Girls' Health*. Family Planning. http://pdf.usaid.gov/pdf_docs/PDACI387.pdf. (accessed March 2007).
- . 2006c. *Issue Brief: Working to Eradicate Female Genital Mutilation/Cutting*. Family Planning. http://www.usaid.gov/our_work/global_health/pop/publications/docs/eradicatefgc.html. (accessed March 2007).

- , 2000. *USAID Policy on Female Genital Cutting (FGC)*. Family Planning. File Name: 200mac_111400_cd22. http://www.usaid.gov/our_work/global_health/pop/techareas/fgc/fgc.html. (accessed March 2007).
- WHO World Health Organization. 2000. Female Genital Mutilation *Fact Sheet No. 241*. <http://www.who.int/mediacentre/factsheets/fs241/en.html>. (accessed March 2007).

MAINTENANCE OF SANCTIONS REGIMES: LESSONS FROM IRAQ

Nathan Lysons

Norman Paterson School of International Affairs, Carleton University

Following the Cold War, the United Nations Security Council has used sanctions as a regular tool for influencing states and maintaining international peace and stability. The literature surrounding the use and effectiveness of sanctions has grown accordingly, but more attention is needed in determining how to create and maintain effective sanctions regimes. Using Iraq as a case study, a number of policy recommendations can be made. This article will argue that in order to address problems of capacity, developed countries should assist in systematically building capacity in countries lacking required enforcement mechanisms, while countries with adequate resources should take measures to ensure effective self-monitoring. To reduce the non-enforcement of sanctions regimes arising from these collateral economic losses, the SC could step up recognition of the negative consequences on affected countries and intensify efforts to provide compensation. The use of targeted sanctions, with defined, but renewable time frames will reduce humanitarian costs, allow greater proportionality, and thus increase the legitimacy of sanctions regimes. In the future, the Security Council should act to shift state interests through both positive inducements and secondary sanctions in order to build and maintain effective sanctions regimes.

INTRODUCTION

With the end of the Cold War, the United Nations Security Council (SC) was freed from ideological bonds and began to use sanctions as a regular tool of coercion. Supporters of sanctions argue that they are justified as a tool which is less costly than war, but more forceful than rhetoric.¹ As sanctions have become more common, the scholarship regarding their use, effectiveness, and nuances has also grown.² While the sanctions literature to date has provided valuable insights regarding the effectiveness of sanctions in obtaining compliance, there has been less research focused on finding best practices to create and maintain effective sanctions regimes. This article uses the sanctions regime authorized by the SC against Iraq as a case study to examine different methods that can be used to make sanctions regimes more effective. The article will begin by briefly highlighting areas that potentially undermine sanctions regimes, and will then look at those issues in the context of Iraq in order to explore factors that contribute to sanction busting. From this analysis, this article will propose a number of prescriptions for enhancing compliance with sanctions regimes. The need for states to find more effective ways to maintain sanctions regimes is real, as threats to international peace and security and corresponding UN sanctions are not abating.³

POTENTIAL WEAKNESSES IN SANCTIONS REGIMES

In the context of restoring or maintaining international peace and security, Article 41 of the Charter of the United Nations gives the SC the authority to employ sanctions as a coercive tool. The SC, however, does have some freedom of interpretation when determining what constitutes a threat to international peace and security. Thus, the SC has applied sanctions regimes in response to a variety of egregious international ills. The following examples are representative of situations that have motivated the SC to establish sanctions regimes. In 1977, the SC passed Security Council Resolution (SCR) 418, which implemented a mandatory arms embargo against South Africa, largely in response to apartheid and its associated humanitarian consequences. Iraq's invasion and attempted annexation of Kuwait in 1990 led to the comprehensive sanctions regime established by SCR 661 against Iraq. The 1994 genocide in Rwanda resulted in SCR 918, which established an arms embargo, among other measures, against that country. Travel and financial sanctions against Afghanistan in 1998, established by SCR 1267, were motivated by the threat of international terrorism. Other sanctions regimes have been established to combat nuclear proliferation, such as SCR 1718 (2006), which targeted North Korea, and those sanctions that have been levied against Iran through SCR 1737 (2006) and SCR 1747 (2007).

As long as the SC is acting within its authority under Chapter VII, the Charter of the United Nations requires that all UN members implement SCRs (Charter of the United Nations, Art. 48). For the most part, scholars tend to assume that states will attempt to fulfill their obligations under international law in good faith.⁴ At the same time, David Cortright and George Lopez, two of the foremost experts on the use of sanctions, note that UN sanctions require state compliance for effectiveness; the UN by itself does not have the institutional capacity to enforce sanctions, although, they argue, it could play a more effective role in coordinating and monitoring sanctions regimes than it currently does (Cortright and Lopez 2005). For these reasons, many of the suggestions which seek to enhance enforcement look to build state capacity for domestic UN sanctions enforcement. Proposed capacity building methods include passing domestic legislation to comply with UN sanctions, and building domestic legal, judicial, policing, and monitoring capacity (Biersteker et al. 2005). As well, in recognizing the impact that sanctions have on non-target states, another proposal is to resolve economic concerns of indirectly impacted states in order to ensure stronger compliance. This approach stems from Article 50 of the UN Charter which gives affected non-target states the "right to consult the Security Council with regard to a solution for those problems." If the SC resolves the economic concerns of the third party states, it is argued, the sanctions regime will be strengthened (Al-Khasawneh 2001, 333).

Beyond strengthening domestic capacity to implement sanctions, the literature also suggests that sanctions which are perceived to be legitimate, addressing real concerns with objective clarity and attainability, will enjoy greater international support (Chinkin 2001, 383). Coupled with this argument is the issue of the humanitarian consequences of sanctions. Sanctions that have egregious humanitarian costs quickly lose international support (Cortright and

Lopez 2002, 32). Although it remains a controversial issue, there are a number of authors who argue that the SC is bound by humanitarian and human rights law, and thus sanctions that have severe humanitarian consequences, or that infringe on customary international human rights, are unlawful. These authors Cortright and Lopez (2002) also point out that states implementing sanctions are often bound by treaty-based human rights regimes, such as the International Covenant on Economic, Social, and Cultural Rights, which arguably preclude adherence to this type of sanction.⁵

Intertwined with issues of legitimacy and human rights is the concept of sanctions fatigue. Over time sanctions, especially those that do not have a perceived end date, tend to lose objective clarity, legitimacy, and force. This is especially true when humanitarian concerns rise, the target adjusts, implementation costs grow higher, and the initial catalyst for the sanction fades in memory (Haass 1998, 205). To prevent sanctions fatigue from occurring, Richard Haass suggests that sanctions should be rapidly implemented and have a strong purpose (Haass 1998, 208).

Despite the recognition that sanction enforcement relies on the “goodwill of (the majority of) States Parties,” (Den Dekker 2001, 111) and the many suggestions of positive reinforcement (capacity building, economic assistance for affected states, etc.), academic scholarship has not yet fully grappled with the fact that the political interests of UN members may conflict with SC obligations, and thus compromise sanction regimes. While it is clear that all UN members are bound to implement SC resolutions, it has been noted that, in practice, the enforcement of sanctions regimes has varied according to national interests and political will (Hull 1997). Political differences among SC members can also negatively impact sanctions regimes even after the implementation of sanctions (Cortright and Lopez, 2002). In light of this, it seems evident that both ‘carrots’ and ‘sticks’ should be used to ensure that UN sanctions are enforced effectively.

THE SANCTIONS REGIME AGAINST IRAQ

The SC introduced comprehensive sanctions against Iraq on 6 August, 1990 with SCR 661 in response to Iraq’s takeover of Kuwait.⁶ SCR 670 further tightened these sanctions on 25 September, 1990 by imposing aviation and maritime restrictions to close loopholes in the regime. Following the First Gulf War, the SC passed SCR 687 on 3 April 1991, which established the cease-fire and also set up eight conditions to be met before sanctions would be lifted. In response to growing humanitarian concerns, the SC passed SCR 706 on 15 August 1991, which authorized the Oil-for-Food program, allowing the sale of up to \$1.6 billion of Iraqi oil over a six-month period, with the money going to a UN escrow account to finance humanitarian imports and war reparations. On 14 April 1995, the SC passed SCR 986 and created a new formula for the Oil-for-Food program that increased the amount of oil that could be sold to \$1 billion every three months. This resolution also shifted the responsibility for distributing humanitarian goods to the Iraqi government. Then, on 20 February 1998, the SC passed SCR 1153, which increased the amount of oil that could be sold to \$5.25 billion every six months. This resolution also allowed the oil revenues to finance critical development needs. SCR 1284, on 17 December

1999, established a new UN Monitoring, Verification, and Inspection Commission (UNMOVIC) to replace the UN Special Commission on Weapons (UNSCOM established by SCR 687). It also removed the cap on oil sales and offered to suspend sanctions for renewable 120-day periods if Iraq cooperated with UNMOVIC and IAEA. Later, in response to humanitarian concerns, the SC established a panel of experts to report on Iraq's humanitarian situation by passing SCR 1302 on 8 June, 2000. Passed on 1 June, 2001, SCR 1352 stated the SC intention to consider arrangements to facilitate civilian trade. The SC used SCR 1409, on 14 May, 2002, to introduce a new procedure to speed up imports. Finally, following the Second Gulf War, on 22 May, 2003, the SC lifted all non-military sanctions by passing SCR 1483. SCR 1483 also formally recognized Britain and the United States as occupying powers.

WEAKNESSES IN THE SANCTIONS REGIME

During the sanctions period there were major leaks in the sanctions regime. One of the main seepage points was Jordan, through which Iraqi exports were smuggled out and "strategic goods" were secretly imported (Naylor 1999, 323). When pressure from the United States forced Jordan to crack down on this illicit trade it was only driven further underground, but the Jordanian government was unable or unwilling to significantly disrupt the flow of smuggled goods (ibid.). Furthermore, Jordan stopped monitoring Iraqi-bound cargo at the port of Aqba in October 2000 (Cortright and Lopez 2002, 34). The lack of domestic capacity for sanction enforcement was an even more obvious issue in Lebanon. The central government was unable to control a number of ports within Lebanon and these militia-operated ports became major shipping points through which goods were smuggled in and out of Iraq. Even when official government control was established in those ports, sanction-busting continued (Naylor 1999, 323).

The Duelfer report estimates that between 1990 and 2003, Iraq was able to generate USD\$10.9 billion in hard currency through illicit means (Duelfer 2004). One commentator asserted that "everyone from the French to the Syrians has quite publicly violated the sanctions in some way" (Friedman 2000). The United States, the greatest proponent of the sanctions, has also been accused of being the largest violator, contributing fifty-two percent of the total kickbacks to the Iraqi government during the sanctions regime (Borger and Wilson 2005). This demonstrates that UN members need to build their capacity in material terms in some instances, such as in the case of Lebanon, and to create more effective monitoring and enforcement methods in other instances, such as in the American case, to ensure compliance with sanctions regimes.

Before the Gulf War, Iraq was a major economy in the Middle East, and many of its neighbors had economic motivations to oppose sanctions. Not surprisingly, a substantial number of nations sought consultation with the SC under Article 50 of the UN Charter. This appeal, the first instance of a broad invocation of Article 50 for economic relief, involved a total of twenty-three countries (Al-Khasawneh 2001, 326). Article 50, however, only gives a right to consultation, not a right to compensation for economic losses. In response to the appeals, the sanctions committee, which was formed by SCR 661 to monitor the sanctions against Iraq, called for international aid in order to lessen the impact

on affected third party states (Al-Khasawneh 2001, 329). To shore up support for the sanctions, some aid was given to front-line states. Under an Economic Action Plan proposed by the United States, highly-affected front-line countries, namely Egypt, Jordan, and Syria, were provided USD\$20 billion in aid, and Egypt also had USD\$7.1 billion of its debt forgiven by the United States (Melby 1998, 115). As well, the SC granted an exemption allowing Jordan to continue importing a small amount of oil from Iraq (Al-Khasawneh 2001, 331). However, these efforts did not fully offset the economic losses felt by these states. The International Monetary Fund estimated that Jordan lost around fifty-five percent of its GNP as a result of the embargo on Iraq and Kuwait (Melby 1998, 114). Turkey's pre-embargo export trade with Iraq was estimated at USD\$8 billion, and Iraq also owed Turkey USD\$800 million, payment of which was suspended when the embargo began (Melby 1998, 114). Egypt estimated its annual losses from the sanctions situation to be USD\$9 billion annually from a decline in tourism, Suez Canal transit fees, worker remittances from Kuwait and Iraq, and the loss of bank deposits in Iraq and Kuwait (Melby 1998, 114). Yemen was also affected by the sanctions regime because its support for Iraq resulted in lost remittances and aid from Saudi Arabia (Melby 1998, 114). Finally, prior to the embargo, forty percent of Lebanon's exports went to Iraq, and worker remittances were estimated at over USD\$150 million (Melby 1998, 114). Thus, despite the aid that was given and the appeals by the Sanctions Committee, the affected states did not receive full compensation for their losses. In fact, on 25 March 1991, a group of twenty-one countries, including Jordan, Lebanon, and Yemen, sent a joint letter to the President of the SC appealing for greater compensation (Security Council document S/22382 in Al-Khasawneh 2001, 330). Even though the SC provided some compensation to affected third-party states for their losses, which it was not necessarily obligated to do, the extent of the economic impact was such that it is not entirely surprising that a number of these countries continued to trade illicitly with Iraq during the sanctions regime.

Initially, sanctions against Iraq had great legitimacy. Boutros Boutros-Ghali, the UN Secretary General during the initial phases of the sanctions regime, noted that it was the first time since the founding of the UN that one member state had attempted to annex another, and that no other crisis in the UN's history had "elicited such attention and action from the Security Council in such a compressed span of time" (Niblock 2001, 97). The goals of SCR 660 and SCR 661 were to force a return to the status quo; Iraq was to withdraw its military forces back to prewar positions and the legitimate Kuwaiti government was to return to authority. These objectives were clearly defined and directed towards remedying a stark violation of international law.

The objective clarity of these resolutions began to erode with SCR 678, which authorized states acting in cooperation with the government of Kuwait to use all means necessary to implement SCR 660 and subsequent related resolutions, and "to restore international peace and security in the area." While the objectives of SCR 660 and 661 were met in March 1991, the SC passed SCR 687 soon after, which maintained the sanctions regime. SCR 687's objectives included the creation of a Middle East nuclear weapons-free zone, and stated

that international peace and stability in the area, as required by SCR 678, had not been restored. The creation of a Middle East nuclear weapons-free zone and the goal of attaining regional peace and stability are regional management issues that could not be solved by sanctioning Iraq alone. Furthermore, Iraq's obligations under SCR 687, such as the requirement that "Iraq undertakes not to use, develop, construct, or acquire" weapons of mass destruction (WMD) in Paragraph 10, introduced even greater ambiguity. This requirement presented potential problems regarding the subjective opinions of a permanent monitoring body. Fears that the monitoring body would be politicized and that the sanctions regime would become permanent further undermined their legitimacy. The goals of the sanctions have been labeled as "new and vague, indeed unattainable, objectives" (Al-Anbari 2001, 374). The Iraqi government attacked the legitimacy of the sanctions on the grounds that Iraq's sovereignty was being violated through the imposed settlement of the border dispute with Kuwait, and the denial of Iraq's "lawful right to acquire weapons ... for defense" (United Nations, Letter from the Minister of Foreign Affairs of Iraq to the President of the Security Council 1991, cited in Niblock 2001, 103). Iraq's attacks on the legitimacy of the sanctions continued throughout the sanctions period. By 1998, international support for the sanctions had weakened to the point that French Foreign Minister Hubert Védrine called them "cruel, ineffective and dangerous" (Global Policy Forum 2002, Ch.3). The argument that compliance by implementing states will weaken if sanctions are not perceived to be legitimate and fair, especially if the sanction duration is considered to be excessive, seems to be supported by the experience of the sanctions regime against Iraq (Chinkin 2001, 383).

Egregious humanitarian conditions caused by sanctions are another factor that can undermine sanctions regimes. Rather than being targeted in nature, the initial sanctions against Iraq were comprehensive, cutting economic activity to all citizens of the country, despite the fact that SCR 661 did allow "payments exclusively for strictly medical or humanitarian purposes, and, in humanitarian circumstances, foodstuffs." In practice, it was up to the Sanctions Committee to determine what constituted humanitarian circumstances and purposes, and this approach did not prevent the Iraqi people's decline to extreme misery and despair (Burci 2001, 146-147). The SC was aware of the growing humanitarian crisis and included in SCR 687, Paragraph 20, a clause that once again allowed food and other goods that were required to meet civilian needs to be traded without limitation. The Sanctions Committee, however, still had to give its approval to transactions involving "material and supplies essential for civilian needs" under a streamlined process. Humanitarian concerns were also evident in SCR 688, which expressed explicit concern for the plight of the Kurdish population, but also included general appeals for humanitarian aid and concern for the human rights of all Iraqis in Paragraphs 2, 5, and 6. Despite this, the appeals were not met, cooperation between Iraq and the UN agencies was tenuous, and Iraq was still under an oil embargo. The end result was a situation where humanitarian problems were not being addressed by external actors and the Iraqi government was unable to use its oil resources to alleviate its population's suffering (Niblock 2001, 47).⁷

The SC, taking note of the continuing humanitarian crisis, adjusted the sanctions regime once again. The much maligned Oil-for-Food program that was part of SCR 706 in 1991, addressed these growing concerns. Paragraphs 1 through 4 authorized Iraq to export up to USD\$1.6 billion in oil over six months, with the revenues going into a UN escrow account in order to finance war reparation, UN disarmament and monitoring operations, and humanitarian programs. The Iraqi government's resistance to SCR 706, 712, and other resolutions, however, prevented the implementation of the Oil-for-Food program (and shifted some of the responsibility for the humanitarian crisis to the Iraqi government) until 1996. In 1996, the adjusted program under SCR 986, which authorized the sale of USD\$1 billion every three months, was finally allowed to begin (United Nations Operation of the Iraq Program, n.d.). The Oil-for-Food resolutions, in response to ever increasing opposition to the sanctions over humanitarian concerns, increasingly relaxed the cap on the sales and purpose of the revenue until SCR 1284 (1999), Paragraph 15, authorized oil sales without a cap, and SCR 1352 (2001), Paragraph 2, expressed intent to explore new methods of civilian trade with Iraq. Despite this movement on the part of the SC, Kofi Annan, the Secretary-General of the United Nations, noted in 2000 that:

The United Nations has always been on the side of the vulnerable and the weak, and has always sought to relieve suffering, yet here we are accused of causing suffering to an entire population. We are in danger of losing the argument, or the propaganda war ... about who is responsible for this situation in Iraq - President Saddam Hussein or the United Nations (United Nations Press Release 2000).

International concerns over the situation of the Iraqi people did have an impact on the sanctions regime. On 10 March 2003, an updated issue brief for the United States Congress was circulated. The brief linked America's difficulty in maintaining international support for the sanctions to international humanitarian concerns, and suggested targeted sanctions (to narrow the impact) as a way of stopping "sanctions erosion" (Katzman 2003). The resignation of UN humanitarian directors Denis Halliday, UN coordinator for Iraq in September 1998, and Hans Von Sponeck, director of the Oil-for-Food program in March 2000, in protest of the conditions in Iraq fueled anti-sanctions protests and campaigns in many countries, including Canada, France, the United States, Britain, and various Arab states (Cortright and Lopez 2002, 32). By the end of 2000, some protest groups, such as Voices in the Wilderness, shifted from demonstrating against sanctions to delivering unauthorized aid. A significant number of flights carrying relief supplies and businessmen landed in Baghdad in open defiance of the sanctions during the second half of 2000 (Cortright and Lopez 2002, 32).

Support for sanctions regimes is also contingent on state interests. As relations in the Middle East shifted, so did states' support for the sanctions regime. Problems with state enforcement due to lack of capacity has been previously discussed, but a number of states have been implicated as either purposefully ignoring sanctions-busters or even profiting from these violations. For example, there are reports that Turkey was actually regulating and taxing

illicit importation of Iraqi oil worth USD\$400 million annually (Katzman 2003, 12). In a move that undermined the spirit, if not the direct letter of the sanctions, Syria and Iraq signed a free trade agreement in 2001.⁸ The Iranian government was accused, by the Commander of the United States Fifth Fleet Vice Admiral Thomas B. Fargo, of being directly involved in sanctions violations in 1997 (*New York Times* February 12, 1997). Despite American laws that require sanctions against countries that violate UN sanctions, Jordan a “key ally in the region,” continued to receive American aid, and U.S. sanctions were also waived against Turkey (Katzman 2003, 12).

Political divisions within the SC also had a negative impact on the sanctions regime. The United States and Britain were ardent supporters of the regime, but as the duration of sanctions increased, so did opposition from Russia, China, and France. In particular, Russian proposals to ease and eventually lift the sanctions were in opposition to British and American plans, and this division hampered the capacity of the SC to act in a timely and effective fashion (Cortright and Lopez 2002, 40-41). Moreover, it was suggested that Russia was motivated by significant economic and national interests in exploiting Iraqi oil fields (Friedman 2000). Similar accusations have been made against the United States government for its initial “tacit approval” of violations by American businesses (Borger and Wilson 2005). The politicization of the SC in this manner undermines both the legitimacy of UN sanctions and the SC itself.

LESSONS FROM IRAQ

Policy makers looking to maintain effective sanctions regimes can draw important lessons from the Iraq experience. Domestic capacity does matter in enforcement. In some cases, actual capabilities to comply need to be heightened. At least some of the illicit trade that went through Lebanon and Jordan could have been prevented with stronger domestic capabilities in those countries. However, other instances of sanctions-busting in Iraq, such as those involving American businesses, cannot be explained in terms of material incapacity. For effective sanctions regimes, uniform implementation willingness and capacity is required, as even a small number of sanction-busters can undermine the whole regime (Biersteker et al. 2005, 62). To remedy problems of capability, participating states lacking implementation capacity (legal, judicial, administrative, etc.) need to be identified early in the sanctions regime and should be offered assistance to undertake sustained capacity building initiatives in order to create the necessary implementation mechanisms (Biersteker et al. 2005, 63). These efforts could take the form of financial assistance, but also require the transfer of technical knowledge and technology in a more systematic and sustained fashion. These states may also require enabling legislation, such as the criminalization of sanctions-busting, to harmonize international legal obligations and domestic law (Ward 2005, 169).

Some countries – such as the United States – still experienced sanctions violations despite the presence of a strong resource capability and the will to pursue compliance with the regimes. In these instances, more effective national monitoring bodies need to be created. Enhanced cooperation across different branches of government and better dissemination of the requirements of

sanctions to affected private actors is required (Biersteker et al. 2005, 59). In this manner, the situation with the American company Bayoil, which reportedly had tacit approval for its actions from the U.S. State Department yet was still charged with sanctions violations at a later date, could be avoided (Borger and Wilson 2005).

The Iraq experience also demonstrated the importance of addressing the economic needs of affected non-target states. Many of the key smuggling routes involved countries that had previously enjoyed strong economic relationships with Iraq, in particular Jordan, Turkey, and Lebanon. Although causation has not been definitively proven, these routes strongly suggest that smuggling was a way for the most affected states to alleviate the economic hardship caused by the sanctions regime. Defections from sanctions regimes over an issue like this could be combated by creating a system of compensation for affected non-target states (Hull 1997). Proposals for a permanent trust fund to compensate states inadvertently injured by UN sanctions have been controversial, however, with developed states expressing concern over funding and fund disbursement (Hull 1997). An acceptable halfway step may be to create an objective body charged with determining which states have been affected and to what degree by the implementation of sanctions, without making binding compensation requirements.

Concerns over the legitimacy of the sanctions against Iraq increased with the length of time the sanctions remained in existence, and the duration of the sanctions reinforced the perception that the sanctions regime was politicized. The perception, following the First Gulf War, that sanctions were no longer just (and therefore illegitimate) reinforced international humanitarian concern over the fate of the Iraqi people. According to many critical observers, the comprehensive sanctions were punishing the wrong people - the Iraqi civilians (Niblock 2001, 194). These issues fueled anti-sanction campaigns, contributed to sanctions-busting at the private level, and ultimately undermined support and effective enforcement at the state level (Cortright and Lopez 2002, 32). To maintain the legitimacy of a sanctions regime, implementing actors need to ensure that it does not violate other principles of international law (human rights law, for example), and also that it is proportional, targeted, and can be periodically assessed (Clapham 2001, 135). Many of these requirements have indeed been written into SCRs. The era of comprehensive economic sanctions seems to have passed, as witnessed by recent targeted sanctions against the Sudan by SCR 1591 (2005), North Korea by SCR 1718 (2006), and Iran by SCR 1747 (2007). Targeted sanctions potentially solve a number of legitimacy issues, such as harmonization with principles of international law and proportionality. Although targeted sanctions do not have the same degree of impact as comprehensive sanctions, they can be effective. North Korea's strong reactions to SCR 1718, which banned the trade of luxury goods to North Korea as well as items that could contribute to its weapons of mass destruction programs, and American financial sanctions against North Korea, which undercut North Korea's ability to illicitly gain hard currency, aptly demonstrate this (Porteus 2007). Because the sanctions effectively constrained Kim Jong-Il's ability to reward and thus maintain loyalty among elite level North Korean

leadership, it is not surprising that North Korea speedily returned to the Six-Party Talks. Nonetheless the SC should, before imposing sanctions, ensure consensus on the type of sanctions to be imposed, assess proportionality and possible unintended consequences, streamline the process for humanitarian exemptions, and create mechanisms for reviewing and lifting sanctions, either through sunset clauses or “negative veto” clauses.

State interests, despite the lack of discussion in the literature, did have a major impact on the Iraqi sanctions regime. Enforcement by Iran, Turkey, and other nations was at times suspect due to these countries’ national interests, while political divisions among the permanent five on the SC also contributed to the erosion of sanctions. One way to minimize the undermining effect of shifts in political support for sanctions would be to ensure that there is widespread initial support for the sanctions, and then allow periodic review of the regime to account for shifts. To solve the problem of maintaining sanctions in the face of differing political interests, positive inducements, such as economic aid as allowed for under Article 50 of the UN Charter, could be offered and sanctions violators could be punished. This did not happen in the case of Iraq, although in 2001, SCR 1343 did impose sanctions against Liberia for violating sanctions, creating a precedent for such action (Cortright and Lopez 2002, 203). For effective enforcement of sanctions regimes it therefore seems prudent to have both positive and negative tools for shifting state interests.

CONCLUSION

Policy makers seeking more effective ways to maintain current sanctions regimes, such as those in effect at the time of writing against North Korea, Iran, Liberia, the Sudan, etc., can learn much from Iraq. To address problems of capacity, developed countries should assist to systematically build capacity in those countries which lack required enforcement mechanisms, while countries with adequate resources should take measures to ensure effective self-monitoring. To reduce non-enforcement of sanctions regimes that arises from collateral economic losses, the SC must recognize the negative impact on affected countries and intensify efforts to provide compensation. The use of targeted sanctions with defined but renewable time frames should reduce humanitarian consequences, allow greater proportionality, and increase the legitimacy of future sanctions regimes. In the future, efforts should be made to shift state interests through positive inducements to build and maintain effective sanctions regimes, but the use of secondary sanctions should not be ruled out as a possibility for ensuring compliance with the regime.

NOTES

¹ This viewpoint has been challenged. Some scholars argue that sanctions have more harmful effects on civilian populations than military action. For an example of this argument, see Al-Anbari 2001.

² Notable works in this literature include David Cortright and George A. Lopez’s 2002 work *Sanctions and the Search for Security: Challenges to UN Action*; Richard N. Haass’ 1998 volume *Economic Sanctions and American Diplomacy*; Peter Wallensteen and Carina Staibano’s 2005 volume *International Sanctions: Between Words and Wars in the Global System*; and Tim Niblock’s

2001 study of sanctions in the Middle East, “Pariah States” and *Sanctions in the Middle East: Iraq, Libya, Sudan*.

³ Proof that UN sanctions are not abating is found in the fact that, according to the United Nations Sanctions Committee website (United Nations Security Council, n.d.) in June 2007, there were twelve sanctions regimes in existence and being monitored by sanctions committees. Furthermore, eight of the sanctions committees were established since 2003.

⁴ Most mainstream international law literature assumes that international law constrains state actions, either by providing the “rules of the game” (in a liberal paradigm), or in a norm building iterative process shaping both law and state interests (in a constructivist paradigm). There are, however some who argue that international law does not constrain state actions (Goldsmith and Posner 2005).

⁵ There are numerous examples of observers who take up this argument in its various forms, for a brief introduction, see Andrew Clapham 2001; Marcos Sassòli 2001; and Iain Cameron 2005.

⁶ The SCRs listed in this section are some of the more noteworthy resolutions, but the list is not exhaustive. A more complete list of sanctions against Iraq can be found at the website UN Security Council Sanctions Relating to Iraq <http://www.casi.org.uk/info/scriraq.html>.

⁷ However, it should also be noted that the Iraqi government would not have necessarily used oil resources to alleviate civilian suffering. Observers accused the Iraqi government of misusing resources gained from the Oil-for-Food program for personal gain and propaganda against the United States and the sanctions regime (Norton-Taylor 1999, and U.S. Department of State 1999).

⁸ To get around the accusations of violating the sanctions regime, Syria claimed that trade would only be in permitted goods (Arabic News.com February 2, 2001).

REFERENCES

- Al-Khasawneh, Bisher H. 2001. Consultation under Article 50 of the United Nations Charter: The Experience of the Middle East. In *United Nations Sanctions and International Law*. Ed. Vera Gowlland-Debbas, 325-334. The Hague: Kluwer Law International.
- Al-Anbari, A. Amir. 2001. The Impact of United Nations Sanctions on Economic Development, Human Rights and Civil Society. In *United Nations Sanctions and International Law*, Ed. Vera Gowlland-Debbas, 371-380. The Hague: Kluwer Law International.
- Arabic News.com*. 2001. Iraq Approved the Free Trade Agreement with Syria. February 2. <http://www.arabicnews.com/ansub/Daily/Day/010202/2001020218.html>. (Accessed March 27, 2007).
- Biersteker, Thomas J. et al. 2005. Targeted Sanctions and State Capacity: Towards a Framework for National Level Implementation. In *International Sanctions: Between Words and Wars in the Global System*. Eds. Peter Wallensteen and Carina Staibano, 57-64. New York: Frank Cass.
- Borger, Julian, and Jaimie Wilson. 2005. US ‘backed ill Iraqi oil deals. *The Guardian Unlimited*, May 17. <http://www.guardian.co.uk/international/story/0,3604,1485546,00.html>. (Accessed March 10, 2007.)
- Cameron, Iain. 2005. Protecting Legal Rights: On the (in)security of Targeted Sanctions. In *International Sanctions: Between Words and Wars in the Global System*. Eds. Peter Wallensteen and Carina Staibano, 181-206. New York: Frank Cass.
- Chinkin, Christine. 2001. Alternatives to Economic Sanctions. In *United Nations Sanctions and International Law*. Ed. Vera Gowlland-Debbas, 381-392. The Hague: Kluwer Law International.
- Clapham, Andrew. 2001. Sanctions and Economic, Social and Cultural Rights. In *United Nations Sanctions and International Law*. Ed. Vera Gowlland-Debbas, 131-141. The Hague: Kluwer Law International.
- Cortright, David, and George Lopez. 2005. A Sanctions Coordinator: Options for enhancing compliance. In *International Sanctions: Between Words and Wars in the Global System*. Eds. Peter Wallensteen and Carina Staibano, 65-74. New York: Frank Cass.
- Cortright, David, and George A. Lopez, Eds. 2002. *Sanctions and the Search for Security: Challenges to UN Action*. London: Lynne Rienner Publishers.
- Den Dekker, Guido. 2001. *The Law of Arms Control: International Supervision and Enforcement*. The Hague: Martinus Nijhoff Publishers.
- Duelfer, Charles. 2004. Comprehensive Report of the Special Advisor to the DCI on Iraq’s WMD. https://www.cia.gov/cia/reports/iraq_wmd_2004/chap2.html. (Accessed 25 March 2007.)
- Friedman, Thomas. 2000. Meanwhile, Back in Iraq. *Global Policy Forum*. November 20.

- <http://www.globalpolicy.org/security/sanction/iraq1/council/russ0011.htm>. (Accessed March 10, 2007.)
- Global Policy Forum. 2002. Iraq Sanctions Humanitarian Implications and Options for the Future. August 6. <http://www.globalpolicy.org/security/sanction/iraq1/2002/paper.htm> (Accessed May 10, 2007.)
- Goldsmith, Jack L. and Eric A. Posner. 2005. *The Limits of International Law*. New York: Oxford University Press.
- Haass, Richard N. 1998. Conclusion: Lessons and Recommendations. In *Economic Sanctions and American Diplomacy*. Ed. Richard N. Haass, 197-212. New York: Council on Foreign Relations Book.
- Hull, Ricard E. 1997. Imposing International Sanctions: Legal Aspects and Enforcement by the Military. Washington: National Defense University Press.
<http://www.ndu.edu/inss/books/Books%20-%201997/Imposing%20International%20Sanctions%20-%20March%2097/contents.html>. (Accessed March 10, 2007.)
- Katzman, Kenneth. 2003. Iraq: Weapons Threat, Compliance, Sanctions and U.S. Policy. *Congressional Research Service*, March 10.
<http://fpc.state.gov/documents/organization/19205.pdf>. (Accessed March 15, 2007.)
- Melby, Eric D.K. 1998. Iraq. In *Economic Sanctions and American Diplomacy*. Ed. Richard N. Haass, 107-128. New York: Council on Foreign Relations Book.
- Naylor, R.T. 1999. *Patriots and Profiteers: On Economic Warfare, Embargo Busting and State-Sponsored Crime* Toronto: McClelland & Stewart Inc.
- The New York Times*. 1997. U.S. Admiral Says Iraq, Aided by Iran, Is Selling Oil Despite Ban. February 12. <http://select.nytimes.com/>. (Accessed on March 27, 2007.)
- Niblock, Tim. 2001. *"Pariah States" and Sanctions in the Middle East: Iraq, Libya, Sudan*. London: Lynne Rienner Publishers.
- Norton-Taylor, Richard. 1999. Misguided Weapons - Sanctions may be an Easy Option but they do not Affect the Leaders they aim to Weaken. *The Guardian*. August 16.
<http://www.globalpolicy.org/security/sanction/sanct2.htm>. (Accessed May 10, 2007.)
- Porteus, Sam. 2007. Smartening Up. *The Standard Finance*. March 26.
http://www.finance.thestandard.com.hk/en/columnists_today.asp?aid=57 (Accessed May 10, 2007.)
- Sassòli, Marcos. 2001. Sanctions and International Humanitarian Law - Commentary. In *United Nations Sanctions and International Law*. Ed. Vera Gowlland-Debbas, 241-248. The Hague: Kluwer Law International.
- United Nations. 1945. *Charter of the United Nations*.
- . 2000. United Nations Press Release, SC/6834. March 24.
<http://www.fas.org/news/iraq/2000/03/20000324-sgsm7338.htm>. (Accessed on 27 March 2007.)
- United Nations Office of the Iraq Program. n.d. *Implementation of Oil-for-Food - A Chronology*.
<http://www.un.org/Depts/oip/background/chron.html>. (Accessed March 27, 2007.)
- United Nations Security Council. n.d. United Nations Security Council Sanctions Committees.
<http://www.un.org/sc/committees/index.shtml>. (Accessed June 6, 2007.)
- U.S. Department of State. 1999. Saddam Hussein's Iraq. September 13.
<http://www.fas.org/news/iraq/2000/02/iraq99.htm>. (Accessed May 10, 2007.)
- Ward, Curtis A. 2005. The Counter-Terrorism Committee: Its relevance for Implementing Targeted Sanctions. In *International Sanctions: Between Words and Wars in the Global System*. Eds. Peter Wallensteen and Carina Staibano, 167-180. New York: Frank Cass.

RUSSIA'S FOREIGN POLICY ACE

Jeremy D. Gordon

Maxwell School of Citizenship and Public Affairs, Syracuse University

As the world's largest gas company with close political and financial ties to the Kremlin, Gazprom has served, and continues to serve not simply as a profit-driven energy company, but also as a strategic tool through which Russia seeks to implement its foreign policy. By employing a mixture of tactics that are both covert and overt in nature, Gazprom has consistently and often effectively served as an asymmetrical instrument of force that the Kremlin has turned to time and again. Since the dissolution of the Soviet Union, Russia has manipulated Gazprom to further two of its most important foreign policy goals – ensuring the rights of Russian communities in the Diaspora and maintaining control over the former soviet republics; regions that Russia refers to as its 'near abroad'. With a deteriorating military and a sense of nostalgia for its old glory, Russia has chosen to adopt this alternative, asymmetric and certainly unconventional strategy in order to maintain its power and influence in a post-Soviet world.

INTRODUCTION

“Gazprom is a powerful political and economic lever of influence over the rest of the world.” Speaking to an audience of businessmen, politicians and shareholders at the 10th anniversary celebration for Gazprom, Russian President Vladimir Putin was paying a flattering tribute to Russia's largest and wealthiest gas company. The significance of his statement, however, touches upon a strategy that stretches far beyond the innocuous compliment that some observers may have taken at face value. Russia's president was essentially describing the core of a very clear and deliberate new aspect of Russia's post-Soviet foreign policy – the use of Gazprom as a strategic instrument for asserting its influence over other countries. As a state-controlled company with tightly woven political links to Russia's executive branch, Gazprom wears two masks: one as a utility company, the other as an asymmetrical instrument used by the Kremlin for political gain. Employing methods that are both covert, such as the acquisition of strategic industries in foreign countries, as well as overt, such as dramatic price hikes and cutting foreign supplies, the Kremlin has used the Russian gas giant to re-assert its position as a dominant regional and world player in a post-Soviet world.

GAZPROM'S DOMINANCE

Since its formal inception over a decade ago, Gazprom has steadily become a monolith in the Russian and world gas markets. Founded by Presidential Decree as the Russian joint-stock company Gazprom (RAO Gazprom) in November 1992 (Gazprom 2006a), today it is the largest gas-producing company in the world and enjoys a monopoly position over the production, refinement and transportation of gas in a country that possesses the world's largest gas reserves (Energy Information Agency, 2007). Gazprom's position in the Russian gas

market is so dominant, however, that it would be more aptly characterized as authoritarian. In the words of one analyst, the gas sector “works largely like a centrally planned Soviet market, where Gazprom takes the role of the Ministry of Gas” (Larsson 2006, 30). A 2004 report by the Organization for Economic Co-operation and Development (OECD) elaborates upon Gazprom’s dominant grip on the market. In examining the fact that Gazprom maintains control over all data on exports and pipeline capacity, regulates the entire process of gas rationing, and is shielded in many ways from both private and public oversight, the report ultimately comes to the marked conclusion that a Russian gas ‘market’ (in terms of competitiveness and capitalist forces) essentially does not exist (Organization for Economic Cooperation and Development 2004).

With Gazprom exerting such dominant control in the Russian gas market, the few competitors that do exist, such as Itera and Eural Trans Gaz, face substantial difficulties in maintaining their independence and financial integrity. With Gazprom in control of 90 percent of all of Russia’s gas and complete control of the pipeline transportation network, any attempt to transit gas within or across Russian territory is only granted with the consent of Gazprom (Larsson 2006). This fact is underscored by the limited ability of Central Asian countries to transport their vast gas resources to the profitable markets in Western Europe without an agreement from Gazprom to use Russia’s pipeline network (Blagov 2006). As long as Gazprom maintains control over the transaction and transport of gas across its territory, each of these companies can be nothing more than “a de facto subsidiary of Gazprom” (Larsson 2006, 149).

Yet the relationship between Gazprom and one of its regional competitors, Itera, runs deeper than just the formal agreement to transport gas. While Gazprom supporters would reference the results of two 2001 investigations¹ that looked into the uncertain relationship between Itera and Gazprom and produced no evidence of wrongdoing, the facts to the contrary are significant. Most of these shadowy connections are described in an article written in 2002 by the Carnegie Endowment, which outlines tax scams, shady second hand gas transactions and the manipulation of gas prices for windfall profits (McGregor 2002).

Gazprom’s dealings with another regional competitor, Eural Trans Gaz, are no less contentious. An August, 2003 article in *Ukraina Kryminalna*, a Radio Free/Ukraine Service publication, is one of several publications implicating Eural Trans Gaz in multi-million dollar speculative deals with Gazprom² (International Helsinki Federation of Human Rights 2006). Gazprom’s entrenched position in the Russian gas market is undeniable. By maintaining its control over the pipeline network, the vast majority of Russia’s gas resources, all transport and reserve information relating to the Russian gas sector, in addition to a suspiciously close relationship with two of its regional gas competitors, Gazprom maintains a position of unparalleled and resounding strength as it sets out to achieve company goals.

GAZPROM AND THE STATE

As Gazprom has expanded its power and influence within the Russian market and abroad, the Kremlin has found an especially close relationship with this energy monolith to be in its own best interests. Besides staking its claim to a 51 percent share in the company, the Russian government has also passed the Gas Supply Law that limits foreign shareholdings in Gazprom's capital to no more than 20 percent (Stern 2005).

The ties between the executive branch and Gazprom, however, run much deeper than majority ownership and restrictions on foreign possessions. The extent of President Putin's personal connections with the Gazprom leadership, for example, is striking. Appointed in 2001 as the Chief Executive of Gazprom, Alexei Miller is one of Putin's closest confidants and an old friend since Putin's days as a city official in St. Petersburg. With three meetings per week scheduled with the President, Miller has the benefit of unparalleled access to Russia's top official (Larsson 2006, 117). The connections between Gazprom and the Kremlin, however, are not limited to this single relationship. Besides Miller, eight out of nineteen of the new management committee members can be linked to Putin (Stern 2005).

Further exploration of Gazprom's Board of Directors also exposes links between the company and the state. Out of eleven members of the Board, two are current ministers (German Gref – Min. of Economic Development and Trade; Victor Khristenko – Min. for Industry and Energy), one is the Head of the Presidential Administration (Dmitri Medvedev), one is an ex-minister (Farit Gazizullin – Min. for Property Relations), and one is the Special Envoy to the President for International Energy Cooperation (Igor Yusufov) (Gazprom 2006b). It is also noteworthy that Russia's current Ambassador to Ukraine, Victor Chernomyrdin, was the former CEO of Gazprom. Ukraine is an especially important transit country for Russian gas, as its pipelines transport 90 percent of Gazprom's exports to West and Central Europe (BBC 2006a). With the establishment of such close political ties to the Kremlin, Gazprom "became very much a part of the state...and accepted its role as an instrument of government policy both domestically and internationally" (Stern 2005: 173).

As a result of its strong political ties to Gazprom, the Kremlin has the ability to use the company as a viable tool for exerting its influence over other countries, especially those that are dependent on Russian energy supplies. This strategic use of energy resources as a political tool is directly related to Russia's post-Soviet foreign policy.

RUSSIA'S POST-SOVET FOREIGN POLICY

Despite its relative weakness following the collapse of the Soviet Union, within the last decade Russia's rising influence has resulted in a foreign policy that has been typically aggressive and often coercive. Having relinquished its superpower status at the end of the Cold War, Russia retains a strong desire to restore its position as a respected and dominant player on the world stage. As part of this attempt to reach the heights of its former power, Russia has asserted an exclusive claim to its sphere of influence, a region that is often referred to as its 'near

abroad' and includes the fourteen states that were once part of the Soviet Union in Central Asia, the Caucasus and Eastern Europe.³ Russia has taken steps to retain a significant degree of authority over the affairs of these countries – some of which (specifically Ukraine and Belarus) are even viewed as an inseparable part of Russia itself (Russell 2005).

Coupled with this Russian Monroe Doctrine⁴ is the need to protect Russians living abroad. According to Presidential National Security Advisor Yuri Baturin, Russia has “a duty as a great power to protect Russians in the near abroad” (Russell 2005, 66). Thus any attempts by an outside power to assert its influence in these areas, or any anti-Russian/pro-Western shift in the foreign policies of these near-abroad countries will be interpreted as an unfriendly policy directed against the Russian State itself. Likewise, any infringement on the perceived ‘rights’ of Russian minorities living abroad will also be perceived as a direct affront to Russia.

One of the principal strategies for implementing this assertive foreign policy is the aggressive use of energy resources. According to one research analyst at the Swedish Defense Research Agency, energy is “a short cut for Russia in its international relations and attempts to be treated with respect. This is one of the most important reasons why energy exports are politically important to Moscow” (Larsson 2006, 50). This politicization of energy is further underscored by official state documents. In the early months of 2003, Russia updated its official ‘Energy Strategy’. Embedded in this document, “in which the foreign policy significance of the energy export [is] elaborated,” is the notion that energy resources are vital not only to the nation’s economic prosperity, but also to its national security (Fredholm 2005, 13). The “militaristic” tone that the document takes is therefore not entirely surprising (Larsson 2006, 65). The Strategy states that energy will be used as a powerful defensive force against possible blackmail as well as an offensive security tool that will be used to prevent against any greater geopolitical threats. Given Gazprom’s ubiquitous influence in the Russian gas sector, its extensive political connections to the Kremlin, and the strategic role of energy in Russia’s foreign policy, Russia has strategically made frequent use of Gazprom as a political instrument.

GAZPROM’S IMPERIALISM

Gazprom has taken a two-pronged approach to gaining leverage over other countries – one is covert and the other is more explicit. The more concealed strategy is one that is premised upon a doctrine that was first published in *Nezavisimaya Gazeta* in early October, 2003 by the prominent Russian statesman and CEO of the Russian electricity giant United Energy Systems, Mr. Anatoly Chubais. The Chubais policy of “liberal imperialism” is guided by the principle of establishing Russia’s dominant influence not by coercion or military occupation, but rather by strategically using state-controlled companies, such as Gazprom, to take possession of foreign industries that are necessary to people’s daily lives. Once these state-controlled companies have become sufficiently integrated into the markets in other countries, the Kremlin will be able to assert

its influence by manipulating the prices and distribution of necessary products through Russian companies. In the energy sector, one of the main ways in which Russia has sought to assert itself “has been to make use of dependence on Russian gas and the growing participation of Russian capital in the privatization of key sectors of these states’ economies” (Grachev 2005, 266).

In line with this aggressive economic tactic, Gazprom has targeted for purchase a host of gas pipeline, distribution and extraction companies in its near abroad. Yet the notion of an energy company expanding to acquire new assets in other countries is not a new or even threatening concept. What makes Gazprom stand out from other energy companies who buy foreign assets has been the forceful and selective manner in which it has chosen to acquire and operate these properties, as exemplified below. As a state-owned company with especially close links to the executive branch, furthermore, Gazprom’s manipulative tactics can be viewed as an extension of the Kremlin’s politics.

Gazprom’s assets in the energy sectors of other countries, especially in the former Soviet Union (FSU), are extensive. In Kazakhstan it holds a 50 percent stake in the gas distribution company ZAO Kazrogaz; in Kyrgyzstan it owns over 85 percent of the gas and oil exploitation company AO Kyrgyzneftegaz; it owns 45 percent of the Armenian gas exploitation company ZAO Armrosgazprom; in Ukraine it holds a 50 percent stake in the gas pipeline operator SP Rosukrenergo (Larsson 2006). In Lithuania, Gazprom owns 40 percent of the gas distribution company Lieutuvos Dujos and in 2003 also acquired the Kaunas power plant. Gazprom owns 40 percent of Estonia’s gas distribution company Eesti Gaz and holds a 50 percent + 1 share stake in Moldova’s gas pipeline operator AO Moldovagaz (Fredholm 2005).

Clearly Gazprom has established a diverse portfolio of foreign assets that allow for ample leverage in the energy sectors of these countries. The terms leading to the acquisition of these and other potential properties have often had links to political issues and goals that have served to further strengthen the Kremlin’s influence. Gazprom’s tactics have included “taking advantage of its customers by allowing them to run up debts that could only be exchanged for equity in infrastructure and energy companies” (Fredholm 2005, 18), raising or threatening to raise prices at times that coincide with political summits, as well as providing favorable treatment for countries or regions with closer diplomatic ties to the Kremlin.

Gazprom’s dealings with Moldova serve as an important example of such strong-arm tactics. Upon realizing that it would be unable to pay off its energy debts in 1998, the Moldovan government was under pressure to come up with other ways to compensate Gazprom for its overdue payments. After much deliberation, the government resolved that the only form of payment besides cash that would adequately satisfy Gazprom was to sell off certain strategic energy assets. In October, the Moldovan government reluctantly approved of a plan that would turn over the entire country’s gas supply system to Gazprom (The Jamestown Foundation 1998).

The timing of this turnover was especially significant. Though Moldova’s debt had been accumulating since 1991, Gazprom’s demand for payment coincided with discussions between Moscow and Chisinau on the withdrawal of

Russian troops from the breakaway region Transdnestria (Socor 2006). In rejecting Moldova's call for the removal of all foreign troops from its territories, Russia subsequently made additional demands to raise the price of gas and to gain control of strategic gas properties in order to force a favorable outcome regarding the matter of troop withdrawal.

Russia has raised the price of gas in order to forcefully acquire strategic properties in the markets of other countries as well. In Belarus, despite its especially close relationship with Moscow since the breakup of the Soviet Union, Gazprom has relentlessly pursued ownership of the pipeline network. As the transit corridor for twenty percent of all Russian gas transported to Europe, possession of this pipeline network would free Gazprom from any third party intervention in the transit of its gas. Ownership would also provide the company with further room to maneuver in Belarus' gas sector. In December 2006, during the heart of Belarus' frigid winter season, Gazprom suddenly threatened to raise the price of gas in Belarus from \$47 per thousand cubic meters (tcm) to a relatively much more expensive price of \$105/tcm (BBC 2006b). Unless Belarus ceded control of its distribution network, which is operated by the state-owned company Beltransgaz, Gazprom threatened to increase the price of gas even further to \$200/tcm (Voice of America 2006). It was not until the day their gas contract was set to expire, January 1, 2007, that Belarus and Gazprom agreed to a new deal. In order to stave off even higher prices during the frigid winter, Minsk was forced to sell a fifty percent stake in Beltransgaz, and will pay \$100/tcm in the first year of a steadily increasing five-year deal (The Economist, 2007). In addition to the transit pipelines, Gazprom now controls all of the local pipelines within Belarus.

Russia has applied these same tactics in Armenia as well. Despite agreeing to pay nearly double the price for its gas import needs, Armenia has signed a deal with Gazprom that sets a price of \$110/tcm through 2008 (Mosnews 2006) – a figure that is still less than half of the \$250/tcm average that is being charged to many European countries. In order to achieve this relatively low price, however, the government in Yerevan agreed to hand over control of significant energy properties. These include ownership of the Armenian section of a planned 40km pipeline that will bring gas to the country from Iran, as well as possession of an Armenian power plant (BBC News 2006c). Staying true to its desire to maintain widespread influence, Gazprom has demonstrated that it will even act with a heavy hand in countries such as Belarus and Armenia that are 'friendly' to Moscow.

Countries which belonged to the FSU that have adopted a foreign policy that is viewed as unfriendly to Moscow have been even more susceptible to Gazprom's merciless policies; this has been especially evident in Georgia. Since the breakup of the Soviet Union in 1991, Georgia has pursued a "cautious but nevertheless consistently Western-oriented foreign policy" (Rondeli 2001, 196). This gradual but deliberate shift away from the influence of its former colonial superpower and towards stronger ties with Western Europe and the United States has constantly irritated Moscow, especially as it runs counter to Russia's foreign policy strategy of maintaining a dominant presence in its near abroad. Georgia's shift towards the west has been manifested in several forms, which include the

country's clear aspirations to join the inherently Western NATO, its cooperation with the United States in advanced army training (the Georgian Train and Equip Program), and its contribution of the most troops per capita of all coalition countries to the U.S.-led forces in Iraq (Civil Georgia 2007). In addition to being one of the founding members of the GUUAM organization, a group of five countries (Georgia, Ukraine, Uzbekistan, Azerbaijan and Moldova) that have sought to develop a new non-Russian center of power in the region, Georgia's deliberate decision to disengage with the Kremlin has been a constant challenge to Russia's foreign policy goals.

The construction of the Baku-Tbilisi-Ceyhan (BTC) pipeline that deliberately bypasses Russian territory, transporting oil from the Caspian Sea to Western markets, has been a source of strong resentment from Russia as well. Led by a consortium of western-based energy companies, this new route for transporting resources from Central Asia to the Western Europe and beyond has effectively undermined Russia's monopoly on the transport of fuels from this region. Georgia's Rose Revolution in 2004 that brought to power a staunchly pro-Western government led by the former Minister of Justice, Mikhail Saakashvili, was a further signal that Russia's influence was becoming increasingly less welcome. Following a post-revolution visit by U.S. President George W. Bush to Tbilisi in May 2005, little doubt remained as to the new direction of Georgia's foreign policy.

The most significant factor in these developments is the growing Western presence in a country that Russia views as naturally part of its sphere of influence. Regardless of strong popular sentiment within Georgia in support of the government's pro-Western foreign policy, Russia has made repeated attempts to assert its will in the country by resisting this policy shift (Civil Georgia 2006a, Stice 2006). Georgia's heavy dependence on Russian gas has provided a convenient point of leverage for the Kremlin to express dissatisfaction and to try and bring Tbilisi back into its orbit. In breaking with one of the central tenets of Russia's foreign policy – maintaining control over it's near abroad – Georgia has found itself the target of repeated threats and gas cuts.

Since the end of the Soviet Union, Georgia has fallen victim to “numerous gas cutoffs...coincid[ing] with special events, such as elections, bilateral negotiations or Russian bombardment of Georgian territory” (Larsson 2006, 228). With Gazprom acquiring control of gas distribution to Georgia in 2003 from Itera, supply disruptions and threats have become even more potent. In November 2005, Gazprom announced that it would be raising Georgia's gas prices from \$USD63/tcm to \$110/tcm. These demands, however, were timed to coincide with a meeting of the CIS Energy Council in Tbilisi in order to force Georgia's hand. Only when Georgia agreed to Russia's demands to join a united energy market did Gazprom issue reassurances that there would be no disruptions (Larsson 2006, 231).

In addition to artificial alterations of price and transit, Gazprom has also targeted the ownership of pipelines as a means for exerting influence in Russia's near abroad. Gazprom has been especially relentless in its pursuit of the main trunk pipeline that transits a majority of Georgia's annual 1.5 billion cubic meter

(bcm) consumption of gas from Russia (Central Intelligence Agency 2007). With discussions dating back to February 2005, Gazprom's attempts to obtain ownership of the pipeline have been consistently rebuffed by Georgian officials (Gularidze 2005). In September 2005 Gazprom suffered a further setback in its efforts to acquire the pipelines as a result of a \$295.3 million Millennium Challenge Corporation loan from the United States. According to the terms of the loan, Georgia would receive \$49.5 million to rehabilitate the North-South gas pipeline through 2010, as long as it remained in the hands of the Georgian government (Millennium Challenge Corporation 2007). The entrance of the United States into Georgia's precarious pipeline politics was a double-blow to Gazprom's aspirations; not only did it serve to strengthen the presence of the United States in a region that Russia has declared as part of its inherent sphere of influence, but it has also threatened Gazprom's monopoly on the region's resources by forcing the company to realize that there is another major player vying for regional influence.

Faced with foreign intervention on its self-declared sphere of influence, as well as the potential threat posed to its predominant regional monopoly on gas supplies and transport, Gazprom's reaction to these events was surely unsettling. The timing of carefully calculated explosions around these two pipelines that cut gas supplies to Georgia in early 2006 was, therefore, highly suspicious. Raising additional concern was the fact that the pipeline was well within Russian territory and under the control of Russian troops in North Ossetia. Immediately following the attacks, Georgian President Mikhail Saakashvili pointed to "numerous threats by Russian politicians and officials at different levels [who wanted] to punish us for basically not giving them pipelines" (BBC 2006d). Despite Russia's promises to get to the bottom of the explosion, no charges have been filed for these destructive acts. Georgia was consequently left without gas during one of its most frigid winters in over twenty years (Ibid.) and was forced to negotiate with Iran to import extra gas at a rate of \$233/tcm— a price well above the \$110/tcm it was paying Gazprom at the time (Civil Georgia 2006b).

Gazprom has used its coercive tactics as a means of influencing politics in Ukraine as well. Ukraine's geographical location in Russia's near abroad, coupled with the large amounts of Russian gas that are transported across its territory to Europe, underscore the country's high level of importance to Gazprom and the Kremlin. The manipulation of gas prices and supplies, however, is nothing new to Kiev. Since the early 1990's Gazprom has attempted to impose its will on Ukraine with respect to certain political decisions. During the summer of 1993, an important meeting was set between Ukrainian President Kravchuk and Russian President Yeltsin to discuss key issues that included the status of the former Soviet nuclear arsenal in the region, Ukrainian debts to Russia and the state of Russia's Black Sea Fleet (Kasianov 1995). Yet just one week prior to the conference, Gazprom cut Ukraine's gas supplies by 25 percent, citing 'non-payment' as justification (Fredholm 2005). The only way that the gas could be returned to its full supply was if Ukraine could find some way to pay these 'debts'. It was understood that if President Kravchuk gave in to Russia's demands, the 'debts' would be forgiven and the gas supply returned to normal (Ibid.). This same coercive behavior was witnessed again in 1995 when Russia

tried to force Ukraine to join the CIS Customs Union. While waiting for a response from Kiev, Gazprom raised the export prices of its gas in an attempt to influence Ukraine's decision in the matter (Fredholm 2005).

Gazprom's political overtures grew increasingly more dramatic in the fall of 2004. In the heat of a presidential election that pitted a pro-Western candidate, Viktor Yushchenko, against a pro-Russian candidate, Viktor Yanukovich, Gazprom threw the full weight of its support behind Yanukovich by promising contracts at the favorable price of \$50/tcm if the pro-Russian candidate was elected (Larsson 2006). Following Ukraine's Orange Revolution in the winter of 2004-05 that brought the Yushchenko government to power, Ukraine's foreign policy direction underwent a decisive shift away from Moscow. Since coming to office, Yushchenko has strengthened Ukraine's resolve to join both the EU and NATO, and has received widespread support from the Bush government in Washington in its pursuit of these endeavors.

With all signs pointing to a deliberate move away from Ukraine's historically close ties to Russia, Gazprom's tactics became more aggressive. In the middle of Yushchenko's first icy winter in office, Gazprom suddenly declared in December 2005 that it would be increasing the price of gas by more than a factor of four – from \$50 to \$230/tcm. If an agreement could not be reached by January 1, 2006, all gas supplies would be cut. After three days with no gas in the beginning of January, Ukraine and Gazprom came to an agreement that resulted in prices jumping from \$50 to \$95/tcm – a price Gazprom claimed could have been reduced if Ukraine had agreed to sell the pipeline crossing its territory into Europe (BBC News 2006e). In response to these forceful tactics, U.S. Secretary of State Condoleeza Rice declared that Russia's actions were “undoubtedly directed by political motives” and that it was “using its energy resources as a political weapon” as a punishment for Ukraine's new pro-Western orientation (Larsson 2006, 209).

The countries in the Baltic region have also been susceptible to the forceful use of gas supplies. Upon gaining their independence in the winter of 1992-93, energy cuts were made to all three Baltic States to send a clear message about Russia's dissatisfaction with their break from the defunct Soviet empire. Not only was Russia finding it difficult to come to terms with the territorial loss of the former soviet republics, but it was struggling to salvage any sense of political allegiance as well. Following calls from Estonia, Latvia and Lithuania for the removal of Russian troops from their respective territories, the Kremlin imposed a four-fold increase in the price of gas as a punishment for their ‘disloyalty’ (Larsson 2006).

Subsequent decisions by some of the Baltic States since their independence have been particularly vexing to the Kremlin. Prior to the 1993 elections in Estonia, the government in Tallinn had been discussing the adoption of a “Law on Aliens” that would address issues such as visas, legal compliance, immigration quotas and employment for non-citizens. The perception in Russia, however, was that these new laws would be detrimental to the lives of the non-indigenous Russian population living in Estonia. In accordance with one of the main tenets of its post-Soviet foreign policy - protecting the interests of Russians living aboard - Russia once again relied on Gazprom to assert its influence. With

Russian officials stating that “the government and parliament...have quite a few levers to make Estonian authorities realize the impermissibility of violating the rights of non-indigenous people” (Larsson 2006,189), gas cuts were seemingly only a matter of time if the laws were approved. With the adoption of the Estonian laws on June 21, 1993 (the “Law on Aliens”), Gazprom proceeded to cut gas supplies to rebuke Tallinn for its political decisions (Fredholm 2005).

In response to accusations that it was raising prices to gain political leverage over the aforementioned countries and others, both Kremlin and Gazprom officials have repeatedly stressed that price hikes “merely reflect market prices” (BBC News 2006b and Radio Free Europe 2007). The Russian argument, however, does not stand up to reason. A blatant disregard for ‘market prices’ is evident in Russia’s policy towards the separatist territories in its near abroad. The behavior of Gazprom in the Transnistria region in Moldova and the Abkhazia and South Ossetia regions in Georgia clearly underscores this point. Each of these regions has close ties to Russia and has been supported by Moscow in various ways in their respective struggles against sovereign host governments. Despite Transnistria’s debt to Gazprom, a figure that has reached more \$1 billion, the breakaway region has received gas “almost gratis for a decade” (Socor 2006). Even during Gazprom’s supply cuts to Moldova in the winter of 1999 and February of 2000, Transnistria, despite being recognized by the international community as part of Moldova, never had its gas supply threatened.

This same form of price discrimination that defies the principles of ‘market pricing’ can be found in Georgia’s rebellious regions as well. Since breaking from the government in Tbilisi shortly after Georgia’s independence in April 1991, Abkhazia and South Ossetia have received significant energy resources from Russia. In the words of President Saakashvili,

There is nothing ‘free market’ or ‘market rate’ behind Russian energy prices...In Georgia, both Abkhazia and South Ossetia, two areas that are outside of our control and whose separatist authorities are directly controlled by Russia, receive natural gas free – hardly a practice free marketers would applaud (President Saakashvili, quoted in Larsson 2006, 211).

This point has been further supported by the construction of a 163 km gas pipeline that runs directly from Russia to South Ossetia’s capital, making the breakaway region independent from Georgian gas supplies (Radio Free Europe 2006). In a region that has a sparse, unemployed population and nearly no industry, it is evident that Gazprom’s decision to begin constructing the pipeline in November 2006 was not financially motivated. Such biased and politically-driven actions in these secessionist regions are best described by the Chair of the State Duma Foreign Affairs Committee in Russia, Konstantin Kosachev. Emphasizing the political role of Gazprom, Kosachev stressed that Moscow “will continue to subsidize energy supplies to its allies” and “apply market principles to those countries [with which] we don’t have an alliance-type relationship” (Larsson 2006, 220). With a deteriorating military and a sense of nostalgia for its old glory, Russia’s energy resources “are the former

superpower's last remaining hope for regaining its world status" (Grachev 2005).

POLICY IMPLICATIONS OF RUSSIA'S FOREIGN POLICY STRATEGY

Using covert tactics to forcefully acquire strategic foreign gas properties and overt tactics of threats, price increases and supply cuts, Gazprom has unquestionably proven itself as a powerful political tool that has staunchly defended the interests and policies of the Kremlin with great dexterity. Yet as Russia continues to employ this asymmetrical strategy for exerting its influence abroad, the forceful use of Gazprom has created important policy implications for both Gazprom customers as well as the international community. On the international level, Russia's successful use of this tactic could set a dangerous precedent by establishing a model that could easily and effectively be replicated by other states with market dominance in the energy sector. Such a standard could adversely affect the numerous countries that rely on foreign energy resources by empowering these national suppliers with the idea that their own energy supplies could be used to help dictate the foreign policies of dependent nations.

Yet for the countries that are currently dependent upon Russia for the majority of their gas supplies, the possibility of reducing Gazprom's influence in the near future remains low. Handicapped in their ability to search for alternative sources by the dint of their close geographic proximity Russia's vast pipelines and resources, these heavily dependent countries face additional difficulties amongst themselves as each weighs its own individual interest against that of the group. While collective action (in terms of negotiating the construction of alternate pipelines, price bargaining, or exploring for non-Russian sources of gas) would be one of the most effective ways of resisting this coercion, such a strategy has been hindered by the fact that each country has its own individual contract dictating the price and quantity of supplies that it receives from Gazprom. Faced with the inability to negotiate price and quantity as a trading bloc or to obtain gas supplies from another region, these countries remain for the near future at the mercy of Gazprom's policies.

While the international community may in fact have an understanding of Gazprom's forcefulness, it is unlikely that any serious action will be taken to dissuade Russia from using this tactic. Russia's position on the UN Security Council, as a member of the G8, and as one of the largest suppliers of oil on a daily basis, together serve to dissuade the international community from pursuing the strong action that would be needed to force an end to this policy. Any action that might limit or sanction Russia's manipulative use of Gazprom could come at the cost of losing Russian cooperation on a host of several other relatively more threatening issues. The need for Russian support on stemming the proliferation of nuclear weapons, international efforts to reduce climate change, counter-terrorism activities, UN reform, and aid to Africa exceeds the urgent desire of many dependent countries to prevent Russia from using Gazprom forcefully.

In the short term, the most effective form of resisting Gazprom is for each

heavily dependent country to diversify its supplies where possible and to develop more efficient energy policies, including the use of renewable sources. While the cost of employing many of the newer and more efficient technologies may be prohibitive for some of the countries in Russia's near abroad, the most realistic policy in the short-term is for these countries to make efficient use of the resources available to them, so as to not feed into Gazprom's influence any further. In the long term, these countries should work to lobby for access to non-Russian sources of gas, specifically in Central Asia and in the Caspian region. If the countries in Russia's near abroad can crack Gazprom's heretofore monopoly on these sources of gas by strengthening diplomatic ties with other resourceful countries and by building new pipelines bypassing Russian territory, Gazprom will face the possibility that alternative sources of energy are viable options threatening its monopoly. If such a threat were to exist, the possibility that Gazprom would moderate its behavior for the sake of keeping its foreign customers could not be so far-fetched. With Russian influence continuing to grow in these areas, however, this will be no easy task.

NOTES

¹ In 2001, two investigations were undertaken to investigate Gazprom's relationship with Itera. One of these was conducted by Gazprom's own auditors, PricewaterhouseCoopers (PwC), while the other was conducted by the Russian Audit Chamber. The results of the PwC audit were released in July, 2001 to the public and concluded that "Russian legislation was not violated" in the relations between the two companies. The results of the Russian Audit Chamber (a body which acts as the government watchdog, similar to the United States' Government Accountability Office) were generally inconclusive regarding this issue.

² See "Murky Deals At Gazprom" *Business Week*, June 21 2004 (http://www.businessweek.com/magazine/content/04_25/b3888073_mz054.htm), and "A Suspect Gazprom Deal," *Alexander's Gas and Oil Connections*, Vol. 8, Issue 3 (<http://www.gasandoil.com/goc/company/cnr30654.htm>)

³ The fourteen former Republics of the Soviet Union are: Azerbaijan, Armenia, Georgia, Belarus, Moldova, Ukraine, Estonia, Latvia, Lithuania, Kazakhstan, Kyrgyzstan, Uzbekistan, Turkmenistan and Tajikistan.

⁴ This term alludes to a policy of non-intervention that came to be known as the Monroe Doctrine, named after former U.S. President James Monroe. Delivered in address to Congress in 1823, Monroe outlined a policy based on establishing distinct spheres of influence in each hemisphere; intervention or colonization by another world power into the sphere claimed by the United States would be deemed as a hostile encroachment upon its self-declared orbit of power. The term is used here in relation to Russia's policy of ensuring that it maintains predominant influence over the territory that consisted of the Former Soviet Union.

REFERENCES

- British Broadcast Corporation News (BBC)*. 2006a. Russia Vows to End Gas Shortage, January 2. <http://news.bbc.co.uk/2/hi/europe/4575726.stm#table> (accessed March 11, 2007).
- BBC News*. 2006b. Belarus Claims Gas Dispute Deal, December 30. <http://news.bbc.co.uk/2/hi/europe/6220455.stm> (accessed March 2, 2007).
- , 2006c. Russia to Control Armenia's Gas, April 7. <http://news.bbc.co.uk/2/hi/business/4888012.stm> (accessed March 10, 2007).
- , 2006d. Russia Blamed for 'Gas Sabotage', January 22. <http://www.bbc.co.uk/2/hi/europe/4637034.stm> (accessed March 21, 2007).
- , 2006e. Q&A: Ukraine Gas Row, January 4. <http://news.bbc.co.uk/2/hi/business/4569846.stm> (accessed March 20, 2007).

- Blagov, Sergei. 2006. Putin and Nazarbayev Ink Deal to Process Gas at Orenberg. *Eurasia Daily Monitor*, July 18. http://jamestown.org/edm/article.php?article_id=2371284 (accessed May 14, 2007).
- Central Intelligence Agency. 2007. *The World Factbook: Georgia*. <http://www.cia.gov/cia/publications/factbook/geos/gg.html> (accessed March 22, 2007).
- Civil Georgia*. 2007. Saakashvili: 2000 Troops in Iraq, at least 100 in Afghanistan, March 3. <http://www.civil.ge/eng/article.php?id=14750> (accessed May 14, 2007).
- , 2006a. Russian MFA Warns Against Georgia's NATO Membership, September 22. <http://www.civil.ge/eng/article.php?id=13618> (accessed May 13, 2007).
- , 2006b. Georgian Government Keeps Silent About Iranian Gas Price, February 1. <http://www.civil.ge/eng/article.php?id=11678> (accessed March 2, 2007).
- Energy Information Administration (EIA). 2007. *World Proved Reserves of Oil and Natural Gas, Most Recent Estimates*. <http://www.eia.doe.gov/emeu/international/reserves.html> (accessed March 13, 2007).
- Fredholm, Michael. 2005. *The Russian Energy Strategy and Energy Policy: Pipeline Diplomacy or Mutual Dependence*. Conflict Studies Research Center, Defense Academy of the United Kingdom, England.
- Gazprom. 2006a. *Before 2001: A History of the Company*. <http://www.gazprom.com/eng/articles/article8517.shtml> (accessed March 17, 2007).
- , 2006b. *Gazprom's Board of Directors*. <http://www.gazprom.com/eng/articles/article8823.shtml> (accessed March 22, 2007).
- Grachev, Andrei. 2005. Putin's Foreign Policy Choices. In *Leading Russia: Putin in Perspective*, edited by Alex Pravda, Ch. 13. Oxford: Oxford University Press.
- Gularidze, Tea and Giorgi Sepashvili. 2005. Of Pipes and Men. *Civil Georgia*, February 24. <http://www.civil.ge/eng/article.php?id=9168> (accessed March 12, 2007).
- International Helsinki Federation of Human Rights. 2006. *Human Rights in the OSCE Region*. IHF Report 2006. http://www.ihf-hr.org/viewbinary/viewdocument.php?doc_id=6868 (accessed March 12, 2007).
- Kasianov, Heorhi. 1995. Russia-Ukraine: The *Imperial Syndrome*. *Perspective 6: 2* Institute for the Study of Conflict, Ideology and Policy. <http://www.bu.edu-iscip/vol6/Kasianov.html> (accessed March 23, 2007).
- Larsson, Robert L. 2006. *Russia's Energy Policy: Security Dimensions and Russia's Reliability as an Energy Supplier*. Swedish Defense Research Agency (FOI) Defense Analysis. March.
- McGregor, Caroline. 2002. *Gazprom and Itera: A Case Study in Russian Corporate Misgovernance*. Carnegie Endowment for International Peace. <http://www.carnegieendowment.org/events/index.cfm?fa=eventDetail&id=468&&prog=zru> (accessed March 10, 2007).
- Millennium Challenge Corporation. 2007. *Georgia*. <http://www.mcc.gov/countries/georgia.index.php> (accessed March 19, 2007).
- Mosnews*. 2006. Georgia Faces Cutoff in Russian Gas Supplies if No Contract is Signed, November 8. <http://www.mosnews.com/money/2006/11/08/gasgeor-gia.shtml> (accessed March 19, 2007).
- Organization for Economic Cooperation and Development (OECD). 2004. *OECD Economic Survey of the Russian Federation 2004: Reforming the Domestic Natural Gas Market*. <http://www.oecd.org/dataoecd/56/4/32389306.pdf> (accessed March 12, 2007).
- Radio Free Europe Radio Liberty. 2006. *Russia: Ossetia Pipeline Adds Fuel to Georgia Fire*. <http://www.rferl.org/featuresarticle/2006/11/13dacbc1-7f2e-4b99-8e88-079ca5a7f35f.html> (accessed March 12, 2007).
- Rondeli, Alexander. 2001. The Choice of Independent Georgia. In *The Security of the Caspian Sea Region*, edited by Gennady Chufirin, Ch. 11. Oxford: Oxford University Press.
- Russell, Wynne. 1995. Russian Relations with the 'Near Abroad'. In *Russian Foreign Policy Since 1990*, edited by Peter Shearman, Ch. 3. Boulder, Colorado: Westview Press.
- Stern, Jonathan P. 2005. The Future of Russian Gas and Gazprom. *Oxford: Oxford University Press*.
- Stice, Nathanael. 2006. *Crisis in Georgia: Frozen Conflicts and US-Russian Relations*. Carnegie Endowment for International Peace. <http://www.carnegieendowment.org/events/index.cfm?fa=eventDetail&id=920&&prog=zru> (accessed May 13, 2007).
- The Economist*. 2007. A Cushioned Blow. http://www.economist.com/daily/news/PrinterFriendly.cfm?story_id=8485325 (accessed August 01, 2007).

-
- The Jamestown Foundation. 1998. *Moldovan Government Caves in to Gazprom*. 4: 191.
http://www.jamestown.org/publications_details.php?volume_id=21&issue_id=1402&article_id=1428 (accessed March 8, 2007).
- Voice of America. 2006. *Russia's Gazprom Threatens Belarus With Huge Gas Price Hike*.
<http://www.voanews.com/english/archive/2006-12/2006-12-29-coa1.cfm?CFID=84855547&CFTOKEN=78433784> (accessed June 12, 2007)
- United Nations. 1987. *Convention Against Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment*. Office of the High Commissioner for Human Rights.
http://www.unhcr.ch/html/menu3/b/h_cat39.htm (accessed April 15, 2007).
- . 2002. *Protocol to the Convention Against Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment*. Office of the High Commissioner for Human Rights.
<http://www.unhcr.ch/html/menu2/6/cat/treaties/opcat.htm> (accessed April 15, 2007).

PROTECTION FOR WHOM? STABILIZATION AND COERCIVE RULE IN HAITI

Jennifer Peirce

Norman Paterson School of International Affairs, Carleton University

This article examines police reform initiatives in Haiti since 1994 in the context of how armed groups, including police, have been used to maintain order and power in the 20th century. Drawing on the model of the protection racket state, it argues that the political links and consequences of local rule by gangs perpetuate violence and conflict. International security and police missions target this ‘criminal’ violence, but have not confronted the political and patronage networks of armed groups and the police force. International pressures for regional geopolitical stability create a ‘low-intensity democracy,’ in which the national government’s prospects for implementing substantial institutional and socio-economic reforms are constrained.

“This is a war. We sit out on our terrace at night and listen to the gunfire.”

Brian Moller, Médecins Sans Frontières staff member in Port-au-Prince
(*Toronto Star*, November 5, 2006)

“Democratization can routinize the authoritarian legacy and ultimately legitimize the silent violence of daily material deprivation afflicting oppressed majorities.”

Robert Fattou, Jr., Haitian scholar (Fattou 2002, 5)

INTRODUCTION

Humanitarian and political leaders talk about the ‘reconstruction’ of Haiti, and numerous countries have joined the project of ‘peace-keeping’ in the country. Haiti is not in the aftermath of a civil war, nor negotiating ceasefires or peace agreements, nor is it even caught in the turbulence of revolution or dictatorship. But the daily violence afflicting Haiti is more prevalent and, in many cases, political, than almost any other country in the hemisphere. More than 8,000 murders and 35,000 sexual assaults occurred in the capital, Port-au-Prince, during the twenty-two months following the overthrow of President Jean-Bertrand Aristide (Kolbe and Hutson, 2006). Prisons are overflowing with inmates awaiting charges from a nearly inert judicial system.¹ Haiti has been described as “the world’s first permanent failed state” (International Crisis Group, 2006), and “a patient [who] failed to recover,” despite past interventions (von Einsiedel and Malone 2006, 154). A major focus of the current intervention, the United Nations Stabilization Mission in Haiti (MINUSTAH), in place since 2004, is controlling crime and violence, primarily through strengthening the police system. This is a complex task, for there is no precedent for just or neutral state law enforcement in Haiti. The relative stability of recent months may be difficult to sustain unless ongoing reforms and activities of the police in Haiti counteract, rather than ignore or perpetuate, the historical patterns of coercive rule in Haiti.

This article contends that the current approach to police reform remains constrained by contradictory local and regional political pressures, and that the emphasis on stability allows the national government minimal opportunity to move beyond coercive methods of governance. The article will apply Charles Tilly's (1985) model of the protection racket state to analyse the use of violence for political power in Haiti, and will suggest that the institutional conditions for law enforcement are distorted. This article also argues that in present-day Haiti, order and power are maintained at two levels: locally, through gang-based protection rackets, and trans-nationally, through "consensual domination" (Robinson 1996, 315). It points to numerous conceptual and operational aspects of police reforms since 1994 that have contributed to, and resulted from, this arrangement. The first section outlines how state agents of violence enforced an exploitative form of 'order' in the 20th century, the political and economic consequences of which remain entrenched. The second section frames gang violence as a form of coercive governance, which national factions can manipulate through neo-patrimonial networks. The following sections propose that international agendas, not the national government, determine the version of 'stability' to be promoted in Haiti. These pressures are contradictory and often prioritize superficial forms of order, to the detriment of more substantial socio-economic reform. The concluding section discusses the opportunities and challenges for police reform and building security, one year into the strong electoral mandate and ambitious projects of the national government under President René Préval.

THE DYSFUNCTIONAL PROTECTION RACKET: THE HISTORICAL CONTEXT OF VIOLENCE AND POWER

Haiti is well acquainted with what Charles Tilly calls "the organized means of violence" (Tilly 1985, 170) – power based on controlling the exertion of violent force. Unsurprisingly, the current stabilization efforts in Haiti attempt to curtail the use of violence by non-state groups, primarily gangs. Tilly's protection racket model posits that functional state institutions can develop over time through a reciprocal relationship: the state offers protection, mainly through force, to its citizenry in exchange for access to resources, mainly through taxation (Tilly 1985). Throughout the twentieth century in Haiti, this dynamic became distorted. Rather than building a mutually beneficial contract with the populace, the Haitian state's use of violence built exclusive protection rackets based on precarious alignments of elite interests, which have exploited the population and stunted economic development. Those in power (officially and unofficially) have used all available means, including violence, to uphold self-serving arrangements against internal threats from other factions.

Tilly proposes that the process of building a legitimate state involves four elements: extraction, protection, war-making, and state-making (Tilly 1985, 182). The case of Haiti shows contradictions and twists on all of these elements. The extraction of resources has been exploitative since the colonial era. The modern taxation system is minimal and still overburdens the peasantry. Agricultural products have traditionally been taxed and exported under

monopoly arrangements, which allowed for the growth of a class of exploitative middlemen, while tax exemptions on imports benefited mainly the elite class (Kumar 1998, 47). Rather than seeking further resources through territorial expansion, as Tilly's model posits, rulers in Haiti have attacked and exploited the lower classes and rival factions, capturing their resources by threat of force or by offering protection against internal enemies.

As a country founded by freed slaves in 1804, and subject to economic and diplomatic hostility from the US and France in the 19th century, Haiti could perhaps have followed Tilly's path of state consolidation by protecting itself from external threats. Instead, divisions persisted until the US military occupied Haiti (1915-1934) and established and armed Haiti's first version of a police force, the gendarmerie, later converted into the military. This provided a means of resource extraction and political control for rulers – for example, a peasant uprising in 1918 was quickly quashed (Robinson 1996, 268). This pattern persisted under the dictatorship of Francois (Papa Doc) Duvalier. Due to the constant threat of military coups, Duvalier governed by channelling wealth through his personal security network, the *Volontaires de Sécurité Nationale* (VSN), also known as the *TonTons Macoutes*. The *TonTons* were groups composed of poor black men that eventually became more powerful than the army and notorious for extreme violence. A major challenge for contemporary police reform efforts is to overcome the legacy of fear and rivalry surrounding these state agents, military and non-military, many of whom now operate in extra-legal networks.

The protection racket that developed in twentieth century Haiti relied on the exclusion and exploitation of the majority of Haitians. The economic elite agreed to pay some taxation to the dictator, in exchange for the protection of their monopolistic business arrangements. The military and private security agents, however, were generally not from the wealthy class. This balance of power therefore began to split when Jean-Claude Duvalier, motivated by his marital ties to the economic elite class and his ambitions to develop assembly factories in Haiti, offered more state resources to the economic elite, which the military networks perceived as a loss of their relative power (Robinson 1996; Snyder 1992). At the same time, the poor majority of Haitians were beginning to mobilize under the *Ti Legliz*, a grassroots movement led by Jean-Bertrand Aristide, a radical priest who called for redistribution of wealth. The breakdown of the alliance between the military and economic ruling classes led to competition for power, for the state no longer held absolute control over the agents of repression. This, bolstered by international pressures against the dictatorship and assistance for electoral procedures, allowed Haiti its first experience with democracy, with the election and eight-month government of Aristide's Lavalas party in 1990-1991. Without the reciprocal economic relationships built by a functional protection racket state, however, Haiti soon returned to "*la politique du ventre*," where rival factions "have been literally eating each other to digest the limited fruits of power" (Fatton 2002, 13). This has left Haiti with a monopolistic and weak economy, a profound distrust amongst sections of society and of the state, and the absence or weakness of functional state institutions – a situation in which neo-patrimonial dynamics

continue to flourish.

NEO-PATRIMONIAL USES OF ARMED GROUPS

Neo-patrimonial rule, “the chief executive’s maintenance of state authority through personal patronage networks, rather than ideology or impersonal law” (Snyder 1992, 279), has been a primary mechanism of influence in Haiti, especially through police, military, and other state agents of enforcement. The Duvaliers purchased loyalty by distributing modest salaries and privileges through military or *macoute* gang networks (Kumar 1997, 37; Robinson 1996, 267). Aristide also maintained rule (in and out of office) in part through the use of armed groups, known as *les chimères*, to intimidate and assassinate those who posed a threat to him (Mendelson-Forman 2006, 23). Aristide’s 1994 decision to dismantle the Haitian army dissolved a major patronage network. Former soldiers were recruited by organized crime and by the Haitian National Police (Mobekk 2001), and much of the current violence in rural areas is attributed to illicit networks of ex-soldiers (International Crisis Group 2007, 4). Even in present times, under a new President and with Aristide exiled in Africa, political factions and positions tend to be identified as pro- or anti-Aristide, suggesting the continuing influence of his persona on Haitian politics. Current patronage networks, however, are significantly less fixed and predictable. The government must compete with other sources of patronage such as rival political leaders, foreign actors, the private sector, and organized crime, for control of these networks. Endemic corruption and divided loyalties within the Haitian National Police can be partly attributed to this circumstance. A regular salary as a police officer is a rarity in Haiti, where jobs are few. Even President Préval’s recent campaign against corruption may not be sufficient to overcome the appeal of other patronage networks.

The pattern of neo-patrimonial politics in Haiti presents two challenges for the Préval government’s attempt to establish meaningful authority and rule of law. First, gangs with shifting alliances have become the de facto governments of neighbourhoods. Second, international actors in Haiti create the contradictory pressures of seeking stability while perpetuating divisions. These challenges will be investigated below.

THE LOCAL PROTECTION RACKET AND PATRONAGE POLITICS: *MACOUTES, CHIMÈRES AND ZINGLENDOS*

This section contends that the dysfunctional version of Tilly’s protection racket state in Haiti has devolved to the local level, now existing on the level of urban gangs. Dismantling urban gangs has been a priority of the UN Stabilization Mission, and the arrest of gang leaders – reportedly 750 have been arrested in 2007 so far (International Crisis Group 2007, 1) – is cited as justification for the increased level of force used in its confrontations in neighbourhoods such as Cité Soleil (Mulet, 2007). The portrayal of gangs as a problem of criminality is, however, overly simplistic, for it does not address the political role of the gangs. In contrast to the Duvaliers’ *macoutes* gangs, and the army units of the military dictatorship of 1986-1990 (including General Cédras from 1991-1994), which

operated nationally with direct funding, present-day gangs are often rooted in particular communities. They generate profit through crime, most notably extortion, kidnapping for ransom, and drug-trafficking (International Crisis Group 2005, 4), and they exert control over delineated territory. In addition to standard crime profits, gangs extract money and resources from the residents of the neighbourhoods they control through coercive taxation, demanding large portions of residents' limited funds,² and siphoning local development project money (International Crisis Group 2007, 6). In exchange, the 'protection' provided by the gangs consists of defending the area from rival gangs and allowing residents to access the basic daily necessities, such as roads, health clinics, and water taps. Rodgers describes a similar phenomenon in contemporary urban Nicaragua where youth gangs are prevalent, and suggests the micro-level parallel with the conventional state:

This is particularly striking in relation to the gangs' use of violence, which like the state's exercise of violence is arguably not only teleologically related to the constitution of social order, but moreover seems to follow remarkably similar modes of organisation. Like the Weberian state at the national level, gangs possess a monopoly over the predominant forms of violence at the local neighbourhood level, and these are moreover considered to be legitimate forms of violence, at least in terms of the local manifestation of the gang, because of their simultaneously protective and informative functions. (Rodgers 2004, 7).

The difference in Haiti is that the local gangs operate not in isolation from a national state, but in a continually shifting patron-client relationship with either the state or opposition factions. Some of Aristide's *chimères* arose from his local political support networks (Faton 2002; von Einsiedel and Malone 2006, 23), and thus the *chimères* have a certain degree of political legitimacy in their respective neighbourhoods. For example, Dread Wilmé, who was killed by MINUSTAH in July, 2005, was portrayed as a dangerous gangster by MINUSTAH chief Juan Gabriel Valdes (Valdes 2006, 10), but is considered a "hero of resistance" by some residents of Cité-Soleil (Haiti Information Project 2005). On the other end of the political spectrum, former army members and *macoutes* have also reconstituted into gangs (International Crisis Group 2005). Because the alliances between gangs and national factions are not fixed, gangs operate primarily for criminal profit but can use their localized monopoly of violence to act as proxy rulers for the highest bidder.³ The currency of this relationship includes not only money, arms, drugs and judicial exemption or leniency, but also political positions, material items, and jobs. The rival ruling factions in Haiti, who continue to view power as a zero-sum game, frequently wield influence through these armed groups in ways that Carey describes as an "international relations security dilemma," where a defensive move by one actor is understood by an opponent as a provocation (Carey 2005, 333).

The public portrayal of these groups as *zinglendos* – criminals – obscures their political impact, but their availability for political mobilization is high, particularly according to the criteria of the "youth bulge" hypothesis (Urdal 2004). Over a third of Haitians are under the age of fifteen (Prest 2005, 12), and young people face bleak prospects for economic survival in Haiti. Furthermore,

many of the youth who go to the US seeking work are deported back to Haiti angry, often with some criminal experience (Fattou 2002, 128; Kumar 1997, 70), and easily recruited to gangs. After the demobilization of the Haitian Armed Forces in 1994, ex-soldiers were left with few appealing economic alternatives and thus joined armed groups, narco-trafficking groups, and to a significant extent, the reconstituted Haitian National Police (HNP) (Mobekk 2001; Mendelson-Forman 2006). The 2006 re-activation of the Disarmament Commission is a sign that the Préval government is responding to past critiques of solely aggressive anti-gang tactics, but the Commission's resources are limited and gang leaders are not eligible for its reintegration program (Associated Press, 2006). NGO-run social and psychological "reinsertion" programs and disarmament efforts are receiving important attention (International Crisis Group 2007, 8). The weakness of economic opportunities, though, continues to push young men to join gangs or to migrate out of the country.

MINUSTAH's tactical decision to target crime where it is most prevalent – in urban slums – while ignoring the political implications of this stance, has created a distrust among Haitians who perceived the HNP, and by extension the CivPol who back them, as agents of anti-Lavalas forces (Donais 2005, 281), as well as an external occupying force (Talentino 2007). The politicization of gangs appears to be shifting, however. MINUSTAH's aggressive operations targeting gangs in Port-au-Prince in December, 2006 was authorized by Préval, a step that UN Envoy Mulet called "political ownership" (Mulet, 2007). Préval also maintains support for a single police force, rather than municipal police forces that are more susceptible to local co-optation (International Crisis Group 2007, 20). Still, accusations of excess force and human rights abuses by UN troops against community residents remain unaddressed (Harvard Law Student Advocates 2005; Howland 2006; Haiti Information Project 2007). Moreover, the volatility of the local context is linked to the geopolitical priorities of Haiti's influential international neighbours.

THE TRANS-NATIONAL RULES: THE TERMS OF STABILIZATION

The following discussion draws on the concept of polyarchy⁴ (Robinson 1996) to contend that the Haitian government's authority exists only under the conditions set by international actors, through trans-national authority shifting from coercive to consensual (Robinson 2003, 53). According to Robinson, the subordinate core-periphery relationship of dependency theory continues to shape the economics of globalization, but the division is not geographical or state-based. Rather, trans-national elites from rich and poor countries form the core that exploits the periphery, and this core seeks a minimized national state apparatus (Robinson 2003, 40). The detrimental effects on the well being of most Haitians, which have resulted from the collusion of Haitian elites with international powers, are undeniable. But the incentives for trans-national neo-liberal rule in Haiti are not as clear-cut as in some of Robinson's Central American case studies, due in large part to the dysfunctional, monopolistic nature of capitalism in Haiti.

The strongest motivation for external actors to promote the “liberal peace project” (Willet 2005) in Haiti is not economic profit; it is ‘stability’ – achieving a minimal status quo to stem the chaos and migration from Haiti. This is especially clear in the issue of enforcing the rule of law and police reform, which have been the central goals of the United States and international interventions since 1994. An important factor in Clinton’s decision to re-instate Aristide by military force in 1994 was the perceived threat of Haitian ‘boat people’ arriving in the United States; Aristide also used this fear as a lobbying tactic (Malone 2006, 155). More recently, rationales for intervention refer to Haiti’s position in international drug-trafficking routes, its potential for economic investment (Macnamara 2005), and diaspora connections (MacKay 2007). Contrary to the logic of neo-liberalism, these interventions have attempted to bolster, not dismantle, some government institutions in Haiti through such activities as recruiting and organizing thousands of police officers, judges, and bureaucrats.

International interventions, such as the example discussed in this article, promote “low-intensity democracy,” defined in the Haitian context as “a variant of democracy divorced from any empowering notions of social justice and participation, [which] privileges civil and political rights over social and economic rights. It is an election- and institution-centred view of democracy with a focus on order and stability” (Shamsie 2004, 1098). For example, although the United States used the rhetoric of democracy to justify reinstating Aristide by force in 1994, his autonomy was limited by his dependence on international forces to implement the law, the conditions of international loans and, when he rejected such conditions, a freeze on American aid and international loans worth US\$500 million in 2000 (Erikson 2005, 84). These factors curtailed the Lavalas socio-economic agenda of 1990 and its potential to spark “upheaval.” Fatton suggests that the simultaneous political elevation and economic deradicalization of the Lavalas party is a contradiction in the logic of the intervention (Fatton 2002, 197). In the longer term, however, it is an effective approach: a low-intensity democracy that co-opts potentially radical factions (on the left and the right) better serves the international goal of stability than the installation of a puppet government.

Past instances of ideological manipulation, such as the CIA covert funding of the FRAPH (Front for the Advancement and Progress of Haiti) paramilitary and the military dictatorship of Cédras (Chomsky, Farmer & Goodman 2004, 6), have faded in favour of the convergence of governance and security discourses with the frame of development and democracy (Duffield 2001; Robinson 1996). After well over a decade of low-intensity democracy existing as a tensely acceptable alternative to the unknown threats of genuine democracy, however, Aristide’s personal dominance over the sub-national factions of power became a cause of complaint following the contested elections of 2000. Despite the many criteria for legitimate democracy, which Aristide’s government was not fulfilling at the time, the issue of electoral fraud – i.e. the basis of procedural democracy – was the main justification with which the international community accepted (some would say pushed) his departure on February 28, 2004. The Transitional Government (also known as the “Council of the Wise”) established in the wake of the coup was chosen by international

forces and drew heavily on leaders of the elite classes. The fact that CARICOM refused to recognize this government, however, indicates that even the “unusually internationalized” (Carey 2005, 349) politics of Haiti are not solely influenced by the United States or elites. Venezuelan President Hugo Chávez’s recent donation of USD\$100 million in petroleum and development assistance, along with the Cuban donation of medical teams, is a deliberate challenge to American influence in Haiti (Joseph and Concannon 2007). Brazil’s high-profile role in MINUSTAH is seen in part as a strategy for gaining credibility as a major player at the UN level (von Einsiedel and Malone 2006, 166). Policy proposals to the Canadian government, meanwhile, have outlined “mentorship” arrangements through which Canadian ministries would effectively manage the corresponding Haitian ministry (Dade and Graham 2004).

These examples of trans-national authority and low-intensity democracy do not add up to a single international agenda. Nevertheless, the analytical frames allow a broader critical assessment of international involvement in MINUSTAH and police reform projects. While noting the pragmatic challenges of specific initiatives (Beer 2004; Donais 2005; Mendelson-Forman 2006) and even recent instances of “progress” (Mulet 2007) is important, the more systemic issues that underlie these projects and their antecedents are equally relevant. The following sections address two elements of international influence: direct government funding and support for the United Nations and policing programs; and other relevant international funding through civil society and organized crime.

INTERNATIONAL INTERVENTIONS: STABILIZATION AND DEVELOPMENT?

Police and military actions are currently the most prominent element of foreign governments’ involvement in Haiti and exemplify the extent of the trans-nationalisation of authority in Haiti. The consequences of this situation are complex, but some of the problems of police projects in Haiti can be attributed to the dominance and limitations of international actors in determining how to address (in)security in Haiti. The former head of MINUSTAH, Juan Gabriel Valdes, identified two levels of stabilization that MINUSTAH seeks to achieve: stabilization of the government so that it will not be vulnerable to coups; and stabilization of the economy so that it will not be vulnerable to “agents of chaos” (Valdes 2006, 13). These goals fit squarely within the models of trans-national authority and low-intensity democracy. Some of the difficulties that the United Nations and police projects have encountered, however, result from the demands, priorities, pressures and limitations of foreign governments. The following examples suggest that the insistence by international actors that police reform projects focus on addressing what they perceived to be threats, and the expectation of expedient results from flawed Haitian institutional arrangements, detracted from building long-term, local stability.

Duffield’s insight that “underdevelopment has become dangerous” (Duffield 2001, 2) is particularly fitting for Haiti, where violent urban slums and illegal immigrants are the emblems of poverty and chaos. The United States-

and United Nations-led interventions of the mid-1990s were among the first experiments in merging security and development discourses (Mendelson-Forman 2006; Robinson 1996). A central critique of the 1994 Interim Police Security Force and the subsequent Haitian National Police focuses on the fact that the model for these projects was military. Members of the former Armed Forces constituted a large portion of the IPSF and HNP, and they retained their military networks of corruption, excessively forceful tactics, and even their location in former barracks (Mendelson-Forman 2006, 23). Says Mobekk, “the change in uniform did not constitute a change in mentality, and six days of retraining could not achieve that objective” (Mobekk 2001, 103). Furthermore, even though some ex-soldiers were vetted from the new police force, the geopolitical focus on fighting narco-trafficking continued to favour military approaches. The United States Department of Justice arranged training of Haitian police officers at military camps in the United States, where the use of military-style techniques was emphasized (Beer 2004, 167).

This approach to international security issues became more prominent with the establishment of a regional police-training program for counter-terrorism, based at the U.S.-run International Law Enforcement Academy (ILEA) in El Salvador.⁵ Police scholar David Bayley argues that the “ILEA paradigm” reverses a fundamental condition for successful police reform, absent in Haiti: prioritizing local security commitments over international ones (Bayley 2006, 72). Some CivPol initiatives promoting “community policing” techniques acknowledged this view, but faced conflicting mandates among international actors and a short time-frame for overcoming the legacy of police brutality (Mobekk 2001, 105; Beer 2004). Recent promotions of community policing appear more promising, although they remain hampered by inadequate local involvement and residents’ ongoing fear of violence by police and gang leaders (International Crisis Group 2007, 9).

The pragmatic limitations of international involvement in Haiti also detracted from long-term institution building. International donors’ excessive focus on quantitative “outputs” skewed past interventions toward training a specific quota of police to the exclusion of developing the judicial system (Mobekk 2001, 107). Judicial reform had already stalled in the flawed National Truth and Justice Commission (CNVJ 1994-1996, addressing the years 1991-1994): documents remained classified, testimonies were documented but with no subsequent prosecution, the report circumvented the judicial institution altogether, amnesty deals were struck with former military members, and funding was focused on training judges rather than pursuing the actual cases at hand (Brody 1999). Full investigations of several of the most politicised arrests and killings since 2000 remain pending (International Crisis Group 2007, 12). Without coordinated and substantial reform and resources for the judicial and penal systems in Haiti, the increased ability of the police to arrest does little to provide stability (Mendelson-Forman 2006, 20). The constant arrival of deportees from abroad adds to the pressure on these systems (Mulet 2007). Short-term project budgets, the onset of donor fatigue, and lack of consistency in the objectives of foreign governments continue to be cited as challenges (von Einsiedel and Malone 2006; Beer 2004; Mobekk 2001). As a result, substantial

institutional reform is diluted by “the transaction costs of demilitarising the police” (Call 2003, 860).⁶

INTERNATIONAL INTERVENTIONS: FACTORS OF DESTABILIZATION

Other sources of international funding can create further tensions and can undermine national institutions; civil society groups are one example of such a funding source. These groups are often at the forefront of grassroots development efforts, but international funding of civil society in Haiti must be assessed critically, for it can exacerbate political divisions. Given the limited scope of the state in Haiti, the title “NGO” or “civil society” encompasses a wide range of organizations, and donors often prefer to fund such groups in order to avoid state corruption (Malone 2006, 167). The increase in “democracy promotion” funds provided by international donors (government and otherwise) has strengthened the political roles of non-governmental organisations in Haiti (Shamsie 2004). The “Democracy Enhancement” project of the National Endowment for Democracy in 1990-1992 was the prototype (Robinson 1996, 288). Its later incarnations were the Democratic Convergence, the opposition coalition around the 2000 elections, and the Group of 184, a civil society umbrella allied with the elite class; both received funds from the International Republican Institute (Engler and Fenton 2005, 51). Fatton argues that external funding of the Democratic Convergence allowed the opposition to reject the compromises Aristide offered in his second mandate (Fatton 2002, 177). Remittances from the Haitian diaspora are a further source of revenue for civil society groups and individuals (Prest 2005, 21-22). While most remittance funding is directed to individuals or humanitarian causes, partisan divisions among Haitians living abroad shape their organization and involvement in international policy discussions (Council on Hemispheric Affairs 2006). The benefits of both development projects funded by NGOs and of remittances in Haiti may not outweigh the potential for international support for partisan elements of civil society to foment internal threats (real or perceived) to the Haitian state and distort the incentives for stability and transparency, thus encouraging *la politique du ventre*.

A second international source of funds and opportunities comes from the trans-national drug trade, which creates incentives and resources that undermine law enforcement. Suspected ties to organized crime operations are especially common against the HNP. These well-funded and well-connected “spoilers” – those who, “because of their interest in retaining or maximising power, income, authority, and position do not want the transition to succeed” (International Crisis Group 2005, 4) – are considered a primary destabilizing factor in the police reform project (Donais 2005). According to Carey, higher salaries for police have not mitigated the drug-related corruption and attrition rates from the police force, suggesting that organized crime offers similar or better prospects (Carey 2005, 354). In 2007, the chief of procurement for the HNP and 252 other officers were removed from their jobs for illegal ties (Mulet 2007). As long as revenues for the drug trade continue, however, “spoilers” of law enforcement

attempts will remain active, whether inside or outside the police force.

The overall effect of international influences on the rule of law in Haiti is contradictory. Official interventions and projects support institutional practices superficially, but do not counter the sources of power and resources that are embedded inside and outside the state. This stance, combined with the influx of funds from organized crime, diaspora contributions, and NGO funding, perpetuates the impotence of the Haitian state. Some advances in the administrative capacity of state institutions have been achieved under Préval's government (International Crisis Group 2007). Without ongoing and meaningful incentives to compromise and share power, however, Carey's prediction is persuasive:

Those coming to power with fear for their security will logically attempt to control the police and paramilitary forces to obtain such political security. To finance that security, corruption becomes necessary, which induces a perverse cycle, where corruption leads to a politicized state, based on corrupt police and even less accountable and more corrupt paramilitary forces (Carey 2005, 355).

DEMOCRATIC DIVIDENDS? PROSPECTS FOR SECURITY WITH PRÉVAL

In 1991, Haiti's first elected government and its promises to respond to the poor majority were not able to overcome the pressures of the predatory state and rule by gunpoint. Thirteen years after the first police reform project, similar dynamics of politicization and corruption continue to distort the rule of law in Haiti. Attempts to "reform" or "reconstruct" the rule of law in Haiti have relied on superficial institutional conditions and have not substantially confronted the politicization of the state's use of violence against its citizens. Donais warns, "one of the clear lessons from the Bosnian experience is that placing responsibility for depoliticizing the police in the hands of those with a demonstrable interest in a politicized police force is a sure recipe for failure" (Donais 2005, 283).

This article has argued that although the outright elite-military collusion of the Duvalier era has passed, exploitative alliances between the politically or economically powerful and armed groups continue to undermine the rule of law. The instrumentalization of local gang-run protection rackets by political factions and organized crime reduces the potential for the national government to develop a neutral, reciprocal relationship with communities through social services or functional police programs in their areas. Furthermore, international actors promote "stabilization" within the bounds of low-intensity democracy and regional concerns about immigration and drug traffic. "Visible" forms of governance and policing, such as training programs and arrests, are given more priority than longer-term strategies for economic, social or institutional development.

These conditions present serious obstacles and limitations to Préval's government and to MINUSTAH. Although an electoral mandate alone does not ensure the legitimacy of the government, after one year in office, Préval still has relatively broad popular and international support, which enables his government

to resist partisan or international pressures and instead focus the approach to security and policing on concrete local and institutional needs. A measure of stability has indeed been achieved in the most violent areas, such as Cité Soleil, illustrated by the fact that foreign leaders (including Prime Minister Stephen Harper in July 2007), NGO staff, and local residents can safely circulate in previously hostile areas, and also by a notable decrease in shootings and kidnappings (*The Economist* 2007; International Crisis Group 2007). Nevertheless, this is unlikely to alter Haitians' perception of MINUSTAH as an "imposed" mission (Talentino 2007) or their resentment sparked by the excessive use of force in densely populated neighbourhoods (Pierre and Sprague 2007). While a reduction in violence allows daily life and local development projects to operate more smoothly, it also reveals the immediacy of the country's economic needs. In the words of one Cité Soleil resident, "It's true the security situation has considerably improved, but you can't eat security. You need food and jobs and schools" (Delva 2007).

Several initiatives could be extended and broadened; for example, disarmament and re-integration programs must do more than provide money for weapons, they must also address the social and political functions of gangs, distorted though they are, for the members and for the neighbourhoods where they operate. The need to demonstrate "early wins" (Mulet 2007) and "peace dividends" (Howland 2006) in poor neighbourhoods must imply more than token projects or compensation for past aggression. Incentives for disarmament can be collective and meaningful to participants, such as neighbourhood peace pacts that set conditions of non-violence for education projects (Boukman 2007). Furthermore, stabilization efforts are often presented internationally as an essential precursor to foreign investment in Haiti, including for example, new apparel factories and tourism facilities (Inter-American Dialogue & FOCAL 2005). The need for more economic activity and less crime is undeniable, but low-wage jobs and the privatisation of state enterprises (most recently the public telephone company, Téléco) are inadequate solutions for entrenched poverty. Private sector development efforts must be more comprehensive and sector-specific in order to collaborate with, rather than detract from, public institutions, service delivery systems, and financial resources (Inter-American Dialogue & FOCAL 2005). A more equitable society may be a distant goal, but redistributive policies and social services should neither be stalled until after "stability" is consolidated nor sidelined by efforts to attract foreign investment with cheap labour.

MINUSTAH has neither the time nor the capacity to be a comprehensive development program, despite Secretary-General Ban Ki-moon's recent statement in support of a one-year extension of the mission (Jacobs 2007). Nevertheless, donor countries could supply more police resources and personnel to the mission, rather than more military resources. Even though international police resources are much scarcer than military personnel, it is crucial that the military tone and role of the mission be scaled back. The recent announcement that Canada will deploy correctional staff to the mission suggests that this emphasis is indeed shifting (Government of Canada 2007). Donor governments could reduce the perception that they are in Haiti for self-serving ends by

demanding better implementation and accountability of the MINUSTAH's human rights obligations (Howland 2006), and faster, more comprehensive cancellation of Haiti's external public debts (Kar and Ricker 2006). In the long term, the national government must use its resources to persuade its citizens and its police officers that participating in economic and political projects and institutions will bring them more benefit than acting as agents of coercion for rival factions, both locally and internationally.

NOTES

¹ According to the International Crisis Group (2007b: 1), there are presently 5,500 inmates, including over 2,000 in the National Penitentiary; only 10% have been convicted. USAID's audit of its justice reform project (2007: 8) finds that the average time in pre-trial detention is increasing -- in 2006 it was 408 days.

² According to GDP and employment statistics, it is reasonable to assume that much of these funds come from the informal economy and from remittances sent by Haitians living abroad.

³ There are still ideological or partisan delineations to this dynamic, but former pro-Lavalas gang alliances, such as in Cité Soleil, have split into rival criminal factions. The International Crisis Group reports that thirty-four gangs remain active in the zone (2007, 2). Paramilitaries often control rural areas on behalf of elite strongmen, while urban gangs align with different factions of Lavalas, but narco-trafficking and organised crime are eroding even these.

⁴ Robinson's definition is, "a system in which a small group actually rules and mass participation in decision-making is confined to leadership choice in elections carefully managed by competing elite" (1996: 49).

⁵ The International Law Enforcement Academies are a network of training institutions designed "to combat international drug trafficking, criminality, and terrorism through strengthened international cooperation," according to the State Department website. Established under Clinton, the Latin American location was opened in El Salvador in 2005, increasing the dominance of such international security issues on the agendas of national governments and police. For the official mandate, see <http://www.state.gov/p/inl/ilea/>

⁶ Call discusses the post-war case of El Salvador; in Haiti the ambiguous nature of the "transition," combined with similar levels of past repressive army-police integration, makes this dynamic more entrenched.

REFERENCES

- Associated Press. 2006. Haiti appoints commission to oversee U.N.-backed gang disarmament. *International Herald Tribune*. September 7.
http://www.iht.com/articles/ap/2006/09/08/news/CB_GEN_Haiti_Gang_Disarmament.php
- Bayley, David H. 2006. *Changing the Guard: Developing Democratic Police Abroad*. Oxford: Oxford University Press.
- Beer, David. 2004. Peacebuilding on the Ground: Reforming the Judicial Sector in Haiti. In *Building Sustainable Peace*. Ed. Tom Keating and W. Andy Knight. New York: United Nations University Press.
- Boukman, Samba. 2007. In Step with Haiti's Samba Boukman (Interview by Daniela Bercovitch). *Comunidad Segura*. July 6. <http://www.comunidadsegura.org/?q=en/node/34855>
- Brody, Reed. 1999. International Aspects of Current Efforts at Judicial Reform: Undermining Justice in Haiti. In *The (Un)Rule of Law and the Underprivileged in Latin America*, eds. Juan E. Méndez, Guillermo O'Donnell and Paulo Sérgio Pinheiro. Notre Dame, IN: University of Notre Dame Press.
- Call, Charles T. 2003. Democratisation, War, and State-Building: Constructing the Rule of Law in El Salvador. *Journal of Latin American Studies* 35: 827-862.
- Carey, Henry F. 2005. Militarization Without Civil War: The Security Dilemma and Regional Consolidation in Haiti. *Civil Wars* 7(4): 330-356.

- Chomsky, Noam, Paul Farmer, and Amy Goodman. 2004. *Getting Haiti Right This Time: The US and the Coup*. USA: Common Courage Press.
- Council on Hemispheric Affairs. 2006. Canada's Aid to Haiti: Commendable, or Making Amends for a Discredited Anti-Aristide Strategy? *Reports: Haiti*. March 30. <http://www.coha.org/2006/03/30/canada%E2%80%99s-aid-to-haiti-commendable-or-making-amends-for-a-discredited-anti-aristide-strategy/>
- Dade, Carlo and John Graham. 2004. *The Role for Canada in Post-Aristide Haiti: Structures, Options, and Leadership.* Ottawa: Canadian Foundation for the Americas (FOCAL) / House of Commons Committee on Foreign Affairs and International Trade. http://www.focal.ca/pdf/haiti_post_aristide.pdf (accessed November 30, 2006.)
- Delva, Joseph Guyler. 2007. Haiti slum residents enjoy new peace, want more. *Reuters*. April 24. http://www.streetnewsservice.org/index.php?page=archive_detail&articleID=1135
- Donais, Timothy. 2005. Back to Square One: The Politics of Policing Haiti. *Civil Wars* 7(3): 260-287.
- Duffield, Mark. 2001. *Global Governance and the New Wars: The Merging of Development and Security*. London: Zed Books.
- The Economist*. 2007. Haiti: A small success for the UN, 2 August. http://www.economist.com/world/la/displaystory.cfm?story_id=9587689
- Engler, Yves and Anthony Fenton. 2005. *Canada in Haiti: Waging War on the Poor Majority*. Vancouver, BC: Red / Fernwood Publishing.
- Erikson, Daniel P. 2005. Haiti after Aristide: Still on the Brink. *Current History*. February: 83-90.
- Fatton, Robert. 2002. *Haiti's Predatory Republic: The Unending Transition to Democracy*. Boulder: Lynne Rienner Publishers.
- Government of Canada. 2007. Canada Announces Deployment of Correctional Services Officers to Haiti. *Canada News Centre News Release 101*. July 30. <http://news.gc.ca/web/view/en/index.jsp?articleid=343079>
- Haiti Information Project. 2007. US Embassy Acknowledges Excessive Force by the UN. *Global Policy Forum*. January 25. <http://www.globalpolicy.org/security/issues/haiti/2007/0125force.htm>
- Haiti Information Project. 2005. The UN's Disconnect with the Poor in Haiti. December 25. http://www.haitiaction.net/News/HIP/12_25_5/12_25a_5.html
- Harvard Law Student Advocates for Human Rights. 2005. *Keeping the Peace in Haiti? An Assessment of the United Nations Stabilization Mission in Haiti Using Compliance with its Prescribed Mandate as a Barometer for Success*. Cambridge: Harvard University and Rio de Janeiro: Centro de Justicia Global. <http://www.globalpolicy.org/security/issues/haiti/2005/0300keeppeace.pdf>
- Howland, Todd. Peacekeeping and Conformity with Human Rights Law: How MINUSTAH Falls Short in Haiti. *International Peacekeeping* 13(4) 462-476.
- Inter-American Dialogue and the Canadian Foundation for the Americas (FOCAL). 2005. *The Role of the Private Sector in Rebuilding Haiti: Dialogue Document from a Roundtable*. September 9-10. http://www.focal.ca/pdf/haiti_dialogue.pdf
- International Crisis Group. 2007. *Consolidating Stability in Haiti*. Latin America / Caribbean Briefing No. 21. July 18.
- International Crisis Group. 2007b. *Haiti: Prison Reform and the Rule of Law*. Latin America/Caribbean Briefing No. 15. May 4.
- International Crisis Group. 2006. *Haiti After the Elections: Preval's 100-day challenges*. Latin America/Caribbean Briefing No. 10. May.
- International Crisis Group. 2005. *Spoiling Security in Haiti*. Latin American/Caribbean Report 13. May.
- Jacobs, Stevenson. 2007. UN Chief Backs Haiti Peacekeepers. *Associated Press / Chicago Tribune*. August 2. <http://www.chicagotribune.com/news/nationworld/sns-ap-haiti-un-peacekeepers,1,1168650.story?ctrack=1&csset=true>
- Joseph, Mario and Brian Concannon Jr. 2007. *Haiti, MINUSTAH and Latin America: Solidaridad?* International Relations Center Americas Program. April 9. <http://americas.irc-online.org/am/4140>
- Kar, Debayani and Tom Ricker. 2006. IDB Debt Cancellation for Haiti. *Foreign Policy in Focus*. December 7. <http://www.fpif.org/fpiftxt/3768>

- Kolbe, Athena and Royce Hutson. 2006. Human rights abuse and other criminal violations in Port-au-Prince: a random survey of households. *The Lancet* 368: 864-873.
- Kumar, Chetan. 1998. *Building Peace in Haiti*. Boulder: Lynne Rienner Publishers / The International Peace Academy.
- MacKay, Peter. 2007. *Haiti: Speech by Minister of Foreign Affairs of Canada at launch of joint Brazil-Canada seminar on MINUSTAH*. February 5. <http://www.reliefweb.int/rw/rwb.nsf/db900sid/EGUA-6Y6THG?OpenDocument>
- Macnamara, W. Don. 2005. Haiti: An Opportunity for Canada to Apply the “3-D” Concept. *Policy Options*. February: 63-67. <http://www.irpp.org/po/archive/feb05/macnamara.pdf>
- Mendelson-Forman, Johanna. 2006. Security-Sector Reform in Haiti. *International Peacekeeping* 13(1): 14-27.
- Mobekk, Eirin. 2001. International Involvement in Restructuring and Creating Security Forces: The Case of Haiti. *Small Wars and Insurgencies* 12(3): 97-114.
- Mulet, Edmond. 2007. *The Journey since the Deployment of MINUSTAH since June 2004 and the Main Upcoming Challenges*. Presentation at the Canadian Institute of International Affairs / Foreign Affairs Canada, June 1, Ottawa, Canada.
- Pierre, Wadner and Jeb Sprague. 2007. Haiti: Poor Residents Describe a State of Siege. Inter-American Press Service. *Inter-Press Service News Agency*. February 28. <http://ipsnews.net/news.asp?idnews=36772>
- Prest, Stuart. 2005. Uprising in Haiti: The Criminal Threat to Canada. *Country Indicators for Foreign Policy*. Ottawa: Carleton University. June.
- Robinson, William I. 1996. *Promoting Polyarchy: Globalization, US Intervention, and Hegemony*. Cambridge: Cambridge University Press.
- Robinson, William I. 2003. *Transnational Conflicts: Central America, Social Change, and Globalization*. Los Angeles: Verso.
- Rodgers, Dennis. 2004. Old Wine in New Bottles or New Wine in Old? Conceptualising Violence and Governmentality of Contemporary Latin America. *Development Studies Institute*. Crisis States Discussion Paper No. 6.
- Savenije, Wim and Chris van der Borgh. 2004. Youth gangs, social exclusion and the transformation of violence in El Salvador. In *Armed Actors: Organized violence and state failure in Latin America*, eds. Kees Koonings and Dirk Kruijt. London: Zed Books.
- Shamsie, Yasmine. 2004. Building low-intensity democracy in Haiti: The OAS contribution. *Third World Quarterly* 25(6): 1097-1115.
- Snyder, Richard. 1992. Explaining Transitions from Neopatrimonial Dictatorships. *Comparative Politics* 24(4): 379-399.
- Talentino, Andrea Kathryn. 2007. Perceptions of Peacebuilding: The Dynamic of Imposer and Imposed-Upon. *International Studies Perspectives* 8: 152-171.
- Tilly, Charles. 1985. War Making and State Making as Organized Crime. In *Bringing the State Back In*, eds. Peter B. Evans, Dietrich Rueschemeyer, and Theda Skocpol. Cambridge: Cambridge University Press.
- Urdal, Henrik. 2004. The Devil in the Demographics: The Effect of Youth Bulges on Domestic Armed Conflict, 1950-2000. *World Bank*. Social Development Papers: Conflict and Reconstruction Paper No. 14.
- United States Agency for International Development (USAID), Office of Inspector General. 2007. *Audit of USAID/Haiti's Justice Program*. Audit Report No. 1-521-007-08P. April 24. San Salvador, El Salvador. <http://www.ijdh.org/pdf/headline5-4-07.pdf>
- United Nations Department of Public Information. 2007. Security Council Extends Haiti Mission Until 15 October, Unanimously Adopting Resolution 1743. February 15. <http://www.un.org/News/Press/docs/2007/sc8956.doc.htm>
- Valdés, Juan Gabriel. 2005. Uso de la fuerza, elecciones y desafíos de la MINUSTAH. Interview by David Alvarez Veloso. *Revista Fuerzas Armadas y Sociedad* 19(1): 3-15.
- Von Einsiedel, Sebastian and David M. Malone. 2006. Peace and Democracy for Haiti: A UN Mission Impossible? *International Relations* 20(2): 153-174.
- Willett, Susan. 2005. New Barbarians at the Gate: Losing the Liberal Peace in Africa. *Review of African Political Economy* 106: 569-594.

This page has been intentionally left blank.

EFFECTS OF FINANCIAL SECTOR FOREIGN DIRECT INVESTMENT ON ACCESS TO CREDIT FOR SMALL AND MEDIUM-SIZED ENTERPRISES IN EMERGING MARKET ECONOMIES: A CRITICAL COMMENTARY

Manuela W. Armenta

Norman Paterson School of International Affairs, Carleton University

This article provides a critical commentary on the literature surrounding the effect of FSFDI on the SME market for access to financing. It begins with a critique of a study by Clarke et al. (2002), on the basis of both methodological and substantive issues. The author goes on to argue that barriers to credit for SMEs often derive from the banking sector's requirement for quantifiable, standardized forms of information that many smaller, semi-formal firms do not have available to provide. For this reason it may be crucial to examine the basis upon which credit is granted rather than simply the quantity of credit lent or its price. More specifically, it is important to determine whether credit is granted using transaction-based or relationship-based methods. The author goes on to suggest that competition, rather than foreign investment *per se*, may be a crucial variable in enhancing the likelihood that domestic banks will service informationally-opaque SME niches. Moreover, the author points out that domestic banks may be poorly managed and subject to capture by special interests, which calls into question how desirable they are as alternative lenders, notwithstanding their willingness to engage in relationship-based, rather than transaction-based lending. Improvements to internal credit and management practices are therefore necessary to increase the ability of these institutions to service all clients, including SMEs, effectively. Technical assistance and capacity building that assist SMEs in overcoming the problem of information opacity may also reduce the likelihood of SMEs being phased-out of existing relationships and greatly enhance access in the long-run.

INTRODUCTION

The 1990's witnessed the lifting of investment restrictions on financial sector foreign direct investment (FSFDI) in many Latin American emerging market economies (EMEs) leading to unprecedented levels of foreign ownership in the sector. As a result, both the structure and stability of affected domestic banking industries changed. Numerous studies focus on the aggregate effects that FSFDI has had on banking sector efficiency and stability; it remains less clear whether the effect of FSFDI has been beneficial or harmful for small- and medium-sized enterprise (SME) access to credit. This article re-examines what is already known about financial systems and the credit constraints faced by SMEs. It then provides comments on existing empirical studies surrounding the positive and negative effects of FSFDI on the ability of SMEs to access credit. The article critiques Clarke et al. (2002) by arguing that their results may not be generalized. Clarke et al. (2002) argue that SMEs are, on balance, better off due to FSFDI because increasing efficiency within the industry assists all firms, including SMEs. Moreover, they argue that domestic banks may develop SME niches to

capitalize on the SME lending opportunities foreign banks eschew. By contrast, this article suggests that the negative short-run effects of FSFDI may overwhelm or diminish the positive medium- to long-term benefits discussed in the literature. Three possible short-run effects are explored: (1) barriers to credit faced by SMEs increase, *ceteris paribus*, because foreign-owned banks focus on transaction-based lending; (2) domestic banks only develop niches that cater to SMEs under certain conditions and may be sub-optimal alternative lenders; and (3) standing lending relationships are harmed or severed and associated costs may be high. This article is not arguing that existing empirical studies are incorrect; rather, it critically examines the assumptions and the data used in some studies to determine how representative their conclusions are for the majority of SMEs in the Latin American region. This article provides a critical commentary on the existing literature in an effort to demonstrate how the absence of certain conditions may lead FSFDI to have net negative rather than net positive effects on SME access to credit.

Methodologically, this article begins with a literature review, examining established theory and evidence on the effects of FSFDI on SMEs' access to credit in Latin American markets. Using Mexico as a mini case study, the article demonstrates how an increase in FSFDI may, under certain conditions, result in a reduction in SMEs' access to credit. The article concludes with a call for further research into the conditions under which the 'net' effect of FSFDI is likely to be positive on SMEs' access to credit in Latin America, and some reflections on the implications of such research for policymakers.

DEFINITIONS, BACKGROUND, AND MOTIVATION

FSFDI Background

According to a report by a working group at the Bank for International Settlements' Committee on the Global Financial System, FSFDI is defined as "international investment that reflects the objective of a resident entity in one economy obtaining a lasting interest in a firm resident in another economy. Hence, it [FSFDI] refers to control rather than a specific form of financing" (Bank for International Settlements 2004, 7). In the Latin American context, FSFDI grew rapidly during the 1990s following the liberalization of highly inefficient financial systems in countries including Mexico, Argentina, Chile and Brazil. The financial markets in these states had long been distorted by interest rate ceilings and directed lending (Todaro 2003, 746). In the wake of economic liberalization and debt crises, banks were often illiquid or insolvent and holding sizable bad debt. With severely undercapitalized banks, governments hoped foreign investment would bring much needed liquidity. FSFDI, via mergers and acquisitions of troubled domestic banks, followed. At the industry level, the theoretical benefits included recapitalization, stability, efficient institutions and spillover to domestic banks (Cardenas et al. 2003, 1-2). It is, however, not guaranteed that all such industry-specific effects occur. It is also not clear that they directly or indirectly support greater access to credit for SMEs. On the contrary, it is possible that in some cases FSFDI simply exacerbates the barriers

to access to credit already faced by SMEs; this question is examined in further detail later in the article.

What are SMEs and why do they matter?

Admittedly, the term SME is somewhat vague; the academic literature contains a number of different definitions of what constitutes a small- or medium-sized enterprise. Various governments define the term differently as well, further complicating matters. In general however, an SME can be described as a business with a limited workforce and payroll, and with limited access to international markets. They are distinguished from micro-enterprises both in size and in degree of formality, though here again boundaries are fluid between categories. In a study conducted for the International Development Research Centre, Lerchs (2002) identified a range of different government definitions. Egypt considered SMEs to include firms employing between 5 and 49 people, while the European Union (EU) included firms employing up to 250 employees. To highlight the degree of confusion, the Mexican government actually employs two different formal definitions of the term, including all firms with up to 250 or 500 employees.

Understanding the characteristics of SMEs and what they offer to an economy explains why barriers to their growth and expansion should be minimized. Lerchs (2002) suggests that small firms in the developing world are often run by individuals living just above the poverty line. The same study suggests that average income levels tend to increase with the size of the firm (Lerchs 2002). These firms are also generally lower-productivity firms, meaning that they may provide relatively lower-skilled employment opportunities, thus potentially promoting inclusive growth and equity (Saavedra 2003). Mexican census data suggest that these firms account for 26 percent of employment. Theoretically, if one could increase productivity among such firms they could become more profitable, thereby creating new employment opportunities, increasing average incomes, and generating economic growth. In smaller firms it is plausible that these benefits would accrue to employees, who on average may be relatively poorer than employees of larger firms. Although tentative, these theoretical linkages suggest a strong case for supporting SMEs as part of a broader pro-poor development strategy.

SMEs and Access to Credit

Small firms have the potential to increase productivity via innovation and expansion. Firm theory, developed by Schumpeter (1968), argues that credit leads to increased firm productivity by enabling the acquisition of technology, implements and additional labour.¹ In addition, Claessens (2006) recently noted that credit makes it easier for firms to capitalize on new growth opportunities. Although factors in the business environment—ease of firm entry and exit, property rights and contract enforcement to name just a few—are important to the survival and development of SMEs, finance accelerates growth by providing the means for firm start-up, expansion, and efficient resource allocation (Klapper, Laeven, Rajan, 2004 as cited in Claessens 2006, 210). Empirical evidence also tells us that barriers to credit negatively affect the growth of SMEs more than the

growth of large firms (Beck et al. 2003). Further testing is required to confirm the theoretical linkages presented above; however, assuming they hold, and taken together with the “pro-poor” linkages discussed earlier, there is a strong case for developing policy initiatives aimed at reducing barriers to credit for SMEs.

Barriers to Credit

Aside from the possible effects of increased FSFDI – whether positive or negative – SMEs in general face considerable barriers to gaining access to credit. A crucial barrier relevant to the discussion of FSFDI and SME credit is information opacity. SMEs are what Berger (2001) refers to as “informationally opaque”. Not only are SMEs, by nature, relatively more risky to finance than large firms, information opacity, or a lack of concrete, accessible and objective information makes it difficult for banks to accurately assess the risk involved with a particular project. Holding constant the level of financial development between states, SMEs tend to have less quantifiable information and the information they do have is often less verifiable in the objective, standardized forms familiar to financial institutions (i.e. credit bureau scores and reliable financial statements). Consequently, the bank has less information about the project than the firm; in economic terms this is known as information asymmetry. This can affect the ability of small firms to obtain financing if, due to information asymmetry, banks compensate for risk by increasing the cost of borrowing (especially if it is to the point where SMEs cannot afford a loan). In some cases, banks may even rule SME lending out completely due to risk and the potential associated costs (Claessens 2006, 222).

EFFECTS OF FSFDI ON SMES’ ACCESS TO CREDIT IN LATIN AMERICA; A REVIEW OF THEORY AND EMPIRICAL EVIDENCE

FSFDI and SMEs – Net Effects

Numerous studies examine the effects of FSFDI on domestic economies and some look specifically at its effect on SMEs. This article argues that the question of net benefit versus net harm depends on a number of context-specific factors and must be analyzed on a case by case basis. For example, the net effect may change depending on whether the analysis focuses on short-term results—particularly during the immediate aftermath of a significant increase in FSFDI— or long-term effects that emerge years, or perhaps even decades after the influx of foreign investment. Further, findings regarding the net effect may in fact change from study to study, depending on how widely one defines the effects. Indeed, one serious shortcoming within the current literature is that most studies rely mainly on econometric analyses; consequently, some effects doubtlessly remain unobserved and unexplained due to data constraints and analytical insensitivity to less quantifiable externalities. Analyzing outliers, both those that perform better as well as those that perform worse than predicted by the model is necessary, especially over time, to determine the context-specific factors that enhance rather than diminish access to credit for SMEs.

Clarke et al. (2002) argue that SMEs are, on balance, better off due to FSFDI because increasing efficiency within the industry assists all firms, including SMEs. Moreover, they argue that domestic banks may develop SME niches to capitalize on the SME lending opportunities foreign banks eschew. Their study looks at the impact of foreign bank presence on firms' access to credit, disaggregating firms according to size to determine if size differences predict different outcomes for access. Using survey data from 3,000 enterprises in 36 developing and emerging economies, the authors use cross-section regression analysis to test whether access to credit is associated with foreign bank presence. To conduct the analysis, they use two dependent variables: borrowers' perceptions of interest rates and access to long-term credit as constraints to growth, controlling for macroeconomic, institutional, and firm-specific factors. They hypothesize that if the potential advantages of foreign bank entry in developing countries outweigh the tendency of foreign-owned banks to abstain from SME lending, then all borrowers, including SMEs, should rate both variables measured as lesser constraints to growth in states with relatively more foreign ownership. The analysis is meant to capture the "net effect" (direct plus indirect) of FSFDI. The regression results indicate that all firms, including SMEs, rank interest rates and long-term credit more favourably in states with higher foreign ownership. These results suggest that, at minimum, SMEs are not harmed by foreign ownership, and may benefit in some cases. Clarke et al. provide two potential causal explanations for the positive finding. First, they cite cross-section evidence that overall sector efficiency (increased competition, stability, efficiency, technological and skill spillover) leads to lower interest margins and overhead costs. Moreover, if these efficiency gains lead to an overall expansion in total lending, then even if foreign banks lend a lower percentage of funds to SMEs, the actual dollar value of funds lent could still be rising. Second, if foreign banks compete with local domestic banks for larger clients, then in response local banks may shift into targeting the SME market.

Both of these theoretical explanations are valid and even plausible, however, neither of them rules out the possibility that many SMEs are harmed by foreign investment. Effectively, the theory and the concurrent results at best suggest that in some cases harm may be offset by benefits for sampled firms. Therefore, the firms may experience harm and benefit concurrently and one more than the other in short-run versus the long-run. This article is not suggesting that Clarke et al. are incorrect in arguing that expansion of overall domestic lending has led to a larger aggregate quantity of funds to lend and reduced interest rates for firms with loans.

Any insight, however, that can be gleaned from the Clarke et al. study can only be applied to the kinds of firms surveyed in the study; recalling that the study only surveyed firms that already had access to credit makes it difficult for the study to shed light on the effect of FSFDI on all SMEs. By including only borrowers, the findings are limited to firms that had successfully obtained and/or maintained credit following the structural changes brought about by FSFDI. Arguably, such a selection bias excludes the firms that are most likely to complain that they are not, from a net perspective, better off, i.e. firms that were either unable to get credit, or that lost access in the FSFDI-infused market. At

minimum, such findings, limit how far the results may be generalized and suggest that the authors have not gone far enough in their testing.

Moreover, it is possible that SMEs do not rely upon long-term loans in the same way that larger firms do; thus the results on this variable may also overstate the benefits to SMEs if this type of credit is less relevant to their growth.

This next section explores why it is plausible that many firms without prior access to credit likely suffered net harm, especially in the short-term. The critical issue is that the information opacity problem remains regardless of whether the overall quality and quantity of lending allocated to SMEs expands. As was discussed earlier, foreign-owned banks rely relatively heavily upon quantifiable data to assess and grant credit and SMEs can not often provide such data (Cardenas et al. 2003, 2). Foreign firms have standards and credit analysis techniques that cannot be circumvented. Therefore, only SMEs that can meet these requirements will have access to credit. This article therefore argues that foreign firm investment raises the already substantial barriers to credit, *ceteris paribus*, for those firms that are *ex ante* outside the borrowing pool.

THREE ISSUES REQUIRING FURTHER CONSIDERATION: INFORMATION, SME NICHES AND ADJUSTMENT COSTS

Information

As noted, SMEs are informationally opaque, a fact limiting their ability to access formal credit instruments. Berger et al. (2001) suggest that one way to circumvent information opacity is via 'relationship lending' wherein the information gathered includes both quantifiable and less quantifiable data. Rather than relying on financial statement information alone, the lender gathers qualitative information about the firm from interviews with the owner and contact with other businesses including suppliers, customers and the local community with whom the firm has a long-standing reputation-based relationship (Berger et al. 2001, 2129). Over time this relationship grows and becomes more valuable as more data is accumulated.

Relationship lending is particularly beneficial because the information collected is used to make decisions on loan renewals, credit increases and renegotiations that tend to become increasingly favourable to the borrower in the medium- to long-term (Berger et al. 2001, 2130). As the relationship strengthens, information asymmetry falls and as a result, so does the risk borne by the lender. Consequently, the price of borrowing, often reflected in the interest rate, falls (Berger and Udell 1995). Berger et al. (2001) also cite a number of empirical studies which found that collateral requirements also tend to fall over time and that the quantity of credit provided tends to rise.

Transaction-based lending, conversely, relies relatively heavily upon readily available, quantifiable data that can be easily verified by financial statements and transparent agencies. Due to their size and level of sophistication, foreign banks may rely upon transaction-based lending techniques rather than relationship-based approaches. Theoretically, the use of transaction-based lending is one of the reasons why asset quality improves in banks acquired by

foreign owners. Foreign owners transfer credit analysis and risk management systems as well as skilled managers from their home markets which are more often than not, developed country economies (Schulz 2006, 12). These systems and staff are sophisticated relative to their counterparts in newly liberalized financial sectors and evaluate credit using computers and credit scoring techniques that expand loan volumes while minimizing risk (Schulz 2006, 26).

Conversely, transaction-based lending is not conducive to absorbing local information and assigning it value, whereas, relationship-based lending is; this is why lenders willing to lend on the basis of relationship-based approaches can be critical to SMEs' prospects for accessing credit.

To further compound the factors discussed above, foreign banks are often headquartered in markets that are very different from host EME markets. Language, cultural, supervisory and regulatory differences combine to increase information asymmetry and the ease of monitoring loans and therefore the costs of lending to firms that rely on relationship-based lending techniques, namely SMEs (Berger et al. 2001, 2131-2). In short, there are a number of structural and cultural obstacles which limit the willingness and the ability of foreign banks to lend to SMEs. FSFDI thus arguably increases the barriers already faced by informationally opaque firms seeking loans.

Berger et al. (2001) confirm this argument. Using firm and bank data from the Central Bank of Argentina, they test whether large banks, foreign-owned banks, and distressed banks face barriers in providing 'relationship lending' (Berger et al. 2001). The study uses loan size, specifically the natural log of the sum of the firm's total loans from all banks, to measure information opacity. For the dependent variables, they used dummies to indicate when a bank was large, foreign or distressed. The findings show that both the large and foreign-bank hypotheses were true; small firms were the least likely to obtain loans from either type of institution. The situation was compounded if the institution was both large and foreign (Berger, et al. 2001).

Due to constraints on firm-level data, we cannot quantify how many SMEs would be affected by these constraints; however, it is clear that affected firms would be harmed through a narrowed range of options. In short, for informationally opaque firms it is moot to question whether the efficiency gains brought by FSFDI expand the aggregate amount of credit lent by foreign-owned banks unless it concurrently lowers the barriers to access to credit. It does not matter to informationally opaque SMEs how much bigger the credit pie gets if they still cannot get a slice. Based on the analysis here, it is fair to suggest that FSFDI increases rather than decreases barriers. The only potentially positive effect for such firms is thus an indirect one; as Clarke et al. (2002) suggest, SMEs may benefit if domestically-owned banks develop niches in SME lending, in order to compensate for business lost to foreign competitors.

The Likelihood of Domestic Banks Developing Niche SME Lending Lines

This section critically examines the likelihood that domestic banks will develop SME niches which could potentially increase the chances of SMEs gaining access to credit as an indirect result of FSFDI. In addition to Clarke et al. (2002),

Berger et al. (2001) also suggest that the consolidation of the banking industry via mergers and acquisitions by large and/or foreign banks,

...may not substantially reduce the overall supply of credit to informationally opaque small business because there may be “external effects” ...That is, although relationship-based small business loans may be dropped by some large banks after M&As, other banks or non-bank lenders may pick up some of these loans if they are positive net present value investments (citation from Berger et al. 2001, 2133).

Mini-Case Study: Mexico

Though potentially beneficial to SMEs, positive “external effects” or indirect effects described may or may not actually occur. On the contrary, the emergence of SME niche lending in domestic banks depends upon a number of conditions, including the existence of a vibrant and viable domestic financial sector in search of new demand for credit. The example of Mexico demonstrates that this condition is not always met post-liberalization in EMEs. Similar to other Latin American EMEs, the post-debt crisis Mexican financial sector was severely undercapitalized. The government allowed foreign banks access to the financial sector at this time to help it recapitalize and to clean up the non-performing portfolios of illiquid banks. Although GDP started to grow again, credit to the private sector as a percentage of GDP did not (Barajas and Steiner 2002). During this “credit crunch” period many non-bank institutions and alternative financing schemes took up the excess demand (Economist, 2002). In a study looking at whether FSFDI improved the productivity of the Mexican banking system, Heiner Schulz (2006) finds that performance on the cost-side improved substantially; provisions for non-performing loans fell (Schulz 2006). Competition in the industry, however, did not improve. The result was higher interest rate margins and larger fee income to banks. In other words, foreign banks imported sophisticated credit scoring models and risk management processes to expand loan volumes while lowering their credit risk. Cutting costs by reducing risk made their business sufficiently profitable such that the foreign banks did not have to compete by lowering the prices of their lending products and services. By 2002, foreign banks controlled 85 percent of the banking industry in Mexico. The market power of foreign banks allowed them to effectively monopolize which firms obtained access to credit—more lucrative, larger borrowers that could afford to pay the fees and represented relatively low risk. As the Economist observed at the time,

Small and medium-sized businesses, to which the government of President Vicente Fox has given high priority, remain desperately short of funding. Even if entrepreneurs can get a small loan, they pay interest roughly 20 percentage points higher than the 8 percent underlying rate (Oct.12, 2002).

With only one domestic bank remaining and increased barriers to credit in place due to foreign banks’ utilization of imported credit and risk techniques, SMEs faced extremely limited alternatives in Mexico.

FSDI in this case did not foster increased competition and the entry of new domestic banks to serve SMEs. Consequently, SMEs faced increased barriers to credit due to a preference by foreign owners both to conduct transaction-based lending and maintain high interest rates and fees. On balance, the evidence suggests that SMEs in Mexico faced higher barriers to credit as a result of FSDI. The Mexican example suggests that FSDI is likely to produce pro-SME, and by extension, pro-poor, outcomes only given the existence of certain conditions; one important condition being competition. Clearly, further research is required to confirm the above insights from the Mexican example, and to identify other factors with the potential to limit the harm to SME access to credit and increase the positive indirect effects of FSDI.

This author acknowledges that Mexico is only one example among many. A number of other studies suggest that other countries have had a more positive experience. Using cross-sectional survey data collected from banks in 60 countries, Jenkins (2000) examines the rate at which commercial banks lend to small firms and their reasons for servicing this market. The study only looks at micro and small business loans; therefore, for the purpose of this paper only the data on small business loans coming from private commercial banks are discussed. In the survey a small business loan is defined by loan size and ranges from USD \$11,000-100,000. The data shows that out of the 148 banks surveyed, 88 make small business loans.² When asked about the top two reasons for making loans to this sector, 49 percent of the banks indicated profitability and 44 percent indicated “the changing market conditions and increasing competition in lending to large/medium enterprises” (Jenkins 2000, 5). The paper references how increased FSDI contributed to increasing competition in developing state banking sectors and how many domestic banks lost large clients to foreign banks and accordingly began to look for new creditworthy customers (Jenkins 2000, 5). Of the banks that did not make small business loans, 40 percent cited “higher administrative costs of making these loans,” and only 17 percent cited risk (Jenkins, 2000, 6). Finally, the study found that newer banks tended to lend a larger portion of their portfolios to small firms than banks that have been in business longer (Jenkins 2000, 8). This finding may support the contention that new banks are established to meet an emerging niche market in smaller firm lending. Jenkins cites the establishment of *Banco de la Pequene Empresa* in the Dominican Republic as an example of an institution specifically established to serve small firms (2000, 9).

Even though the Jenkins (2000) study was not confined to developing states in Latin America and did not specifically state that SMEs were being catered to because of the inflow of FSDI, the loose connection was however made. The fact that domestic banks cater to SMEs because of “changing market conditions” does provide plausible evidence that the theory may hold in some countries. However, given Jenkins’ reliance on a single cross-sectional survey of 60 countries, including developed, developing, and transitional economies, considerably more research is necessary to confirm the findings. In particular, research using panel data is necessary both to confirm that the findings hold within countries across time, and to identify any other potential outliers from the theory. This task will bring researchers closer to understanding that conditions

such as healthy competition are established to make FSFDI relatively more positive for SMEs and to avoid cases like Mexico. In-depth case studies involving the collection of firm-level data are also necessary to determine the causal linkages in the case of Mexico.

The Desirability of the Domestic Bank Offsetting Effect

If in some cases domestic banks are an alternative, it is important to examine their quality as lenders. Over the short-run, for example, it appears that domestic lenders are sub-optimal alternatives. Borrowers obtain value from strong banking relationships with lenders that know how to select profitable investments and know how to monitor projects to ensure that they remain profitable. Borrowing funds involves an on-going contract with a lender and entails obligations and commitments on both sides. Borrowers often have to meet minimum financial reporting requirements and report on their on-going position regarding cash flow, liquidity, profitability and solvency. Skilled lenders provide expertise and therefore value to their clients through this iterated relationship. If domestic banks are inefficient, poorly operated, lack human capital and are vulnerable to special interests capture, they may be unwilling or unable to effectively and objectively evaluate which borrowers are most creditworthy and develop the types of relationships discussed.

There is strong evidence to suggest that in developing countries domestic banks are relatively inefficient at allocating credit; indeed, this is one of the core arguments underpinning efforts to increase FSFDI. A number of studies employing cross-country regression analysis find foreign banks in developing countries to be more efficient than domestic banks (Schulz 2006 cites Claessens et al. 2001 and Claessens and Lee 2002). In a paper for the World Bank, International Monetary Fund and Brookings Institution, Pomerleano and Vojta (2001) argue in favour of developing states welcoming foreign bank investment; however, in doing so, they also acknowledge that:

This trend toward a two-tiered segmentation of banking services in emerging markets may result in competitive disadvantages for small and medium-size enterprises, whose growth prospects may be constrained by the service capacity of domestic banks. This issue certainly requires domestic banks to improve the capacity to serve SMEs (Pomerleano and Vojta 2001).

In other words, assuming domestic lenders are available as alternatives; SMEs may be forced early on to rely on relatively inefficient bankers. Although it may be argued that the domestic bank alternative is better than no alternative, it is critical to articulate the notable disadvantages for SMEs. Policy to improve the sound management of these institutions can be added to the previously noted need to ensure competition in the industry.

As discussed, prior to the 1990's domestic banks in emerging markets often operated under less than free market conditions. Pomerleano and Vojta (2001) outline areas where domestic banks could focus improvements including; the establishment of new policies and procedures for loan portfolio management, asset/liability management, information collection and monitoring systems as well as the recruitment and retention of appropriately trained staff.

The evidence that FSFDI results in spillover in these areas is scant and mixed; however, it suggests that spillover takes place within acquired banks rather than between foreign and domestic institutions (Schulz 2006). Though Pomerleano and Vojta (2001) note that there has been some transference in the area of management information systems, their examples cite mainly Asian cases rather than Latin American ones. Although this is not to say that such transfers do not take place in the latter region as well, important differences exist between the two regions that may limit how far such findings can be generalized. For example, foreign investors tend to control a much larger proportion of the banking sector in Latin America, thus limiting the number and capitalization of domestic banks ostensibly competing with one another for SME business. The potential for spillover effects will be a function of time and also the level of competition in the market. As a result, domestic banks may only be in a position to effectively meet the credit needs of SMEs once critical domestic banking sector reform, including institutional, regulatory and human capital development occurs. Spillover may occur in the long-run; however, this time horizon suggests that SMEs may continue to suffer from high barriers to credit which leave them capital deprived, or forced to engage in borrowing relationships with sub-optimal lenders until then.

FSFDI and Portfolio Clean-up

It has been established that FSFDI may in some cases increase the barriers faced by SMEs attempting to secure access to credit and force them to seek lending relationships with potentially sub-optimal domestic banks. This article has not, however, examined the effect of FSFDI on SMEs that may have already been customers of newly acquired or merged foreign-owned banks. The fact that foreign investment took place in Latin American EMEs post-crisis, when banks were carrying significant levels of non-performing debt on their books, leads one to expect that significant portfolio re-evaluation was necessary. Clearly, portfolio clean-up and significant risk reduction is prudent to rebuild and stabilize financial industries; however, questions remain whether credit-worthy SME customers with standing banking relationships were dropped or phased-out of foreign-owned portfolios due to an inability to meet transaction-based criteria. Further, questions remain regarding the extent to which such SMEs were able to obtain new banking relationships with remaining or *de novo* domestic banks. Further research is necessary to determine whether these short-run disruptions existed, whether they had a significant effect on the viability of affected SMEs in Latin American EMEs, how long they took to resolve themselves, and levels of attrition. To appropriate a Keynesian metaphor: How many SMEs lived to see long-run benefits of FSFDI? A thorough assessment would require access to reliable panel data that also tracked the identity of firms, making it possible to follow their movement over time. A substantive amount of qualitative data in the form of interviews with both banks and SMEs would be necessary to establish the explanation for why firms were or were not phased-out, for example, profitability issues or a lack of sufficient information.

To this writer's knowledge there are no existing studies that have used data of this nature to assess this question. Certainly, it is plausible to assume that

there are cases in which SMEs are phased out of existing lending relationships for lack of sufficient quantifiable data; it is equally plausible to assume that when this occurs, it is highly costly for affected firms. For these reasons, further research on this subject is warranted.

Cardenas et al. (2003) argue that foreign subsidiaries improve the conditions of acquired Latin American EME banks by “drastically overhauling portfolios.” Schulz (2006) also discusses how theory suggests that asset qualities of acquired banks improve through the transfer of credit workout schemes and risk management techniques from home markets, resulting in lower risk portfolios. Berger et al. (2001) assert that distressed banks may ‘cut-off’ informationally opaque firms.

A paper commissioned by the Bank of Canada studied the experience of two major Canadian commercial banks, Scotiabank and CIBC. Both of these institutions have multi-national operations with investments in various Latin American EMEs (Murray, 2004) and are thus exemplars of FSFDI. The study finds these banks to be highly risk averse (Murray, 2004). In order to maintain the desired low-risk profile, a number of checks and balances are put in place. Scotiabank’s formula, for example, involves sending all credit decisions back to their Canada-based headquarters specifically to avoid local decision-making biases (Murray 2004, 7). This can be interpreted to mean that Scotiabank wants to ensure that loan officers are not relying on ‘relationship-lending’ type information to make lending decisions. CIBC also cited concerns that local loan officers were too “optimistic” when granting credit, though the report does not outline their mitigating strategy (Murray 2004, 4).

The cases of Scotiabank and CIBC suggest that clients, who previously relied on relationship-based lending, would likely be unable to meet new transaction-based criteria if asked to provide it. In this instance it is plausible that they would be deemed ‘high risk,’ and phased-out as part of any loan restructuring that either bank was involved in. We already know that SMEs rely relatively more heavily upon ‘relationship-based’ lending and would therefore be the most likely of any pre-existing clients to encounter difficulties in having their loans renewed for lack of sufficient information. Admittedly, given the banking sector’s reluctance to share data on such decision-making processes, it would be extremely difficult to confirm that this phase-out process takes place; however, on the basis of the foregoing it seems plausible and worthy of further study. To recap, potential methods of inquiry could include repeated surveys of small, medium, and large-scale enterprises and lenders both before and after the takeover of a domestic bank. This would help determine the extent to which firm size and its use of relationship rather than transaction-based lending affected its ability to retain a relationship with the bank.

The Potential Costs to SMEs of Losing a Credit Relationship

The discussion of relationship-based lending specifically noted that the bank-borrower relationship gains additional value over time. It is difficult to ascribe a precise quantifiable value to the information that is accumulated; however as discussed earlier, it is at least partially reflected in the reduction of interest rates, access to larger loans, lower collateral requirements and fees. SMEs that lose

such a relationship with a lender lose all of the intangible value inherent within it because the value is non-transferable. A number of writers, notably De Soto (2000), have highlighted the difficulties that entrepreneurs in emerging markets face in securing collateral for credit; standing relationships constitute social capital in a very real and tangible sense, and the effects of their loss should not be underestimated.

A loss of credit can pose serious impediments to firm survival. Firms rely on working capital loans to finance receivables and inventory during the business cycle. Extended or complete withdrawal of this type of credit could, depending on the business, lead a firm to encounter problems with paying suppliers for goods necessary to fulfill contracts and purchase orders. Depending on the length of time associated with the disruption, firm survival could be a concern. Finally, establishing a new banking relationship also implies additional costs separate and distinct from those mentioned above. Search fees, new loan set-up fees, collateral registration and legal fees are all part of the costs borne by a firm while establishing new credit with a bank (Berger, et al. 2001). In addition to the loss of the intangible 'relationship' and its accumulated stock of information, costs associated with disruption, search, and set-up, must also be factored into the equation. In summary, for SMEs, which are *ceteris paribus* undercapitalized when compared with larger firms, displacement from an existing lender and the process of finding and establishing a new lending relationship pose significant costs and represent potential threats to firm survival. Again, the extent to which the process of adjustment actually takes place post-FSFDI requires substantiation; however, at minimum it is implied by some of the literature. Certainly, the potential costs as outlined are substantial enough to merit further investigation.

ANALYSIS

There are a number of implications that stem from the arguments presented. Notwithstanding the question of the net effect of increased FSFDI on SME access to credit, it is clear that information opacity and SMEs' reliance on relationship-based lending act as serious barriers to credit. Under certain conditions, these obstacles are aggravated by an influx of FSFDI. Moreover, even in cases in which SMEs gain increased access to credit due to domestic banks developing SME niches—to the extent that the banks must rely on these lenders to obtain financing, they may be at a disadvantage compared to larger firms capable of accessing foreign-controlled banks. In cases where domestic banks are less efficiently managed, such inequality will tend to persist until reforms are made.

It is ironic that as banks become more sophisticated, they tend to move from more relationship-based to more transaction-based lending practices. SMEs that rely on relationship-based lending will thus continue to suffer, forced to find sub-optimal domestic banks to obtain credit. It is therefore necessary to overcome some of the constraints of information opacity by supporting the development of financial infrastructure such as credit bureaus that SMEs may register with, while concurrently fostering competition within the financial sector so that there are market incentives in place to encourage the development

and entry of new domestic banks that will develop SME lending niches. To prevent the emergence or persistence of duality in the financial sector, competition is also necessary to force domestic banks to improve their management practices, specifically their credit granting policies, risk management techniques and levels of human capital; otherwise they will continue to operate sub-optimally. As this process unfolds it will be crucial for the government to provide the necessary institutional tools and support to SMEs, as discussed above, so that they may reduce their reliance on relationship-lending. In summary, initiatives focused on both macro as well as micro-economic reform and support is required concurrently.

Policies aimed at reducing the costs to SMEs when they are forced to switch lenders seem more limited. The ideal way to reduce such costs is to avoid incurring them; in other words, minimize the number of SMEs that rely solely on relationship-based forms of information. In addition, strengthening domestic banks may assist those SMEs that are ultimately forced to switch by broadening both the quantity and the quality of the lending options available to them.

Admittedly, during any process of adjustment when non-efficient industries transition to efficient market-based systems, short-term costs are necessary for medium to long-term efficiency gains. Therefore, a final observation is that the costs to SMEs will likely be greater in the short term; however, if broad ideas such as the ones outlined above are developed into effectively executed policies sooner rather than later, they would likely mitigate the magnitude of harm done to SMEs' prospects for accessing and maintaining credit in the context of rising FSFDI. Ideally barriers fall more than they grow, allowing efficiency gains to trickle down to more firms in the economy.

CONCLUSION

This article provides a critical commentary on the literature surrounding the effect of FSFDI on the SME market for access to financing. It starts by critiquing a study by Clarke et al. (2002), which uses cross-country data to argue that FSFDI bring benefits to all firms. The authors argue that FSFDI creates efficiency gains in the financial sector that result in an expanded pool of credit, at reduced prices. Further, Clarke et al. theorize that when foreign firms poach larger, more sophisticated clients, domestic banks may develop SME lending niches; hence, SMEs also benefit from FSFDI for this reason. This article highlights the problems with this and other studies. Methodologically, it highlights the difficulties with relying solely on cross-country data from borrowers who already have access to credit. More fundamentally, barriers to credit for SMEs, often derive from the banking sector's requirement of quantifiable, standardized forms of information that many smaller, semi-formal firms do not have. Consequently, researchers and policy makers need to look more closely at the basis upon which banks grant credit, be it transaction-based lending or relationship-based, rather than simply looking at the price and quantity of funds lent. If increased FSFDI simply translates into a move toward more transaction-based lending, then, in the short-run, barriers to credit for informationally opaque firms may actually rise, *ceteris paribus*.

Noting the mini case study of Mexico, this article demonstrated that competition may be a crucial variable in enhancing the likelihood that domestic banks service SME niches. However, the potential inefficiency of domestic banks as sources of credit calls into question how desirable they are as alternative lenders. Improvements to internal credit and management practices are necessary to increase the ability of these institutions to service SME clients, among others, effectively. Finally, a more thorough examination of the dynamics involved with FSDI is warranted. Looking at whether or not SME clients with standing banking relationships are forced to switch lenders post-FSDI is critical if one is to effectively measure its costs to this sub-sector of firms.

Policy must be directed towards improving the market conditions within which SMEs operate. Institution building that assists SMEs in overcoming the problem of information opacity may not assist in increasing access to loans with foreign-owned banks. It may however reduce the likelihood of SMEs being phased-out of existing relationships and enhance access to better managed domestic institutions with skilled lenders. This is not to say that domestic banks should increase efficiency through a move to transaction-based approaches, but to assist those domestic banks which have chosen to target SMEs as clients to be in a position to provide higher value service to them. Promoting competition in the financial industry is another way to encourage enhanced financial services to all sectors of the economy, including SMEs, and is therefore an important sister policy to institutional strengthening.

NOTES

¹ This type of capital and technology acquisition must be sharply contrasted with the type that occurred during the neo-liberal reform period. The reforms caused technology and capital acquisitions to leap frog the capability sets available in the economy. In the SME example, the slower progression of capital acquisition, often coupled with on-the-job training and/or increases in education and skills training may only occur over the medium-term.

² For this statistic the data provided by Jenkins (2000) did not disentangle which banks were private and which were state-owned, however, given 98 banks were private, it is known that at least a portion of the 88 were private.

ACKNOWLEDGEMENTS

The author would like to express her sincere gratitude to those who assisted in preparing this article for publication. This article benefited enormously from the insightful comments and invaluable editorial work done by Stewart Prest. The author would also like thank Patrick Cuenco, Kristen Thomasen and Alexandra Dostal for their exceptional editing work and support throughout the process.

REFERENCES

- Bank for International Settlements. 2004. *Foreign direct investment in the financial sector of emerging market economies*. Working Group established by the Committee on the Global Financial System. Basel, Switzerland: Bank for International Settlements Press & Communications.
- Barajas, A., & Steiner, R. 2002. Why don't they lend? Credit Stagnation in Latin America. *IMF Staff Papers*, 49(0), 156-184.

- Beck, Thorsten, A. Demirguc-Kunt and Vojislav Maksimovic. 2003. *Financial and Legal onstraints to Growth: Does Firm Size Matter?* World Bank.
- Berger, Allen N., Leora F. Klapper and Gregory F. Udell. 2001. The Ability of Banks to Lend to Informationally Opaque Small Businesses. *Journal of Banking and Finance* 25: 2127-2167.
- Berger, Allen N. and G.F. Udell (1995). Relationship Lending and Lines of Credit in Small Firm Finance. *Journal of Business*. 68: 351-382.
- Cardenas, F., J. P. Graf and P. O'Dogherty. 2003. *Foreign Banks Entry in Emerging Market Economies: a Host Country Perspective*. Banco de México Paper. The Bank for International Settlements (BIS). <http://www.bis.org/publ/cgfs22cbpapers.htm>.
- Claessens, Stijn. 2006. *Access to Financial Services: a Review of the Issues and Public Policy Objectives*. August 2, 2006. World Bank: Oxford University Press.
- Claessens, S., A. Demirguc-Kunt and H. Huizinga. 2001. How Does Foreign Entry Affect Domestic Banking Markets? *Journal of Banking and Finance*, 25:891-911.
- Claessens, S. and J. Lee. 2002. Foreign Banks in Low-Income Countries: Recent Developments and Impacts. World Bank Manuscript.
- Clarke, George R.G., Robert Cull and Maria Soledad Martinez Peria. 2002. Does Foreign Bank Penetration Reduce Access to Credit in Developing Countries? Evidence from Asking Borrowers. *Development Research Group*, World Bank U.S.A.
- De Soto, Hernando. 2000. *The Mystery of Capital: Why Capitalism Triumphs in the West and Fails Everywhere Else*. New York: Basic Books.
- The Economist*. 2002. Finance and Economics: Won't Lend; Mexican banks," October 12, 2002, Vol. 365, Issue 8294, p. 72-3.
- Jenkins, Hatice. 2000. Commercial Bank Behavior in Micro and Small Enterprise Finance. *Development Discussion Paper* No. 741. Harvard Institute for International Development. U.S.A.
- Klapper, Leora, Luc Laeven, and Raghuram Rajan. 2004. *Business Environment and Firm Entry: Evidence from International Data*. Policy Research Working Paper 3232. Washington, D.C: World Bank.
- Lerchs, G. 2002. *The Study of an Operational Definition for Micro, Small and Medium-sized Enterprises in Egypt*. Vancouver: Egyptian Ministry of Foreign Trade, IDRC, CIDA.
- Murray, J. 2004. *Direct Investment in EMEs: the Experience of Canadian Banks*. Bank of Canada Paper. The Bank for International Settlements. <http://www.bis.org/publ/cgfs22cbpapers.htm>.
- Peek, Joe and Eric S. Rosengren. 1996. Small Business Credit Availability: How Important is Size of Lender?. *Financial System Design: The Case for Universal Banking*, Anthony Saunders and Ingo Walter, eds. Burr Ridge, IL: Irwin Publishing.
- Pomerleano, Michael and George J. Vojta. 2001. *What Do Foreign Banks Do in Emerging Markets? An Institutional Study*. The World Bank, International Monetary Fund, and Brookings Institution 3rd Annual Financial Markets and Development Conference, New York, NY April 19-21.
- Saavedra, Jaime. 2003. "Labour Markets During the 1990s." In Pedro-Pablo Kuczynski and John Williamson (eds), *After the Washington Consensus. Restarting Growth and Reform in Latin America*, Washington D.C.: Institute for International Economics, 263.
- Schulz, Heiner. 2004. *Foreign Banks in Mexico: New Conquistadors or Agents of Change?* Wharton Financial Institutions Center Working Paper No. 06-11 University of Pennsylvania. April 22, 2006.
- Schumpeter, Joseph A. 1962. *The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest, and the Business Cycle*, trans. Redvers Opie. Cambridge, Massachusetts: Harvard University Press.
- Stiglitz, Joseph. 1998. "The role of the financial system in development." (Paper presented at the 4th annual Bank Conference on Development in Latin America and the Caribbean, June 29).
- Strahan, Philip E. and James P. Weston. 1996. Small Business Lending and Bank Consolidation: Is There a Cause for Concern? *Federal Reserve of New York: Current Issues in Economics and Finance*, II, 1-6.
- Todaro, Michael P., and Stephen C. Smith. 2003. *Economic Development*. Eighth Edition. Massachusetts: Addison Wesley.

**COLONIAL TERRORIST, AMERICAN HERO:
TWO HISTORIES OF THE ORIGINAL AMERICAN
INSURGENCY**

Daniel Doktor

School of International and Public Affairs, Columbia University

Two books on the American Revolution – one history and one novel – offer some surprising insights into the nature of America’s first encounter with insurgency: it’s own. As these the books demonstrate, the American rebels, ably led by George Washington, adopted insurgent-style tactics to defeat their British counterparts. The lessons learned (and ignored) by the British, and the experiences of both armies offer valuable insights to contemporary observers of the United States’ efforts to combat insurgencies in Iraq, Afghanistan and elsewhere.

Washington’s Crossing

David Hackett-Fischer

(2004. New York: Oxford University Press.), 564

Pages.

The Glorious Cause

Jeff Shaara

(2002. USA: Random House, Ballantine.), 680 Pages.

* * * * *

Jeff Shaara introduces his novel *The Glorious Cause* with a note to the reader that attempts to capture the grandeur of the American Revolution. It was, he writes, “in many ways our first civil war, and eventually [it] became the first true world war” (Shaara 2002, viii). While perhaps true, his novel and a recent Pulitzer Prize winning historical examination, *Washington’s Crossing* by David Hackett-Fischer, together reveal that the Revolution provided still another first—America’s first insurgency. To the typical contemporary American observer of international relations, the term *insurgency* evokes an ideologically driven foreign menace. In 1776, the founding fathers and their followers fit precisely into such a description in the minds of the British superpower. These two books offer a portal to that time and to the founding myths of the United States. Their authors tell the stories of Washington, Hamilton, Lafayette, and Knox, and reveal unknown back-stories with a historian’s glee.

Jeff Shaara’s 680-page tome provides a macro perspective, depicting the whole of the revolutionary war, while Hackett-Fischer’s more focused account examines the lead-up and immediate aftermath of George Washington’s fateful amphibious assault on Trenton, New Jersey. The fighting force that carried out the risky attack at Trenton is skillfully characterized by Hackett-Fischer as a motley crew: part militia, part army, and all rebel. His book uses this nascent Continental Army as a prism through which to see the birth of a nation.

Conversely, Shaara's novel traces the birth of an army, from small bands of uncoordinated attackers to a unified, well-drilled, and disciplined force. Both books lend credence to Mao Tse-Tung's famed assertion that "Guerilla hostilities are the university of war" (Tse-Tung 1961,73).

While the political backdrop of the revolution (from the Boston Tea Party to the Declaration of Independence) is well known, the works of Shaara and Hackett-Fischer together expose a second, more immediate type of politics. They describe the imperative of local support in an insurgency. For example, each battle was not only fought *against* the British army, but also *for* public opinion. General Washington's victory at Trenton defeated only 1,500 opposing soldiers, and even then these soldiers were German mercenaries (the *Hessians*), not the polished Red Coats of the English Empire. The importance of that victory, however, was not merely military but also emotional, as it demonstrated to the fledgling American nation that the war could actually be won. Recognizing the insurgency as a domestic phenomenon has important implications for contemporary U.S. policy, as policy-makers struggle to understand and properly react to maturing insurgencies in Iraq and Afghanistan.

* * * * *

The back cover of *The Glorious Cause* ironically evokes the centrality of George Washington's Christmas Night crossing of the Delaware River as the central development in the Revolutionary War. The novel that precedes the back cover, however, seems minimally interested in the details of the said crossing. What constitutes the crux of Hackett-Fischer's history garners only three lines in Shaara's novel, "With a sharp cry, growing into a long high cheer, the troops began to charge into Trenton. [Line Break] The Route was complete, the Hessians completely stunned by the surprise assault" (Shaara 2002, 158). The rest is left to the reader's imagination. This style, which forgoes detailed descriptions of battles and tactics, marks the most notable departure of Jeff Shaara's writing style from that found in his father Michael's civil war masterpiece, *The Killer Angels*. This absence of detail certainly disappoints in terms of the richness of the reading experience, but those so interested can look to a historian-come-story-teller like Hackett-Fischer who revels in precisely such details.

Hackett-Fischer, despite (or perhaps because of) limiting his focus to a relatively short timeframe (the winter of 1776-1777), provides a fantastically readable and surprisingly (for its scope) epic snapshot of the American nation in transition. This painstakingly researched volume demonstrates, through detailed analyses of tactics and maneuvers, how George Washington came to order the Christmas night attack across the Delaware River. Interestingly, he describes how the American insurgents were not like today's insurgents, who are characterized by a relative dearth of technology. On the contrary, he writes that while "the standard practice in European armies during the 19th century was to use two or three 'battalion guns' for every thousand infantry, the Americans advanced on Trenton with seven or eight guns for every thousand muskets, a very large proportion" (Hackett-Fischer 2004, 223). Henry Knox, the artillery

captain for the revolutionary army plays a central role in both books, and demonstrates that insurgents are not defined by the weapons they wield.

The American Revolution, as described by Hackett-Fischer and Shaara, reveals several other vital insights into insurgents and counter-insurgency. While the term “asymmetric warfare” normally refers to an imbalance of fire-power between two foes, these books shed light upon a secondary asymmetry – one of unequal *motivations*. From the British perspective, it was a war of limited aims (to return a colony to the empire’s rule), and one fought by soldiers and officers driven by honor. The American rebels, by contrast, fought a decidedly *unlimited* war for their survival as a self-determined nation. Hackett-Fischer contrasts Europe’s “aristocratic elites, who thought of war as a nobleman’s vocation and a pursuit of honor” with Americans who “fought not for the sake of fighting, but for the sake of winning” (Hackett-Fischer 2004, 370). Shaara provides the color in his depiction of one of the many rants of British General Clinton who criticized that pyrrhic victory at Bunker Hill (termed Breed’s Hill by the British) and claimed “there was the perfect opportunity to go around, cut the rebel retreat from behind...but of course, there would not have been such...*pageantry*” (Shaara 2002, 40). The British took a straight-down-the-middle approach at Bunker Hill, and despite winning the battle, the heavy losses they sustained reverberated throughout the war and colored the strategies of the Commander-brothers, General William Howe and Admiral Richard Howe.

For the American rebels, fighting and winning was no small task in this battle of David and Goliath proportions. Both Hackett-Fischer and Shaara evoke the landing of the British Navy in New York harbor and the awe that such a magnificent display of force inspired in the local population. Hackett-Fischer puts the numbers of the attacking force at a full *two-thirds* of the British army, complimented by nearly 30,000 troops-for-hire from Germanic States (Hackett-Fischer 2004, 33). Their various landings and movements throughout the colonies during the war went largely unimpeded. Yet, among the most important lessons these books teach is that a war is not won through troop movements alone.

How did the ragtag American rebels, whose bloodied feet stained the snow-bound trails upon which they marched, defeat the most skilled army in the world? Hackett-Fischer and Shaara both describe how the British contributed to their own defeat through their behavior and biases. The British Army, and to an even greater extent, the highly professional Germanic troops who accompanied them, held a profound disrespect for their enemy. Hackett-Fischer writes of the Hessians, “they despised the American language of liberty and freedom as the cant of cowards, traitors and poltroons” (Hackett-Fischer 2004, 59). To these European troops, the rebels were not only unworthy opponents, they were war criminals as well. “By eighteenth century European rules of war, civilians in arms could be put to death,” writes Hackett-Fischer (2004, 180). Yet American militias for much of the Revolutionary War were little more than civilians in arms. As even the American General Charles Lee refers to them in *The Glorious Cause*, they were an “outrageous band of amateurs, these farmers” (Shaara 2002, 430). Such disrespect yielded an ill-advised dismissal of the enemy and a reasoning that seemed to dismiss the fighting capabilities of the Americans

simply because they employed different techniques.

The British dismissal of American troops was compounded by the fact that while force-projection was the logistical feat that allowed the war to occur, the force projectors (i.e. the policy-makers) in London scarcely had any connection to the war itself. When British General Lord Cornwallis returned to London on leave during the war, he was astonished by the detachment of the government from the foreign war. “Now that I am here,” Shaara quotes Cornwallis, “I begin to understand how very far away [the war] is. It is simply not *real*. To the ministry, even to His Majesty, it seems that America is only some place on a map” (Shaara, 2002, 475).

That place on the map played host to a palpable strategic tension within the wartime army. Before being sent to America as Commanders, the aforementioned Howe brothers were “moderate Whigs” who “did not support strong measures against America,” and instead “sympathized with American demands for rights” (Hackett-Fischer 2004, 70). They favored an overall strategy of *pacification*, which was accompanied by the minimally aggressive, slow moving tactics of occupation and conversion of the people, including invitations to take an oath of allegiance to the Crown in order to avoid punishment (Hackett-Fischer 2004, 161). From a theoretical standpoint, this strategy could have worked by eventually convincing the local population that the British army was generally benevolent and, more importantly, not going anywhere. However, this strategy represented only one pole of the invaders’ behaviour.

The other pole was what the Germanic troops called *Shrecklichkeit*. This was war by punishment—a philosophy that routinely denied the typical mercy granted to captured soldiers, allowing instead for summary executions and abuse of the local population. As Hackett-Fischer writes, “plundering became pillage, so pillage became rape” (Hackett-Fischer 2004, 179). Shaara offers a more haunting perspective from the American spy, Nathan Hale, hidden deep within a British camp in New York, who hears the screams of a young woman being raped and sees, among the British officers, “smiles, reacting to the screams of the woman only as some shared experience” (Shaara 2002, 83). As a result, before the American colonials could be pacified by the British strategy, they were radicalized by the German mercenaries’ tactics. Thus, the conflict between the two poles of British strategy provided for the ultimate failure of both.

This destructive tension between the tendencies of the Howe brothers and those of the Germanic officers reflects the differences between counter-insurgency and anti-insurgency strategies. The Howe brothers understood that counter-insurgency hinged on the necessity of recapturing the sentiments of the American population. The Germanic anti-insurgent tactics hoped to extinguish the insurgents one-by-one. Yet, as the British experience in America demonstrates, a poorly executed anti-insurgency can ruin even the best-laid counter-insurgency plans.

While Shaara is no military historian and his prose largely avoids discussions of military philosophy, he does contribute one particularly informative insight about the centrality of land occupation in the European way of war. In *The Glorious Cause*, Shaara effectively paints the British leadership as obsessed with the holding of territory. As Washington discovered, General

Howe believed “defeating your opponent is best accomplished not by defeating his army, but by capturing his capital” (Shaara 2002, 239). In a land the size of the United States in 1776 (essentially the east coast of the current United States), such a strategy was doomed to failure. Hackett-Fischer describes how Howe attempted to create an “extended chain” of garrisons from the Delaware River to the Hudson River: “the area was less than five percent of New Jersey and a tiny fraction of the continent, and yet it stretched the resources of the British Army to the limit” (Hackett-Fischer 2004, 188). Even the most powerful armies can be spread only so thin. Observers of insurgency, such as journalist Robert Taber, agree; “the ‘clear and hold’ strategy is always doomed to failure because the government ... cannot hold many such areas without dangerously scattering its forces” (Taber 2002, 86). Washington responded to the tactics of the British by fashioning a guerilla-style strategy, which placed minimal value on land. Soon after capturing Trenton (following the Delaware crossing), Washington quickly abandoned it to keep his army on the move. Hackett-Fischer posits, “he resolved to keep his army in being, but it would be a ‘retreating army,’ defending what it could, yielding when it must” (Hackett-Fischer 2004, 102). While to some the American Revolution represented the birth of the modern West, the tactics employed by General Washington had a distinctively eastern bent—a yin and yang character that evokes Mao’s famed lines, “the enemy advances, we retreat; the enemy camps, we harass; the enemy tires, we attack; the enemy retreats, we pursue” (Quoted in Taber 2002, 17). At their foundations, all insurgencies are marked by a profound flexibility.

Washington’s flexible style of leadership was epitomized by emphasis on an open exchange with his fellow generals—a crucial element of Hackett-Fischer’s history of the Delaware crossing. In a feat of research, Hackett-Fischer offers insight into the details of the various tactics proposed during these exchanges. For example, Horatio Gates, according to the author, advocated a “Fabian Defense,” whereby “Americans should avoid a big battle, retreat into the interior...and wear down the European armies by slow attrition” (Hackett-Fischer 2004, 78). This is precisely how Taber characterizes guerilla war: *campaigns* of attrition in place of *battles* of attrition (Taber 2002, 98). General Lee, a near villain in Shaara’s account, advocated still another style of attack, one premised upon guerilla-style lightening attacks of small groups against the stronger British ally, followed by subsequent retreat (Hackett-Fischer 2004, 78). Hackett-Fischer asserts that Washington and the American Revolutionaries adopted a hybrid approach, assimilating both strategies when necessary. Sun-Tzu – who said that the ideal attack is one against your enemy’s *strategy* – would have certainly approved (Sun-Tzu 1963, 77).¹

For even the casual observer of international relations, it is difficult to read Hackett-Fischer and Shaara’s accounts and not think of the current experiences of the American military juggernaut in its struggles against the comparatively primitive Iraqi insurgency. When Shaara’s General Cornwallis returns home to England and confronts the disconnect between war abroad and life-as-usual at the English court, an observer of contemporary American politics might find an obvious corollary in the policy of domestic tax cuts during a time of foreign wars. The comparison is even more evident for Hackett-Fischer who

chooses his examples to illustrate the connection; following the Declaration of Independence, revolutionaries “pulled down a handsome equestrian statute of George III...cut off its head and carried the body through the town” (Hackett-Fischer 2004, 29). Written in 2003-04, Hackett-Fischer knows such an example can and will evoke the famed images of a Saddam Hussein statue pulled down by American forces following the fall of Baghdad.

What can the experiences of the Revolutionary War teach American policy-makers and military officials about how to better execute the war in Iraq? The implications are not promising. Will more troops solve America’s problems in Iraq? Not likely, according to these two stories of the American Revolution. Howe’s system of garrisons across New Jersey only served to weaken the capacity of his army while doing little to truly cover the vast colonies. Might the American Army be able to make a better effort to win hearts and minds? Judging by Shaara and Hackett-Fischer, the time has likely passed for that. The anti-insurgency tactics of *Shreklichkeit* employed by the Germanic tribes and the practices of forage, plunder, and eventually pillage that were necessary to support a huge army eliminated the possibility of a successful execution of Howe’s *pacification* strategy. Given the abuses at Abu-Grahib and Haditha, among others, as stains on the reputation of the American military in Iraq, the American efforts there may have passed the point of no return. While it is would be unwise to assert that each and every lesson of the American insurgency of the 1770s applies to the Iraqi insurgency of the 2000s, the underlying themes are informative of both mistakes made as well as opportunities still available.

There is one final opportunity exposed by these histories of the American Revolution that does not appear to be of great likelihood for the United States at this point: co-opting the potential foreign alliance of the domestic insurgents--Iran in the case of Iraq. Some of Jeff Shaara’s most engaging portrayals detail Benjamin Franklin’s war-within-the-war as he works desperately in Paris to achieve a formal alliance between France and the American colonies. Skillfully playing off tensions between the British and the French, Franklin eventually secures the French alliance and the scales of battle are tipped in favor of the Americans. Arguably, Britain could have offered the French a deal to stay out of the Revolutionary war. Today Iran, by most accounts, offers material and logistical support to Shiite insurgents in Iraq. Should a full-out civil war engulf the country and threaten the region, Iran might choose to formally ally itself with Iraqi Shiites. Given this potential, convincing Iran to permanently withdraw its support, however unlikely, may be the only viable option for the United States to successfully weaken the Iraqi insurgency.

With regards to viability, the question that both Shaara and Hackett-Fischer leave us with is perhaps the most disconcerting: can an insurgency *ever* be defeated? Does a war on *insurgency* itself, or even on *insurgents*, make sense strategically? On this final question, the evidence is inconclusive. What would have happened if British General Clinton had his way and the strategy had been to relentlessly attack Washington’s army? As it was, Washington and company were only a hair’s width away from extinction. On the other hand, had his strategy proven successful, and defeated Washington, these books suggest that another hero might have taken Washington’s place in the history books,

regardless of the talents and charisma of America's favorite founding father. Either way, these two books illuminate the march of history, which we ignore at our own peril.

¹ Sun-Tzu. 1963. *The Art of War* Trans. Samuel B. Griffith. Oxford University Press, 77.