Security Challenges

Volume 10 Number 3 (2014)
ARTICLES

Paul Dibb and John Lee
Why China Will Not Become the Dominant Power in Asia .......................... 1

Rick Nunes-Vaz, Steven Lord and Daniel Bilusich
From Strategic Security Risks to National Capability Priorities ...................... 23

Mick Ryan
After Afghanistan:
A Small Army and the Strategic Employment of Land Power ...................... 51

Steven Paget
Old but Gold: The Continued Relevance
of Naval Gunfire Support for the Royal Australian Navy ........................... 73

Paula Hanasz
Hydro-hegemony and Water Conflicts in South Asia ............................... 95
Editors’ Note

2014 was an unusual year for analysts of Asia’s regional security environment. Despite regular periods of local tension and no sign of an end to key disputes, global security attention has seemingly shifted back to the Middle East and, perhaps more surprisingly, to Europe. For Australian and regional policy-makers and analysts this unusual dynamic presents an opportunity to take stock and re-assess underlying judgements for 2015.

Paul Dibb and John Lee offer a provocative challenge to the common wisdom by arguing that China will not become a dominant power. Examining the economic, demographic and regional challenges facing China, they argue China is far weaker internally and regionally than often assumed. Meanwhile Paula Hanasz examines how India’s access to and control over water supplies in South Asia affects the regional political and security dynamic.

Turning from regional to operational concerns, three papers in this edition call for re-thinking how Australia provides for its security. Rick Nunes-Vaz, Steven Lord and Daniel Bilusich detail how strategic risk management can provide a framework for identifying national capability priorities. Mick Ryan looks at the lessons from the Australian Army’s busy decade from East Timor through to Afghanistan. Finally, Steven Paget argues for the enduring relevance of Naval Gunfire Support.

Finally, this is the last volume edited by Stephan Frühling. The other editors want to thank him for the outstanding contribution he has made to the journal over the last several years. He has done more than anyone else to turn Security Challenges into the publication it is today. We wish him all the very best with his new projects.

Best wishes to all our readers for a safe and happy festive season.

Stephan Frühling  Peter Dean  Andrew Carr  &  Iain Henry
Managing Editors
November 2014
Why China Will Not Become the Dominant Power in Asia

Paul Dibb and John Lee

The belief that China will soon become the dominant power in Asia is based on assumptions that its continued and rapid economic rise, and its emergence as a regional peer of America's in military terms is all but assured. Such a belief underpins arguments that a fundamental strategic reorganisation of Asia is inevitable, and that it will be necessary and perhaps even desirable to concede to China significant 'strategic space'. Dependent largely on linear extrapolations about the future, such arguments ignore the implications of China's economic, social and national fragilities, its lack of major friends or allies in the region as well as the considerable military deficiencies and challenges faced by the People's Liberation Army. With the Defence White Paper due for release in 2015, the government should bear in mind that planning for an era of Chinese dominance in the region—or even its emergence as an American strategic peer in Asia—would be premature if not improbable. Australia should not design its defence force for war with China, but it should be able to counter Chinese coercion and contribute to Allied military operations if necessary.

In his seminal work The Rise and Fall of the Great Powers, Paul Kennedy states there is a very clear connection in the long run between an individual Great Power's economic rise and fall and its growth and decline as an important military power. He makes two further related points: first, the power position of the leading nations has closely paralleled their relative economic position and, second, it is not certain whether the existence of 'rising' and 'falling' powers in an anarchical world order must always lead to war. Applying these important judgments to Australia's defence and foreign policies regarding the rise of China has led to strongly opposing views. There are those who consider that the inevitable rise of China must result in that country becoming the naturally dominant power to which the United States must concede strategic space and acknowledge China's 'legitimate strategic interests'. There are others—including the authors—who believe that China's endless rapid rise economically is far from inevitable and perhaps even unlikely and that its military power will continue to lag seriously behind that of America.

The argument that China will emerge as Asia's pre-eminent power is based on assumptions that its economic and military capacities are expanding and improving at such a rate that regional dominance is all but assured. Yet,

---

2 Ibid., pp. xxiv, 537.
the sustainability of China’s rapid economic rise and capacity to embark on the path towards becoming an advanced and resilient political-economy, in addition to its ability to become a genuine military superpower wielding proportionate regional influence, is widely assumed but rarely analysed in any depth, at least in the Australian literature. In examining the factors that go towards the development of Chinese national power—and its ability to use it to achieve national objectives—predictions about a Chinese superpower with the ability to dominate Asia would be premature, if not improbable, in our view.

The Limitations of China’s ‘Economic Miracle’

When a book by Harvard University’s Ezra Vogel entitled Japan as Number One became a bestseller in 1980, the Japanese economy had been growing continuously at around 10 per cent per annum for two decades since 1960. Some fifteen years later and with Japan well into its lost decade of zero or negligible economic growth, Paul Krugman penned a seminal article entitled ‘The Myth of Asia’s Miracle’. In the article, Krugman offered the simple but irrefutable economic argument that “Economic growth based on expansion of inputs, rather than on growth in output per unit of input, is inevitably subject to diminishing returns.” Observing that this was true of the Soviet Union, Krugman applied the same economic logic to explain the problems faced by Japan from the 1980s onwards. The observation was not so much that growth in these inefficient economies had been illusory but that rapid growth based on ever more finite reserves of capital and labour inputs becomes more and more inefficient, and is therefore ultimately unsustainable.

As any economist would confirm, there are only three ways to generate economic growth: adding more capital inputs, more labour inputs, or using capital and/or labour more productively. In essence, Krugman was simply reaffirming the reality that although countries have different approaches in terms of emphases on the drivers of growth, the laws of economics as described above apply universally to all political-economies.

Is China simply doubling down on recent Japanese mistakes? This comes down to how China has achieved growth, especially since the mid-1990s. Most commentators focus on the spectacular success of China’s export sector and the emergence of China as the ‘world’s factory’. The export-manufacturing sector is indeed significant and a vibrant component of the economy, employing possibly fifty million people directly, and another one hundred million people indirectly. But the greater contributor to Chinese growth is domestically funded fixed-investment which was behind around 40

---

4 Ezra Vogel, Japan as Number One: Lessons for America (Mass: Harvard University Press, 1979).
6 Ibid., p. 63.
per cent of growth throughout the late-1990s and the previous decade. In 2009, and due to the massive stimulus ordered by the government, more than 70 per cent of growth was the result of fixed investment.\(^7\)

To put China’s reliance on fixed investment in context, during the periods of rapid industrialisation in Japan, South Korea and Taiwan in the 1950s and 1960s, fixed investment as a proportion of GDP was at around 30 per cent or below; and only for brief periods of several months did levels touch 35 per cent.\(^8\) In contrast, Chinese fixed investment as a proportion of GDP jumped from a more sustainable 35 per cent in the 1980s, to 38 per cent in 1999, and to 45 per cent in 2004.

According to official World Bank data, which is generally considered to be conservative in its estimate, the figure is now 49 per cent of GDP, the highest in the world with the exception of highly distorted economies of Equatorial Guinea, Mongolia and Bhutan. To explain these extraordinary numbers in another way, it is estimated that over the previous decade, China’s official GDP has increased by 162 per cent. Of this 162 per cent increase, additional labour inputs have contributed about 6 per cent. But an enormous 135 per cent can be attributed to fixed investment. This means that total factor productivity (TFP)—using capital and/or labour more productively—contributed only 20 per cent out of the 162 per cent GDP increase. In other words, increasing levels of fixed investment has been behind more than 80 per cent of China’s GDP growth over the past decade.\(^9\)

Even if these figures deviate slightly from the truly accurate one (which is unattainable given the unreliable nature of China’s official statistics), one can say with confidence that Chinese reliance on the levels of fixed investment to drive and sustain growth is unprecedented in economic history. It is therefore unsurprising that the Chinese economy’s deployment of capital which is used to fund fixed investment is now widely viewed as unsustainable.

The evidence for this is ample. For example, the amount of capital input needed to produce one additional dollar of output (i.e., capital-output ratio) increased from 2:1 in the 1980s to about 4:1 in the 1990s, and was well over 5:1 in 2011 according to OECD figures.\(^10\) The capital-output ratio estimate for 2012 was 5.5:1 meaning that a capital input of $5.50 achieves only $1.00

---


As economic logic insists, and the development experiences of other East Asian countries show, capital-output ratios at this level depict an enormously wasteful and capital-inefficient economy that is not sustainable.\footnote{See John Mauldin, ‘China’s Minsky Moment?’, \textit{Forbes}, 22 March 2014.}

Undertaking such an unprecedented capital-intensive model of growth has other deleterious ramifications. Of particular concern is that the enormous level of capital inputs used to generate growth has meant that national corporate debt levels have risen from 147 per cent of GDP at the end of 2008 to over 250 per cent at the end of June, 2014.\footnote{See Jamil Anderlini, ‘China Debt Tops 250% of National Incomes’, \textit{Financial Times}, 21 July 2014.} To put this in some context, the expansion of debt—which has been used to finance capital inputs—from US$9-10 trillion in 2008 to US$20-25 trillion in 2014 exceeds the size of the total American commercial banking system.\footnote{See Satyajit Das, ‘China’s Debt Vulnerability’, \textit{EconoMonitor}, 9 April 2014.}

Such an unprecedented increase in debt-financed capital has also created systemic problems for the entire financial system in China. In 2011, an investigation by the Chinese Academy of Social Sciences put the debt to asset ratio of Chinese firms at 105 per cent, the highest amongst the twenty major economies studied.\footnote{See Dexter Roberts, ‘China Bad Debt Could Spark Global Growth Slump’, \textit{Business Week}, 9 May 2014.} That China’s overall debt to GDP ratio was then 169 per cent and is now over 250 per cent some three years later means that corporate debt is now undoubtedly far higher than it was in 2011. Given that so much of the fixed investment is wasteful which is reflected in the rapidly rising capital-output ratios, there is almost universal agreement that the official non-performing loan (NPL) ratio of 1 per cent—which has not changed for a decade—is not credible. Instead, most independent analyses conservatively place the NPL figure to be at least 5 per cent, meaning the NPLs on the balance sheets of state-owned banks—whose liabilities are ultimately government liabilities—could be around US$2 trillion.\footnote{See ‘Corporate Debt Reaches “Alarming” Levels’, \textit{China Daily}, 18 May 2012.}

This brings us back to Krugman’s warning about growth models based predominately on ever-increasing levels of capital investment: they eventually run out of steam and no economy has ever escaped the so-called ‘middle-income’ trap through ever increasing capital (or labour inputs). With an export sector no longer capable of driving rapid Chinese GDP growth as it did in the 1990s and earlier this century, all medium-term indicators point to a consensus that the ability of fixed investment to pick up the slack is
nearing its end. This means the rapid growth rates of the 1990s and the previous decade are no longer possible.\(^\text{17}\)

Even if the current Fifth Generation of Leaders manage to successfully transition their economy from one led by investment and exports to one driven by domestic consumption, such a transition would involve severe short- and medium-term disruptions to the economy with precarious risks for the regime. Additionally, such a transition would mean the winding back of the privileged access to capital and opportunity afforded to state-owned-enterprises in favour of a currently suppressed private sector—undermining the primary strategy used by the Chinese Communist Party (CCP) to ensure its continued relevance in the country’s rapidly industrialising economy. Besides, even if reforms were successfully undertaken and domestic consumption were to become the primary driver of growth, GDP growth rates would still slow significantly from the recent past as it is almost impossible, and indeed unprecedented, for consumer-driven economies to grow at the rapid rates enjoyed by China over the past few decades.

Bear in mind that approaches based on linear extrapolations of Chinese absolute economic size expressed in a GDP figure to inform sensible assumptions about Chinese power are also superficial—as it was with Japan, a country with a far more resilient political-economy, in the 1980s. In isolation, GDP size and growth rates offer no decisive indication of how a country is actually faring. Remember that the Soviet Union officially tripled in size from 1950 to 1973, yet its economic model was fundamentally flawed as we realised in hindsight. GDP is essentially an accountant’s tool used to document final economic activity within a country in any given year. But GDP does not measure whether economic activity is productive, profitable or even commercially irrational.

Indeed, the self-criticism previously offered by then Chinese Premier Wen Jiabao that the country’s economic model is “unstable, unbalanced, uncoordinated and unsustainable”\(^\text{18}\) is offered on the basis that much of the fixed investment activity is unproductive, not needed or significantly under-utilised, and cannot be justified by any commercial logic. Even so, activity such as the building of uninhabited housing or an increase in steel-making capacity which will not be utilised is counted as ‘economic growth’ under measurements of GDP and is erroneously treated by many outside commentators as if such activity contributes positively to the accumulation of China’s national strength.

\(^{17}\) For example, see World Bank, China 2030: Building a Modern, Harmonious, and Creative Society (Washington DC: World Bank, 2013).

In reality and as the next section argues, how a country generates growth matters. An economy driven by mandated growth targets rather than commercial logic and/or merit, and artificially dominated by a relatively small number of highly protected and privileged State-Owned Enterprises (SOEs), leaves the country ill-equipped to deal with many of its domestic challenges. These challenges are far more intractable and serious than is generally recognised in the Australian discussion about Chinese power and its future place in Asia. Nowhere is this more apparent than in the country’s lack of structural and policy preparedness vis-à-vis the inevitable aging of its society—an approaching challenge which could become the most significant millstone around the neck of the country’s leaders over the next two decades.

**China Growing Old Before it Becomes Rich**

Although Japan is often considered the ‘grandfather’ of East Asia, much more attention should be paid to China when it comes to estimating the future power balance in the region. Currently, China has a population of about 1.35 billion people, but this number is expected to shrink slowly by around 2030.

More important is the ratio of working-age people to those over sixty-five, the latter considered aging or formal retirees. In the 1980s, the proportion of the working-age population (fifteen to sixty-four years) was more than 73 per cent of the overall population. Currently at about 68 per cent, the working-age population is expected to decline to about 65 per cent in 2020, and 60 per cent in 2035. In 2015, it is expected that the absolute size of the labour force will begin to shrink as more people leave than enter the workforce.\(^{19}\)

The significance of these numbers become apparent when one compares the proportion of working-age people with formal retirees. When China embarked on reforms in 1979, there were about seven working-age persons to every retirement-age one. Today, the ratio is about 5.5 to 1. Current projections suggest that by 2035 there will be barely more than two working persons for every retiree. In other words, the old-age dependency ratio will be more than double over the next two decades.\(^{20}\)

The age profile of the working population also matters. Studies show clearly that most workers are at their most productive and innovative from their late twenties to their mid-forties.\(^{21}\) This has been the basis for China’s ‘demographic dividend’, the massive productivity generated by the combination of declining fertility levels and a mass of young workers entering

---

\(^{19}\) See World Bank, *China 2030*, p. 8.

\(^{20}\) Ibid.

the workforce with relatively light familial responsibilities or burdens.22 In fact, one study estimates that the effects of a favourable working-age population alone accounted for between 15-25 per cent of per capita GDP growth in China from 1980-2010, and that this advantage is all but exhausted.23

Yet, while there are currently around 120 million Chinese people sixty-five years or older, by 2035, there will be around 320 million with the overall population only about one hundred million larger than it is today. Even within the working population, in 2035 there will be 1.5 older workers (fifty to sixty-four years) for each of their younger counterparts (fifteen to twenty-nine years old), which is the direct opposite of the current situation.

Moreover, for a number of reasons it is extremely unlikely that these trends can be arrested or reversed. For a start, China’s aging population is largely the result of a dramatic increase in average lifespan, which has increased from under sixty-five years in 1980 to the current seventy-five years. Fertility rates have also declined, from 2.63 children per woman in 1980 to about 1.5 in 2011. Wealthier cities such as Shanghai which has a reported fertility rate of only 0.6, the lowest of any major city in the world, provide evidence that emerging Chinese middle classes, like Western counterparts, are choosing lifestyle and career opportunities over larger families. While actual future figures could vary slightly from the trend lines, little can be done about China’s aging demographics over the next few decades. Even if the one-child policy were to be relaxed or abolished, the aging trend would not be reversed to any appreciable degree for several decades.

By 2030, China’s age demographics will resemble that of Norway or the Netherlands today. (Incidentally, the United States is the only great power with favourable age demographics leading up to the middle of this century.) But in assessing national preparedness to meet the aging challenge, it is almost certain that unlike aging advanced economies, China will be the first major economy in history to grow old before it grows rich (or even moderately rich). This should cause us to look at more than China’s GDP size and growth, which although impressive, has nevertheless been underpinned by a political-economic model that has left the country woefully unprepared for its inevitable aging for a number of reasons.

First, China’s state-dominated political-economy means that the best economic opportunities are reserved for the approximately 140,000 SOEs. Likewise, around three-quarters of formal finance (mainly bank loans) are given to the SOEs rather than the millions of private firms, meaning that

22 Demographer David Bloom, who coined the phrase ‘demographic dividend’ has argued that around one third of economic growth in East Asia from 1960-90 has been the direct result of favourable demographics. See David E. Bloom, ‘7 Billion and Counting’, Science, vol. 333, no. 6042 (July 2011), pp. 562-9.

SOEs dominate investment in every major sector of the economy except for export-manufacturing. The consequence of such a structure is that the revenues of SOEs have been rising at rates far exceeding GDP growth, while household and disposable income has been rising at rates significantly below GDP growth.

The suppression of household and economic opportunity in favour of SOEs is reflected in World Bank Development Indicators data showing that while investment and household consumption growth was rising at similar rates in the first fifteen years of reform (1979-94), investment growth has consistently and significantly exceeded household consumption since the late 1990s—a period when the CCP decided that SOEs were to reassert dominance in all ‘strategic’ and ‘important’ sectors of the Chinese economy through privileged and cheap access to capital, subsidy regimes, and through other regulatory and tax concessions.

The point is that the state-owned sector rather than private households have been the primary beneficiaries of economic growth and development over the past fifteen years, a reversal of what occurred from 1979-1989 when households took the lead in economic activity. It is therefore not surprising that household consumption as a proportion of GDP in China is around 33 per cent, by far the lowest of any major economy in the world. In a country where up to four hundred million people still live on less than US$2 per day, according to World Bank estimates, the national bias towards the state corporate sector rather than households leaves a poor society tragically unprepared for its own aging.

Second, a GDP per capita figure (derived from dividing national output by the number of people) offers little indication of how wealth is actually distributed throughout the country. When considering measurements of income distribution such as the Gini coefficient (where zero denotes perfect equality and one denotes perfect inequality), China has gone from being the most equal society in all of Asia to the least equal within a generation. Its Gini coefficient has risen from 0.25 in the 1980s to 0.38 in the 1990s, to around 0.5 currently. As a comparison, the World Bank figures for India, the United States, Japan and Russia are 0.34, 0.43, 0.38 and 0.42 respectively.

---

27 For the differences between economic models from 1979-1989 compared to the mid-1990s onwards, see Yasheng Huang, Capitalism with Chinese Characteristics (New York: Cambridge University Press, 2008).
Moreover, the link between suppressed household income on the one hand and dangerous levels of income inequality, on the other, is unmistakable. In a system where around 140,000 SOEs receive the lion’s share of economic opportunity and capital at the expense of millions of private firms and tens of millions of informal businesses, a small number of well-connected and well-placed ‘insiders’—generally those with political connections or ties with the CCP or SOEs—benefit disproportionately from the current growth model.28 Revealingly, and from 1979-1989 when household and private initiative drove economic growth, China’s Gini coefficient was stable at around 0.3. It was only after the mid-1990s with the re-emergence of the state-led model that inequality increased dramatically.29

The point about inequality of income and national wealth is that China’s suppressed household sector will exacerbate the country’s unpreparedness for an aging population since many retirees will be in a far worse and more vulnerable financial position than they otherwise could be. This unpreparedness is further exacerbated by the reality that only around one third of all urban residents and less than 5 per cent of rural residents have some form of central, provincial or local pension fund.30 Although the current pension scheme covers a minority of citizens, the consensus amongst researchers is that the state’s pension liability amounted to about US$2.9 trillion in 2013.31 Other reports estimate that the state’s pension liability could amount to US$10.8 trillion over the next two decades (or almost 40 per cent of GDP based on a generous assumption of 6 per cent GDP growth each year).32 As the following section will argue, such domestic vulnerabilities and shortcomings will have a profound impact on the development of Chinese national power which is poorly appreciated in the Australian discussion.

**National Frailty and Domestic Fragility**

Many countries have made the transition from very-low income to middle-income status (approximately US$15,000 per capita), and China at less than US$10,000 per capita still faces obstacles to get there. But only around

31 See ‘Pension Gap to Hit $2.87 Trillion’, *China Daily*, 14 June 2012.
thirty countries have made the jump from middle-income to high-income status. Except for a small number of oil-rich Middle Eastern states, those economies that have achieved this all have the same characteristics that do not apply to contemporary China. These include institutions required for sophisticated commercial interactions such as rule-of-law, strong property and intellectual property rights regimes that are reliably enforced, and an independent judiciary needed to fairly and reliably resolve commercial disputes between commercial actors. Unlike modern-day China, advanced and innovative economies also have political-economies that largely allocate economic opportunity and credit based on commercial merit rather than on policy or political grounds. In contrast, it is not economically feasible to achieve sustainable growth, let alone become a high-income advanced economy, by adding ever increasing levels of capital and labour inputs.

In addition to these serious flaws in the overall political-economic model that will invariably entail a structural slowdown, reform and domestic concerns will increasingly increase the burdens on the public purse. Currently, the CCP devotes a large share of government finances to enhancing Chinese national power and influence through double-digit increases for its military, the People’s Liberation Army (PLA), and its paramilitary, the People’s Armed Police (PAP). For example, in the 2014 budget, US$241.5 billion was officially spent on the PLA and PAP or 11.4 per cent of the total budget. If the true figure for military spending as estimated by organisations such as the Stockholm International Peace Research Institute is used, the proportion of the budget spent on defence and national security is 14.6 per cent (8.8 per cent on the PLA and 5.8 per cent on the PAP).33

To mitigate what demographer Nicholas Eberstadt has called China’s “slow-motion humanitarian tragedy”,34 there will need to be a much larger share of the budget allocated to public goods such as social security and unemployment benefits, and healthcare which constitutes 10.5 per cent and 6.1 per cent of the 2014 budget respectively.35 This is in addition to enormous and mounting burdens on the public purse such as existing pension liabilities and NPLs hidden in the books of state-owned banks, both mentioned earlier.

Additionally, there will need to be a massive transfer of national wealth and opportunity from SOEs to private businesses and households in order for those without adequate pensions—the majority of the people—to dramatically raise their incomes and look after themselves and their families.

36 China’s Ministry of Finance figures.
as they age. The financial role of the state in managing the aging demographic will be exaggerated by the unique phenomenon wrought by the legacy of China’s one-child policy known as the 4-2-1 problem: in a large number of cases, four grandparents and two parents will be looked after by one child. All this will necessitate dramatic and possibly destabilising change to the Chinese political-economy. Such reforms will also erode the dominant role SOEs currently play in all key sectors of the economy, and subsequently the capacity of the CCP to selectively use SOEs to advance state power and objectives.

Finally, it must be noted that the legitimacy of the CCP depends first and foremost on improving the living standards of its citizens. This means that while developing a powerful military and pursuing an expansive foreign policy may well reinforce national pride, the regime cannot simply do so at the expense of ignoring pressing domestic problems and frailties that are worsening in important respects. Although the regime has so far survived, officially reported instances of mass unrest have grown to above 180,000 according to 2011 figures, rising from just a few thousand in the mid-1990s—proof that rapid growth has worsened rather than tempered national instability. With a slowing economy, and with structural economic and social deficits becoming worse rather than better, China is a large but fragile power ruled by a vulnerable Party which cannot afford any major economic or foreign policy disasters. These do not appear to be strong foundations for the emergence of the dominant state in Asia—an argument reinforced by the following examination of China’s strategic isolation and military inadequacies that will be difficult to rectify in the foreseeable future.

**China as the Lonely Asian Rising Power**

China has very few powerful or influential friends in Asia. For a country with such a large population and the world’s second largest economy it does not have many close bilateral relationships. In her book, *China: Fragile Superpower*, Susan Shirk describes China as strong abroad but fragile at home. This strikes us as being incorrect: in our view, China is certainly fragile domestically but it is also a lonely power when it comes to acquiring real influence in Asia. A listing of China’s friendships in the region reveals that only North Korea and Pakistan can be counted as countries with which it has a strong relationship. But what sort of trust can Beijing have in Pyongyang not dragging it into an unwanted war with South Korea? As for Pakistan, it is constantly teetering on the edge of becoming a failed state and risks a conflict with India that certainly would not be in China’s interests.

---

For centuries in the past, Imperial China was feared and respected as the
dominant power in Asia, as Susan Shirk has correctly observed. But that
was all a very long time ago when China faced no real competition until the
arrival of European colonial powers in the nineteenth century. China now
operates in a highly competitive regional environment against such major
powers as the United States, Japan and India. Of late, many Southeast
Asian countries have become increasingly concerned about China’s
assertiveness and several of them have taken steps to align themselves
closer to the United States. Not even Russia can be counted by Beijing as a
long-term friend, let alone an enduring ally.

The following sections examine in turn China’s strategic situation in
Northeast Asia, Southeast Asia and South Asia. We focus on geopolitical
relationships that may threaten or undermine China’s influence and capacity
to wield power in Asia.

Northeast Asia is a critical area of geopolitical confrontation for China. In
this neighbourhood, it faces the risks of war on the Korean Peninsula, a well-
armed Japan that increasingly identifies China as a threat, and a United
States that has strong forces based in both Japan and South Korea. China’s
relations with Japan in particular are now tense and there is a growing risk of
miscalculation and perhaps even military conflict. Tokyo is building up its
naval and air assets in response to what it perceives as territorial
aggressiveness by Beijing. Japan is strengthening its military relations with
both the United States and Australia and assisting countries such as
Vietnam and the Philippines to resist China’s territorial claims in the South
China Sea. While it is true that China is developing comprehensive military
capabilities to deny the United States unchallenged military operations in its
maritime approaches, Japan and the United States have the capacity to
deny China freedom of movement in the East China Sea and effectively
blockade its naval forces from breaking out into the deep ocean. The Shinzo
Abe Government has reinterpreted the Japanese Constitution to exercise
collective self-defence, which will be permissible when there is a danger of a
country such as the United States that has a close relationship with Japan
coming under armed attack. Although clear limitations will remain on the use
of force by Japan, it is obvious that China’s threatened use of force is
causing Tokyo to strengthen considerably its military situation. In the worst
case, and if Japan loses confidence in the US alliance, this may eventually
provoke Japan to consider taking the nuclear weapons option it has long
denied itself. That would certainly not be in China’s strategic interests.

Some commentators believe that Washington must concede strategic space
and acknowledge “legitimate strategic interests” to a China that will become
the naturally dominant power in the region. This is a highly contested

proposition that asserts the United States and its allies must accommodate China as a power-sharing equal as the only rational response to its rise to power. That is a proposition which we reject because, as argued throughout this article, we do not accept that China is an endlessly rising power that has no serious domestic and external constraints. Neither do we accept that America is a power in terminal decline. Rather, we are of the view that the United States will remain the world's dominant military power for the foreseeable future. Of course, it will be important that Washington demonstrates to China that it will not accept Chinese use of military force towards America's allies and friends in Asia.

On the Korean Peninsula, China is trying to have the best of both worlds strategically but that is an unworkable proposition. Beijing's current gamble is that it can retain North Korea as an ally that, if necessary, will go to war to prevent a unified Korea that is an ally of the United States. But, at the same time, Beijing is trying to develop close relationships with Seoul and playing on the Republic of Korea's (ROK) enmity towards Japan. In the end, China will find that if it cannot restrain Pyongyang's dangerous unpredictability, the ROK will have no option but to continue to rely upon the threat of US retaliation against North Korean aggression.

China and Russia are currently embarking upon a new phase in their relationship that is centred on their resistance to what they see as unfettered American power dominating the current international order. Both China and Russia are leagued together in their suspicions of the United States and the West generally. They share in common a distrust of Western invented international norms and rules of behaviour. Russia and China have become more assertive recently. China is using coercion backed up by the implicit threat of military force in the East China Sea against Japan over the Senkaku/Daiyu Islands and in the South China Sea against Vietnam and the Philippines. For the foreseeable future these two large authoritarian countries are threatening to destabilise the international system. Russia needs its alignment with China and China needs Russian energy and resources. But the rapidly growing inequality in power between China and Russia and their long common border—let alone their huge ethnic and cultural differences—does not auger well for a future peaceful relationship.

Southeast Asia is a region that China traditionally has seen as a sphere of influence as evidenced by its 1,000 years of suzerainty over Vietnam, its ethnic influence in Thailand, the presence of over 25 million overseas Chinese in Southeast Asia and its attempts in the 1960s and 1970s to proselytise communism in the region. Today, China is the major trading partner of most Southeast Asian countries. The region developed ASEAN as a loose-knit multilateral community in the wake of the Vietnam War in an attempt to avoid being dominated by the major powers. It is now a community of ten countries and over 600 million people and it has devised the ASEAN way of making decisions by consensus and avoiding
provocation. Lately, however, it has found itself being pushed around by an increasingly assertive China. Beijing has demonstrated its imperial haughtiness by reminding ASEAN members that they are small countries and China is a big country. Thus, it will not consent to negotiating with ASEAN as a group over the South China Sea so that it can pick-off small regional countries one by one. As a result of this overbearing Chinese pressure, several countries in the region have moved towards closer relationships with America.

Geopolitically, Southeast Asia looms as a crucial chokepoint through which China's huge imports of energy and resources must be transported by sea. According to the Pentagon, approximately 84 per cent of China's oil imports now transit through the Strait of Malacca and the South China Sea.\(^{39}\) China's international trade (exports plus imports) accounts for almost 50 per cent of its GDP and is a highly significant component in its quickly rising living standards. But international trade also imposes great vulnerabilities on China: war with the United States would see China at catastrophic risk of its sea lines of communication through Southeast Asia being blockaded. President Hu Jintao identified this as being China's Malacca dilemma, where Chinese oil imports traverse vulnerable waterways.\(^{40}\)

In South Asia, China does not have close relations with India and instead it has developed relationships with countries on India's periphery—Pakistan, Bangladesh and Sri Lanka. However, India is by far the dominant power on the sub-continent. Kissinger argues that India will be a fulcrum of future regional order based on its geography, economic potential and tradition of sophisticated leadership.\(^{41}\) India aspires to be accepted as a great power and China is seen as its arch rival.\(^{42}\) These two great Asian powers could not be more unlike: India is the world's largest democracy whereas China is an authoritarian communist country; India is set to outstrip China in population size and, unlike China, its demographic base is young with all this implies for India's future economic growth. These two powers have long-standing border disagreements along the Himalayas. And as China extends its naval power into the Indian Ocean this will challenge what New Delhi sees as its natural strategic primacy in the ocean named after it. China and India are two nuclear powers with very different histories and cultures. China's increasing military build-up is causing India to respond, including by cautiously developing its relationships with America as well as an eastward looking policy towards ASEAN. As Harris notes, the relationship between


China and India lacks warmth and depth and there are serious points of friction and underlying mistrust.\(^{43}\)

Overall, China's poor relationships with the United States, Japan and India do not augur well for its ability to shape the future regional order. Moreover, Beijing's increasing aggressiveness and harsh attitudes towards its pre-emptive territorial claims in the region run the risk of miscalculation and conflict. This risk, coupled with Beijing's inclination to challenge established international norms of behaviour, is a suitable point to turn to China's military build-up and an examination of its strengths and weaknesses.

**Is China’s Military Power Exaggerated?**

China undoubtedly has developed substantial military capabilities in the last twenty years or so. America's overwhelming demonstration of superior conventional military power in the first Gulf War in 1991 was a great revelation to Beijing. Since then, it has aimed to develop technically advanced military forces capable of fighting and winning short duration, high-intensity regional contingencies.\(^{44}\) The Pentagon's latest report to Congress on China's military makes it clear that preparing for potential conflict in the Taiwan Strait “remains the focus and primary driver” of China's military investment.\(^{45}\) Preparing for contingencies other than Taiwan include potential contingencies in the South and East China Seas and a range of missions beyond China's coast, including sea lane security, counter piracy, peacekeeping, humanitarian assistance and disaster relief. The report concludes that China's military investments provide it with a growing ability to project power at increasingly longer ranges.

Although China has developed potent military capabilities to make it hazardous for US forces to operate in the approaches to China, the fact remains that Beijing could not enforce a full military blockade of Taiwan or attempt a full-scale amphibious invasion of that island.\(^{46}\) The Pentagon also observes that limited logistical support remains a key obstacle preventing China's navy from operating more extensively beyond East Asia, particularly in the Indian Ocean.\(^{47}\) In addition, "it is not clear whether China has the capability to collect accurate targeting information and pass it to launch platforms in time for successful strikes against targets at sea beyond the first island chain".\(^{48}\)

It is instructive that the Office of the Secretary of Defense makes such cautionary remarks. The fact of the matter is that China's forces still lag

---

\(^{43}\) Ibid., pp. 159, 160.

\(^{44}\) Office of the Secretary of Defense, *Annual Report to Congress*, p. i.

\(^{45}\) Ibid.

\(^{46}\) Ibid., pp. 54-5.

\(^{47}\) Ibid., p. 38.

\(^{48}\) Ibid., p. 32.
considerably behind those of the United States in overall resources, technology and experience. In our view, China is twenty years behind the United States in high-technology weapons and sensor development. China is not a military superpower and will not become one until it develops the capability to project decisive military power anywhere on the globe. Presently, China is a regional military power entirely without any modern combat experience and with major deficiencies in doctrine, human capital and training—and particularly the complexity and realism of joint operations. China's ability to develop a powerful military is also seriously constrained by the fact that its own technological level remains relatively low and that its only source of foreign arms is Russia.49

The most comprehensive recent analysis available in the public domain of China's military capabilities versus those of the United States has been written by Aaron Friedberg, Professor of Politics and International Affairs at Princeton University.50 Unlike many commentators, Friedberg is not inclined to exaggerate China's military capabilities whilst acknowledging the growing threat to the American position in East Asia. For example, he cites a survey by the US Office of Naval Intelligence describing China's capabilities in the acquisition of targeting information essential for anti-submarine warfare (ASW) as marginal.51 China's navy has begun to invest in the underwater sensors, dedicated fixed-wing aircraft, helicopters and surface vessels necessary to locate and track enemy submarines, but it has yet to address its shortcomings in ASW. This is an important deficiency given America's big advantage in terms of tracking other submarines and the difficulty all other countries have of detecting US submarines. China's conventional submarines are relatively easy to detect and its nuclear boats possess little ASW capability.52 Without major improvements in ASW, the Chinese Navy would be an easy target for US (and Japanese) military forces. China's military would be hard-pressed to prevent hostile submarines and unmanned underwater vehicles from operating close to its shores and destroying its surface fleet.53 It also remains unclear how capable of joint coordination China's different services are in operations over water. Integrated operations between a highly regimented and rigidly structured Chinese Air Force and an immature and sea-based Navy would require technological and service-culture innovations, as well as exercises less carefully scripted than has been usual, to develop the requisite interoperability and inter-

51 Ibid., p. 29.
53 Aaron Friedberg, Beyond Air-Sea Battle, p. 37.
service coordination. In promoting officers and selecting leaders, the Chinese prize loyalty to the Communist Party and reliability over independence and initiative. In the meantime, the United States is pressing ahead with technological game changers, such as unmanned undersea vehicles (UUVs) for reconnaissance, surveillance and strike that could radically change undersea warfare to China’s huge strategic disadvantage.

There are similar gaping deficiencies in China’s air defence capabilities against any technologically advanced enemy. As Friedberg points out, China’s ability to detect and intercept ballistic missiles or stealthy aircraft and cruise missiles appears to be limited. Moreover, the United States is working on technological advantages that will make China’s task of air defence even harder—they include a new low observable penetrating bomber and long-range precision strike with very high-speed hypersonic glide vehicles. Such developments would greatly increase the expenditure that Beijing would have to devote to both active and passive defence measures.

In addition, China’s reconnaissance network remains vulnerable to attack by a sophisticated foe: over the horizon radar (OTHR) antennas are large and fragile structures that are crucial to China’s ability to track and target American aircraft carriers. China probably does not yet have the kind of detailed, real-time targeting information it needs to make the DF-21 anti-ship ballistic missile into an effective weapon. In addition, the serious growth of China’s area-denial capabilities is unlikely without sea-based aviation and land-based, over-water, mid-air refuelling capability, in addition to some means of coordination and defence (for example, an AWACS equivalent). These are all vulnerable military capabilities for China to develop.

Some observers argue that China will be able to use nuclear weapons against US forces in the region with impunity because the United States dare not escalate the conflict to large-scale military action against China because of the risk to US cities. This seriously underestimates American war-fighting culture, its possession of a much wider range of tactical and strategic nuclear weapons than China, and the fact that China is one of the most vulnerable countries in the world in terms of its population density to nuclear attack.

55 Ibid., p. 378.
56 Friedberg, Beyond Air-Sea Battle, p. 37.
57 Ibid., p. 139.
58 Ibid., p. 38.
59 Ibid., p. 48.
60 Erickson et al., China’s Future Nuclear Submarine Force, p. 242.
61 White, The China Choice, p. 80.
Space does not permit us to delve into further detailed analysis of China’s military deficiencies, but the examples given above demonstrate that China is unable to assert dominant military power and even in its approaches it has distinct vulnerabilities. As Beijing extends its strategic reach it does not rank as a regional—let alone global peer—competitor of America. In the event of major conflict it would face a correlation of forces that included the United States, Japan and Australia and perhaps India. And, short of major conflict, strong external balancing against China appears to be the new norm in the region, especially in Northeast Asia.62

China is a continental power and, as Robert Ross has argued, China’s maritime power will be limited by the constraints experienced by all land powers, including the geopolitical sources of the repeated failure of land powers to secure maritime power.63 His main thesis is that land powers confront internal threats that impose severe resource constraints in developing maritime power, whereas the geographic circumstances of maritime powers offers them enduring internal border security and ready access to the sea. It is a telling point, in this regard, that China continues to spend as much on internal security as it does on its defence build-up.64 In the Cold War, another land power, the Soviet Union, practised the same sort of access-denial capability to reduce the challenge of US carrier-based aircraft to its territorial security that China is now implementing. China has a surface fleet without organic air power and nuclear-powered submarines that remain relatively noisy.65 Its land-based air capabilities are insufficient to enable China to project decisive power in even the relatively near waters of maritime East Asia.66 China is not capable of challenging US dominance of regional sea lanes or the security of America’s strategic partners in maritime Southeast Asia.67

Nationalism and increasing concerns in China about its dependence on international trade for both food and energy security are driving its naval build-up. But China has the longest land borders in the world, which it shares with fourteen neighbouring countries, and it must simultaneously pursue land and naval capabilities. Failure to develop both land and sea power will constrain China’s power role in world affairs.68 In this analysis,

64 Ground forces continue to make up approximately two-thirds of China’s defence spending. The People’s Liberation Army is also responsible for the survival of the Chinese Communist Party.
65 Andrew S. Erickson, ‘China’s Modernization of Its Naval and Air Power Capabilities’, in Tellis and Tanner, China’s Military Challenge, p. 68.
66 China’s air forces have very little aerial refuelling capability and its bombers are antiquated.
68 Ibid., p. 73.
the constraints on China’s maritime capability will be similar to those encountered by France, Germany and Russia when they too sought maritime power. Unlike China, America can choose to prioritise maritime power in its defence spending and ensure that US naval pre-eminence continues to underpin the broader Asian order. This is a task that is not beyond the technological or innovative capacities of the United States.69

None of this is to underrate the potential challenge to regional stability from China’s military modernisation. But neither is it to succumb to the current fashion of exaggerating China’s military capabilities. Despite its many achievements, China is still a weak state and, as Andrew Shearer points out, its transition to exercising influence as a sea power has provoked region-wide balancing behaviours.70 These behaviours include hard-power balancing through the acquisition of more modern military equipment and soft power balancing by politically resisting Chinese coercion and with a shift recently from softer to harder forms of balancing through new naval and air acquisitions. As time goes on, neighbours around China’s periphery may also feel compelled to field similar capabilities in order to address the growth in Chinese long-range strike assets.71 Ongoing requirements of China’s naval and air forces to secure Chinese near-seas priorities make it highly unlikely that a force that is still modest in size will be able to sustain a robust top-end footprint in the far seas, no matter how much its capabilities improve.72

Conclusions

In our view, China may soon be approaching the zenith of its power as its economy encounters serious structural impediments and demographic barriers to growth. This will also have important implications for the opportunity costs forgone of ever-increasing defence expenditure in a technological arms race against the United States, which Beijing cannot hope to win. Our analysis portrays a China in which worsening domestic problems will remain the leadership’s highest priority and addressing such concerns will take up an increasing share of economic resources and national wealth. The Communist Party leadership will struggle to keep a lid on growing popular discontent, which may have implications for its very survival.

We have also described a lonely power that has very few friends in Asia. Although China’s world view of itself is shaped by strong historic impulses of a hierarchic order with itself at the apex, very few countries in the region

70 Andrew Shearer, ‘Southeast Asia and Australia: Case Studies in Responding to China’s Military Power’, in Tellis and Tanner, China’s Military Challenge, pp. 241-75.
72 Erickson, ‘China’s Modernization of Its Naval and Air Power Capabilities’, p. 96.
Paul Dibb and John Lee

appear willing to concede to China the status of the dominant power. Indeed, it is more likely that countries such as the United States, Japan and India will concert together—either directly or indirectly—against an increasingly aggressive China.

In military terms, China's Achilles heel is that it lags at least twenty years behind the United States in key high-technology areas. The fact that China has no experience whatsoever of modern war and its military hierarchy depends crucially on loyalty to the Party means that China's actual war-fighting prowess must be in serious doubt. Moreover, China's military build-up is causing a classical response in kind as countries such as Japan, India and some Southeast Asian countries acquire advanced maritime military forces in order to check China.

In summary, as The Economist observes: China needs Western markets, its neighbours are unwilling to accept its regional writ, and for many more years the United States will be strong enough militarily and diplomatically to block it.73

What does all this mean for Australia's national security planning and the forthcoming Defence White Paper in 2015? First, the most important point to make is that any suggestion the United States should move to one side in Asia to make strategic space for China should be rejected. China is not now or foreseeably a strategic peer of America's and any move by Washington to concede China's so-called legitimate strategic interests would smack of appeasement; and offered unnecessarily and for little conceivable gain. So, when Beijing proclaims that the entire South China Sea is a core strategic interest, a term traditionally reserved for Chinese claims over Taiwan and Tibet, China's maritime expansionist ambitions should be firmly resisted.

Second, Australia does not need to structure its Defence Force for war with China. Beijing is not developing the conventional forces with which to invade or directly attack Australia. But we should develop the high-technology naval and air assets necessary to contribute to any Allied conflict in the region where we might need to make a contribution or where Australia needs to help resist Chinese military adventurism. Developing these capabilities will further complicate the strategic and operational environment for a still isolated China, which will in turn place further constraints on, and likely encourage greater caution from Beijing. In Northeast Asia, this would suggest niche contributions from us in such areas as submarines and air power. Our Army cannot make a difference to conflict outcomes in Northeast Asia. Closer to home, however, we could make a much more substantial contribution by having the capability to blockade the straits of Southeast Asia in the event of a serious war in Northeast Asia involving the United States.

73 The Economist, 23 August 2014, p. 9.
Third, short of military conflict Australia must be able to resist Chinese coercion—whether by military or other pressures—with regard to our own direct security interests, including if necessary our economic security. We also need to be capable of countering coercion in our region of primary strategic interest—particularly Southeast Asia. It is in Australia's crucial strategic interests for Southeast Asia to avoid being dominated by China geopolitically or becoming a Chinese security domain. Southeast Asia forms a strategic shield to Australia's vulnerable northern approaches and Canberra needs to place high priority on strengthening its relations with Southeast Asian countries, particularly in the defence arena, and to help them resist Chinese coercion.

Paul Dibb is emeritus professor of strategic studies at the ANU’s College of Asia and the Pacific. He is a former deputy secretary of Defence, director of the Joint Intelligence Organisation and head of the National Assessments Staff for the National Intelligence Committee. On behalf of DFAT, he represents Australia at meetings of the ASEAN Regional Forum’s expert and eminent persons group. paul.dibb@anu.edu.au.

Dr John Lee is a senior fellow at the Hudson Institute in Washington DC, an adjunct associate professor at the University of Sydney, a visiting senior scholar at the Institute for Southeast Asian Studies in Singapore, and a Director of the Kokoda Foundation. He is current providing advice to the American and Japanese governments on economic and security futures in Asia. johnleeemail.
From Strategic Security Risks to National Capability Priorities

Rick Nunes-Vaz, Steven Lord and Daniel Bilusich

Since 9/11, many western nations have re-framed their national security decisions in terms of strategic risk management. All have undertaken risk assessments, but valid translation into capability priorities has been abdicated largely by transferring priorities directly from risk magnitudes. Treatment priorities should be determined from risk reduction benefits in relation to costs, but a method for assessing risk reduction effects has been broadly elusive. This article shows, in a pragmatic way, how treatment options are generated, and how the capabilities that contribute most to risk reduction can be identified. These should be priority targets for the investment of limited resources.

Over the last ten years many western nations, for example, the United Kingdom, Canada, the Netherlands, the United States and Australia, have adopted risk as a central part of their national security decision-making and prioritisation. Each has initiated a process of strategic national security risk assessment that feeds and informs discussions on capability priorities, in turn, informing resource allocation decisions. There has been

5 Attorney-General's Department, Guide to Australia's National Security Capability (Barton, ACT: Commonwealth of Australia, Attorney-General's Department, 2013).
commentary on the broad structure of this process (Figure 1), which translates essentially to the well-known steps in an integrated risk management or risk governance cycle.

**Figure 1: National security risk methodology**


However, the step from risk assessment to national security capability priorities is not well bridged and, in practice, relies on face-to-face discussion between experts and stakeholders without a tailored, designed process to guide those deliberations. The Dutch take this approach:

Based on the risk assessment of all the scenarios analysed, an investigation is conducted to find out which capacities [capabilities] are already available and which of these could contribute to a reduction of the impact or the likelihood ... The capability analysis takes place in a working group that includes all relevant experts and interests...

Their process is typical in its use of risk-scoring, exemplar scenarios and expert panels. An exemplar scenario is intended to represent a class of ‘what ifs’ about the future, for example, a chemical attack in a city’s subway transport system. Expert panels generate risk scores by assessing the likelihood of each scenario and its impact on objectives. The same panels may then discuss capability requirements for addressing high-scoring risks.

---


10 Bergmans et al., *Working with Scenarios, Risk Assessment and Capabilities in the National Safety and Security Strategy of the Netherlands*, p. 11

11 Mennen and van Tuyll, ‘Dealing with Future Risks in the Netherlands’.
This largely common approach, is potentially deficient in a number of important respects:

- The validity of decisions based on the kind of risk-scoring seen in national assessments, using a risk assessment matrix, a probability-impact grid, or a likelihood-impact diagram is now much disputed and largely discredited. The key issue here is a failure to account for the many uncertainties within the assessments.

- There is often a mismatch in level of detail between the risk assessment process and capability analysis, despite some recognition of the need to match levels of detail: “the incident scenario must be so specific that it is possible to deduce from it which capabilities will have to be brought to bear in that scenario”. The methodology must support traceability between capabilities and their risk-reducing effects.

- It is generally possible to treat a risk in several different ways, for example, by deterring, preventing, or disrupting an attack, or by protecting its potential targets from harm. Alternative approaches usually reflect differing security philosophies, e.g., a political preference for prevention. However, the risk-reduction effectiveness of available alternatives, and hence the ability to compare their benefit-cost is not, methodologically or practically, well understood or supported. It is commonly beyond the cognitive reach of experts to make comparative evaluations in the absence of a systematic method to guide their thinking.

- A capability is generally ascribed higher value, and hence higher priority, if it contributes to risk reduction in several scenarios, particularly if there are many such scenarios. For example, if the process identifies four cyber and two terrorist scenarios, then cyber-related capabilities may gain prominence (and priority) because they

---

appear more often in the analysis. While there is clearly additional value in capabilities that contribute broadly, it is important that the balance of exemplar scenarios appropriately reflect their relative weights in future possibilities, in order to avoid a potentially unwarranted bias of priorities.

- Often the role of exemplar scenarios is misunderstood, leading to an inability to create an appropriate balance of scenarios across the spectrum of threats, as discussed in the last point. Exemplar scenarios represent mutually exclusive portions of the future containing many possible pathways to impact. Typical resource limitations in the assessment process mean that each and every possible configuration of events leading to impact cannot be considered as a separate scenario. For practical purposes, exemplar scenarios should therefore be seen as classes of pathways. The danger inherent in such compression is that experts may assess them literally rather than as broad expressions of future possibilities.

It is known that judgments of experts and stakeholders relating to capabilities and priorities are strongly influenced by the particular method chosen for risk reduction assessment.\(^1\) It is therefore very important to report the practice of assessment leading to capability prioritisation to enable constructive critique and improvement.

This article reports a method for identifying which of a nation’s (existing or proposed) capabilities provide disproportionate value (in risk reduction terms) in the treatment of a spectrum of national security risks. Such capabilities perform critical functions in the national security architecture, which means that deficiencies or vulnerabilities associated with their roles are the logical targets for investment of limited resources. The method is founded on the risk standards,\(^2\) but is intended to address the methodological deficiencies noted above.

The approach utilises a construct termed a risk pathway, which is a more detailed version of the commonly known ‘bow-tie’ diagram,\(^3\) and is a more pragmatic form of engineering approaches like coupled fault trees and event trees.\(^4\) Risk pathways are constructed to a level of detail that supports

---


assessment of the roles of capability, although not as individual contributions. Work elsewhere, associated with the concept of security-in-depth (SiD),\textsuperscript{21} has shown that the only way to manage the complex interdependencies of capabilities and assess their risk reduction contributions, requires that they be considered in ‘packages’. High-level concepts familiar in the national security lexicon such as prevent, prepare, respond and recover,\textsuperscript{22} are related to, but not quite correct as choices for these packages, as we discuss in the next section.

Following a brief overview of the security-in-depth framework, ‘Risk Assessment Using Risk Pathways’ illustrates the construction of example risk pathways that are appropriately matched to the needs of capability planning. ‘Risk Evaluation’ highlights inadequacies of the risk matrix for determining risk treatments and capability needs, while ‘Risk Treatment’ discusses how this is much more effectively achieved. From the identification of required capabilities ‘Identifying Capability Priorities’ sets out the principles by which particular capabilities that are critical to risk reduction may be identified. The discussion then addresses the issues associated with aggregating capability priorities across risk pathways in order to gain a sense of strategic priorities. Finally, the article concludes with the advantages of the advocated approach over current practice.

**Brief Overview of ‘Security-in-Depth’**

The security-in-depth (SiD) framework\textsuperscript{23} is based on a hierarchy from *security controls* (the physical, technical, procedural elements of capability) that perform or contribute to *security functions* (higher-level constructs or ‘security verbs’ that include detection, alert, response, delay, neutralise, etc.), that, in appropriate combinations, constitute *security layers*. A typical security layer includes detection, alert and response functions because detection capabilities in the absence of a response, or response/neutralisation systems without a cue to act, are impotent in risk management terms. Security layers, as integrated sets of functions, are the smallest meaningful aggregation of capabilities that can stop harmful events or diminish their consequences.

Figure 2 shows a bow-tie diagram that conceptualises all possible pathways from threats to consequences. Superimposed onto the bow-tie are the

\begin{quote}
Attorney-General’s Department, *Guide to Australia’s National Security Capability*, p. 13.\textsuperscript{23}
\end{quote}
seven layers of the strategic SiD framework.\textsuperscript{24} Consequences (on the right side of Figure 2) are represented in two ways. Immediate impacts (or effects) are measured in terms such as lives lost, dollars incurred, disruptions to or losses of essential services, etc. However, depending on the resilience of physical, economic, infrastructural or social systems, these effects may or may not escalate into impacts of national security significance. For example, soon after the London bombings of 2005 the transport network was still operational through most stations (infrastructure resilience), and London commuters were still willing to use public transport (social resilience) in the face of remaining uncertainties.\textsuperscript{25} Without such resilience, national impacts would have been far greater.

![Figure 2: The Security-in-depth framework](image)

The seven strategic layers of the security-in-depth framework (in blue), acting to reduce the probability of threat-initiated events and their potential to generate harm. The layers are deliberately orientated horizontally or vertically. Vertical alignment indicates a passive layer that requires pre-positioned capability. Horizontal alignment indicates an active layer with potentially many moving parts. Enabling capability and arrangements are represented within the construct called ‘prepare’. All elements are described in the text.

Each security layer should be seen as an integrated set of functions, performed by many inter-dependent controls or capabilities, and individual capabilities can contribute to functions in more than one layer. The layers are aligned in sequence, that is, if ‘shaping’ does not resolve the threat, then

\textsuperscript{24} This is the strategic version of the security-in-depth (SiD) framework, extended by three layers (shape, resilience and investigate) when compared with the version published in ibid.

it may be 'deterred'. If not 'deterred' then the attack may be 'prevented', and so on. If the first six layers fail, the 'investigations' layer may help identify perpetrators or their associates in order to reduce risks associated with future events.

This layered construct is valuable for its completeness, that is, for illustrating the full set of opportunities for potential intervention and treatment in sequential stages from threat emergence to national impact rather than, for example, focusing on one aspect such as prevention.

A further dimension of the SiD construct involves a concept called 'prepare' (in Figure 2). Preparation manages the security risk that arises from internal failures of the security enterprise itself, rather than its failure to manage external threats (or hazards). Such failures may arise from poor organisational structures and arrangements, a failure to perform or deliver a role through poor resourcing, unreliable systems, etc., or potentially through the malicious actions of insiders.

Preparation in the SiD framework represents everything within and associated with the enterprise that must align in order for security to perform effectively. It is sub-divided into 'action', 'management' and 'policy' levels (note the distinction from 'layers') and is addressed in more detail elsewhere. The concept represented by 'prepare' should be considered an enabler rather than a layer because, even if it performs perfectly, it does not reduce security risk (the criterion used to define security layers). If it performs less than perfectly it has a negative effect on the enterprise’s ability to reduce security risk.

**Risk Assessment Using Risk Pathways**

For comparison with the stages of Risk Assessment, Risk Evaluation and Risk Treatment in the national and international standard the following three sections are titled to match.27

**STRATEGIC OBJECTIVES**

The Standard defines risk in terms of impact on objectives. National security objectives are defined or articulated within a nation's suite of strategic documentation, and are commonly couched in terms of limiting

---


- physical harm
- social/psychological harm
- economic harm
- reputational harm
to the nation or its citizens, and
- violations of sovereignty or territorial integrity.

Strategic risk assessment is then an examination and evaluation of threats to those objectives.

**Sources of Risk**
Horizon scanning, scenario analysis and other techniques are typically used to anticipate, identify and assess all potential sources of risk (i.e., threats). National security strategies commonly identify these risk sources to include:

- State-based conflict
- espionage
- terrorism, and
- organised crime.

Depending on a nation's view of national security, they may also include natural hazards, such as:

- infectious human, animal or plant disease
- flood, fire, earthquake etc., and
- industrial accident.

**Articulating Risks in Terms of Pathways**
To reveal the contributions of national security (NS) capabilities to the reduction of risk to objectives requires the articulation of national security impacts from each risk source in the form of a risk pathway. A bow-tie

---

*United Kingdom; US Government, The National Strategy for Homeland Security; Kevin Rudd, Speech by the Prime Minister to the Parliament: The First National Security Statement (Canberra: Department of the Prime Minister & Cabinet, 2008).*

*29 ‘National security impacts’ is shorthand for impacts to national security objectives.*
diagram is known to be insufficiently detailed to support this.\textsuperscript{30} Figure 3 shows a generic risk pathway,\textsuperscript{31} associated with a malicious threat such as crime or terrorism. Once the actors’ intent is formed, they must acquire capability, formulate a plan and conduct an act. Their actions are, in general, intended to benefit a cause (shown in the lower ‘return loop’). The act may itself be of national security significance or it may trigger a cascade of effects that generate impacts of national security significance.

![Diagram of National Security Risk Event]

**Figure 3: National Security Risk Event**

A generic representation of a national security risk event as a pathway from the emergence of threat actors to the generation of national security impacts.

**REPRESENTING THE SPECTRUM OF THREATS: COMPLETENESS**

Flexibility and adaptability are present and valued in capabilities that play broad emergency management roles. Other capabilities, however, though much less flexible are equally important in other scenarios, such as sensors for detecting specific harmful chemical or biological agents. To understand the relative priorities of such diverse capabilities first requires a comprehensive set of risk pathways to represent the scope of potential problems.\textsuperscript{32}

The challenge in developing such a set, however, lies in the difficulty of ensuring that it is comprehensive, yet not repetitive. As a primary objective of the whole process is to gain a sense of capability priorities, it is important not to over-represent some concerns (pathways) relative to others. To achieve this, the set of risk pathways should be developed by all relevant...


\textsuperscript{31} We use the words ‘pathway’ and ‘scenario’ interchangeably from here, understanding that these scenarios are of the specific type, specifying a risk source and how it leads to impacts.

stakeholder agencies with three over-riding guiding principles as noted in other scenario studies:\textsuperscript{33}

- Each risk pathway must be a complete sequence from the emergence or appearance of the threat or hazard through to its generation of harm.

- Each risk pathway must be mutually exclusive.

- Each scenario should translate into a distinct pathway inasmuch as each differs in a meaningful way from others, with regard to capability and the scope of the national assessment. It is possible to generate many attack pathway variations, even though each would stress treatment capability in largely the same manner. Judgment is needed to maximise coverage while minimising redundancy with respect to capability-needs assessment. Subsequent assessment of any particular pathway requires consideration of all the variants ‘compressed’ within its representation, rather than taking a literal view.

\textbf{LEVEL OF DETAIL}

The level of detail in Figure 3 is, in practice, too coarse to be useful. The risk pathways are expanded in relation to each source of risk and the expansion continued to a level of detail that matches the understanding or articulation of risk treatment (national security) capabilities. In practice there will be iteration between pathway detail and the matching capability discussion, as we address in ‘Risk Treatment’ below.

An example pathway representing an unspecified terrorist attack is shown at the foot of Figure 4. Each step in the pathway represents an opportunity for intervention and the application of security capability. Two elements of the pathway, that is, ‘motivated actors’ and ‘acquire means’ have been expanded in the upper part of Figure 4. There are two implications of expansion in this manner. The first is that expansion reveals additional opportunities to intervene and manage, in these cases, the probability of the pathway progressing to completion. The additional detail is useful if there are capabilities that can be applied or developed to reduce these probabilities. The second implication is that some sub-pathways can be developed and re-used, in similar form, in a number of different pathways. We have found a modular approach to pathway construction useful in practice.

Specifically, we have found two types of modular sequence useful. The radicalisation and weapon development modules are examples of what we call precursor sequences. They are invoked whenever a complex process remains implicit within an element of the risk pathway.

![Diagram of Terrorist Attack Pathway]

**Figure 4: Terrorist Attack Pathway**

A plausible representation of a terrorist attack pathway (lower part of the figure), with two precursor sequence modules (i.e., weapon development and radicalisation) essentially expanding the target elements of the main terrorist pathway. Only the terrorist pathway is a risk pathway because it includes a defined event and its consequences. The dashed elements in the terrorist pathway are intended to represent activities that may or may not occur.

The other type of pathway module is a ‘disabling sequence’; an example of which occurs in Figure 5. The disabling sequence allows us to represent compromise, disruption or failure of the security machinery itself (represented in blue without detail in Figure 5). Compromise may occur as the result of deliberate or malicious interference (or deliberate inaction) arising, for example, from sabotage or corruption by organised crime or ‘insider’ agents, as shown. It can also arise from a failure in preparation (as
we discuss below in the section ‘Preparation’) when, for example, a critical
detector fails due to inadequate maintenance, and the need for a
replacement was never considered. Even though disabling sequences
represent ‘attacks’ on the machinery of security, their expression as
pathways reveals the nature of opportunities to intervene to preclude or limit
the potential for compromise.

A third type of pathway module involves follow-on or indirect impacts from
events (with two such modules illustrated in Figure 5). Events from several
sources of risk, such as state-conflict, state fragility, terrorism or pandemic
may lead to follow-on impacts from mass migration or civil unrest, for
example. Again, the expression of follow-on effects as pathways reveals the
opportunities to contain, protect against, or be resilient to their evolution into
national security impacts.

Risk Evaluation

Common practice, at this stage, would see subjective risk evaluation of each
pathway, in which experts assess the likelihood (roughly, the chance that
each scenario will occur, or how frequently it is expected to occur in a future
time period), and the band of consequences that most closely represents the
risk’s impact. By this process each scenario is allocated into a cell of a
matrix representing risk magnitude.

However, it is inappropriate to assign priority to a capability based on the
magnitude of risk of a pathway that invokes it, that is, a high risk pathway
implying high priority capabilities. Firstly, while there may be a strong
imperative to mobilise, innovate and develop resources to tackle the,
possibly existential, risks in the high likelihood, high consequence corner of
the matrix, greater risk reduction for the same cost is generally the guiding
principle in government decision-making. It is the reduction in risk from the
use of particular treatments, not whether the treatments address a high risk,
which is most relevant to prioritising capabilities.

37 Australian Government, Office of Best Practice Regulation, Best Practice Guidance Note: Decision Rules in Regulatory Cost-Benefit Analysis (Canberra: Department of Finance and Deregulation, 2009).
A more detailed example of pathways representing adversary-initiated state-based conflict (centre) leading to national security consequences. Friendly, defensive capabilities are represented without detail in blue. The adversary has employed two additional (modular) sequences at top, that is, a cyber capability development precursor sequence which supports another sequence intended to disable defensive systems. Similar modularity is seen in the lower part of the figure in the form of two follow-on pathways (community violence and mass migration).
migration) triggered as indirect effects of the conflict. Many other indirect effects could be explored. Every step in a sequence provides opportunities for countermeasures and hence the identification of national security capability needs.

In fact, it is easy to manipulate, whether intentionally or inadvertently, the magnitude of risk that is attached to a scenario when using the risk matrix approach. This can be done by changing scenario detail. More specific scenarios correspond to a smaller proportion of the future, which means lower likelihood and therefore lower risk. A high risk pathway in the matrix approach may be broken down into several lower probability or lower consequence sub-pathways. This is a common tactic when risk management is focused on compliance. One can always make a project comply with risk tolerance limits by increasing the specification of relevant scenarios. This is just one aspect of the ambiguity risk matrices generate because they provide no information on the issue of risk aggregation. Appropriate aggregation would show that the low risk sub-pathways add up to the high risk parent pathway, so the assessment should be independent of scenario specification. As scenario specification does have a direct bearing on the risk assessment process, translation to capability priorities requires care to manage risk magnitude and risk aggregation appropriately. As already noted, priority should relate to risk reduction rather than risk magnitude, as discussed in the next section. The aggregation problem is discussed in the section ‘Identifying Capability Priorities’.

A further point, although still related to aggregation, is about inter-dependency between national security threats. Capabilities usually treat sources of risk, and those sources may appear in several risk pathways, for example, anti-virus software treating cyber threats that occur in both ‘terror’ and ‘crime’ pathways. To assess the value of such capabilities we must be able to examine all relevant pathways and aggregate (in some sense) their contributions to the treatments that reduce risk. The inter-relationship between threats means that the capabilities required to treat a particular risk (represented by one marker in a risk matrix) should not be considered in isolation from their role elsewhere (associated with other markers in the matrix).

We therefore advocate that evaluating the risk magnitude of pathways (scenarios) is not useful at this stage. As distinct from current practice, we do not risk-score pathways until they have been developed to the resolution required to judge the needs and values of risk treatments.

Risk Treatment

Figure 6: Risk Treatment

Threats and hazards represent sources of risk with respect to strategic security objectives. Treatment of risk requires the performance of functions or tasks which require the application of capabilities.

Risk treatment (Figure 6) is achieved through the effective application of national security capabilities (e.g., Customs’ maritime patrols), performing required functions (e.g., law enforcement). A sense of priorities comes from understanding the relative contributions to risk reduction that can be achieved, and this is assessed using the SiD framework.

Allocating Capabilities to Risk Treatment

Figure 7 shows an expansion of the generic pathway relating to terrorism at the foot of Figure 4, truncated where the terror act occurs. A notional border has been added, distinguishing offshore from onshore elements of this particular pathway, and four regions are identified (international, border, national and local) representing different contexts for security intervention.

41 Attorney-General’s Department, Guide to Australia’s National Security Capability.
A more detailed representation of the cross-border movement of terrorists and their capability, highlighting the way that context focuses the interpretation of relevant capability.

The figure shows the layers of security that might be applied to treat the risk. For example, the explicit pathway component marked ‘radicalise’ might be treated using capabilities that contribute to ‘shaping’ the relevant offshore communities. Multiple instances of security layers are often needed, so a hypothetical overseas terrorist may be deterred or prevented prior to their overseas departure (where intelligence sharing and cooperation are strong); at the border (immigration officers); within the national system (federal police); or by the local security associated with the target itself (gates, guards, etc.).

Figure 8 shows an example of the way that expansion of the pathway proceeds to support understanding of capability needs. The ‘develop capability’ element of Figure 7 requires the terrorists to acquire expertise, acquire materials and then manufacture the capability. The lower part of Figure 8 shows examples of capabilities that might prevent each of the steps. Prevention requires the successful performance of detect, alert and respond functions which, in turn, demand the successful operation of potentially many integrated capabilities, some examples of which are shown in the figure. By increasing the level of detail in the pathway, commensurate detail can be developed in the array of potential treatments, from which the implications for capability can be assessed.
An expansion of ‘develop capability’ from Figure 7, with example capabilities that might be used to treat the elements. Note that the prevent layer requires the performance of all three functions of detect, alert and respond and each, in turn, requires appropriate capabilities, only some of which are shown for illustration.

Other layers, beyond prevent, are then examined in similar manner. Expansion of pathways should ensure that the resolution of pathway elements matches the resolution of capabilities being considered. In this way, the expansion and discussion of risk pathways is intimately tied to the discussion of capability needs, and should not be separated as a sequential process. The net result is a layering of capabilities along the pathway between the sources of risk and impacts.

Figure 9 shows a further illustration of the use of pathway sequences to identify capability needs. The disabling sequence illustrated in Figure 5 shows an adversary’s use of cyber capabilities to compromise national defence systems involved in state-based conflict. The disabling sequence itself provides opportunities for treatment (countermeasures) as shown, in partially developed form, in Figure 9.

---

Figure 9: Cyber Countermeasures

Example countermeasures associated with shaping, deterring, and preventing sabotage or foreign interference by cyber means. These are countermeasures intended to reduce the probability and the potential impacts of adversary action.

LAYERS AND THEIR FUNCTIONS

Shape, prevent, contain and investigate layers (the horizontal layers in Figure 2 and Figure 7) all rely on the successful performance of variants of detect, alert, respond and impede functions. A successful investigate layer, for example, relies on the detection of clues (evidence), piecing enough parts of the puzzle together to warrant raising the alarm, and some interdiction capabilities.

In general, the passive layers (deter, protect and resilience) rely on single, context-specific functions. Deterrence largely derives from the perceived effectiveness of security (ignoring the contribution from the severity of penalties, if convicted),\(^{43}\) which can be manipulated using real or purported capability. Protection is highly context-dependent, so protection from crime, bombs, fire, floods, disease, etc., all require very different and very specific capability sets. Suitable protections can be identified by asking “what are we protecting, and from what?”

Some aspects of resilience are passive, such as the social resilience of London commuters and their use of public transport after the 2005 bombings. This kind of social resilience is developed in the system prior to an event and derives from people’s perceptions of risk, which is also subject

---

to enhancement or manipulation, through communication. Active components of resilience are based on detect, alert and respond functions. For example, resilience in the London transport system relied on detection of non-functioning components and re-routing to restore operations to the highest possible level.

**PREPARATION**

Primary security capability relies on the availability and quality of enablers such as training, communications, information systems, logistics, etc. These enablers will often be as important to the effective performance of security as the security systems themselves. Preparation also involves alignment between management and policy and the needs of security risk management. A mismatch, for example, between a role and the authority to conduct that role, or its foundation in law, may disable or limit the effectiveness of security. Thus failure of enabling systems, collectively termed ‘preparation’, can substantially compromise security. Preparation is closely related to the notion of enterprise or intrinsic risk.

Figure 10 shows a sequence of elements for which inadequacy or failure of any one may lead to compromised risk treatment capability. The items shown are illustrative only, and are not intended to be exhaustive. This kind of breakdown and analysis, which is related to fault tree analysis, can be used to understand the context and implications of failures of enterprise risk management.

**Figure 10: Enterprise Risk**

A non-exhaustive sequence representing elements associated with intrinsic or enterprise risk. The failure of one or more elements leads to compromise of security.

In the feedback arrow in Figure 10 we see another role of ‘prepare’, which is to promote and enable self-awareness, contextual sensitivity, and the ability and propensity to monitor and assimilate lessons from both internal and external events, in order to identify the drivers for development.
Identifying Capability Priorities

INVESTMENT WITHIN LAYERS

The processes in the previous section are intended to lead, through cross-departmental and stakeholder consultation and discussion, to a determination of treatment options, in the form of a list of capabilities (and their enablers) required to manage the risks of all relevant pathways. The purpose of this section is to guide, using security-in-depth, the identification of critical capabilities from what may be a long capability list. A critical capability is one that a) when absent or damaged ‘substantially’ impairs the overall system’s performance, and b) should be considered a target for investment because each increment of enhanced performance has high payoff in overall risk reduction effect. Performance should be judged in the context of the whole layer.

The security-in-depth framework works on the basis of risk reduction contribution. In SiD, within any single (multi-function) layer investment should be focused on the security function that performs least well.44

To make these assessments it is not always necessary to know precise performance figures for the functions. Consider the case of preventing the cross-border passage of contraband through an airport or seaport. Ports are designed to ensure that there is a high probability of both raising an alert and successfully interdicting the threat when a true-positive detection is made. Successfully preventing the passage of contraband therefore often relies most heavily on detection of the threat, and the prevent layer is therefore compromised inasmuch as detection probability remains comparatively low. If these assumptions are correct, the prevent layer’s effectiveness would be improved by investing in detection.45 Domain experts are often able to assess which function, in any given layer, is comparatively weak, and they or others can usually identify how to best enhance its performance within a defined budget.

Further examination shows that there are two ways to lift the performance of the detection function, and hence two candidate strategies for investment.46 The first is to improve the performance of detection systems in situ, perhaps by increasing the percentage of goods that are screened or by investing in more advanced screening technologies to reduce the number of false negatives and false positives. Alternatively, detection might be improved by enhancing the probability that cueing information will be received prior to the contraband’s arrival at the border. One way to do this is to raise the quality

45 Assuming that each dollar spent generates a similar incremental improvement in each function.
46 Assuming that performance is not currently limited through compromised enablers like training.
of intelligence feeds into border operations, which might be achieved by enhancing information sharing arrangements with other nations.

Given the assumptions stated or implied in this example, the security objective of contraband interdiction implies that enhancement of detection capabilities is critical. Decisions on the best way to achieve that enhancement should then be resolved using cost-benefit principles.

**INVESTMENT ACROSS LAYERS**

Given that it is possible to identify the most cost-effective ways to enhance a security layer with respect to a given risk pathway, the next question is, which layer(s) deserve(s) greater investment. It is known that investment should be directed, counter-intuitively, to the layer which is already the best performer. The layers are sequential, independent risk reduction systems which means that, theoretically, effective security only requires one layer to be successful. If prevention is successful then protection, containment and resilience are, in that instance, not required. Requiring only one successful layer means that investment should be directed to the layer that is most likely to achieve the overall security objective, that is, to the best performing layer. This conclusion, however, carries a number of caveats.

The first caveat is that layers perform differently with respect to different threats and scenarios. For example, choosing to protect against violent extremism requires that all potential targets be protected (because an intelligent adversary can learn which remain relatively unprotected). This implies that a strategy focusing on protection will either be very costly or relatively weak. Contrast this with a bio-threat, for which protection (e.g., vaccination) may be much the most effective strategic approach.

The second caveat notes that a perfect layer is often difficult to achieve and investment should be sensitive to the returns that can be achieved, particularly where they progressively diminish. For example, deterrence may perform quite well against terrorism but statistics indicate that some attackers will remain undeterred. Rather than trying to further enhance deterrence, which will climb in cost and remain ineffective against the most determined adversaries, it is more beneficial to invest in (an)other relatively effective layer(s).

For some individual risk pathways all layers may be compromised, to some extent, and it may be difficult to identify the 'best' target for enhancement. In this case, several layers may contribute to risk reduction, not necessarily equally but perhaps comparably. In some cases, it may be possible to understand trade-offs between layers (the value of focusing investment in

47 Nunes-Vaz et al., 'A More Rigorous Framework for Security-in-Depth'.
some layers rather than others) using robust quantitative analysis.\textsuperscript{49} However, national assessment currently falls into the category of being so complex that even subjective opinions about the sequential actions in risk pathways, and ontological uncertainty about modelling the risk pathway interactions, are likely to lead to inaccurate assessment. Strategic decision-makers may also be less than comfortable with such a step. Since it can be difficult to assess these trade-offs, we consider it more useful to assess all seven layers and ask which functions, in each layer, should be considered the primary targets for investment. The decision to focus investment into a particular layer might be made for reasons other than risk, for example, for political, social or economic reasons, such as the choice to protect particular iconic targets against the potential effects of terrorism. Freed from a requirement to prioritise between layers, each layer should be examined to identify its weakness in order to determine critical capabilities (and their enablers) on a per layer, per risk pathway basis. Illustrative results from such an analysis are shown in Figure 11.

<table>
<thead>
<tr>
<th>SHAPE</th>
<th>DETER</th>
<th>PREVENT</th>
<th>PROTECT</th>
<th>CONTAIN</th>
<th>PASSIVE RESILIENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>detect</td>
<td>alert</td>
<td>respond</td>
<td>detect</td>
</tr>
<tr>
<td>Agency-C Capab-3</td>
<td>Agency-B Capab-5</td>
<td>Agency-D Capab-11</td>
<td>Agency-E Capab-15</td>
<td>Agency-G Capab-17</td>
<td>Agency-C Capab-16</td>
</tr>
<tr>
<td>Agency-B Capab-14</td>
<td>Agency-D Capab-14</td>
<td>Agency-D Capab-14</td>
<td>Agency-E Capab-8</td>
<td>Agency-G Capab-2</td>
<td>Agency-C Capab-14</td>
</tr>
</tbody>
</table>

**Figure 11: Risk Scenario**

The results of analysis of (hypothetical) scenario-X, in terms of the capabilities in each layer required to manage its risk. Critical capabilities are shown in red, noting that they lie inside the functions that were identified as relatively weak in their layer context. \textsuperscript{49} A red capability may represent weakness in an existing system, or it may represent a missing capability that the analysis has exposed. Note that enablers are also included in the diagram (principally, although not exclusively, inside ‘prepare’).

In general, the analysis will identify ‘packages’ rather than individual capabilities that are disproportionately important to a layer’s effectiveness. Figure 11 shows such capabilities in the seven security layers, and the enterprise enablers associated with ‘prepare’, from an analysis linked to hypothetical ‘Scenario-X’. Those critical capabilities might already exist but may be under-performing, or they may not yet exist. This kind of analysis, as an output of stakeholder discussions, becomes an input to decision-making.

DETERMINING STRATEGIC CAPABILITY PRIORITIES

The previous section outlined a process by which critical packages of capabilities are identified in each layer, for each particular risk pathway. Pathways are themselves components in a de-aggregated view of the whole strategic security landscape.

Attempting to aggregate findings by adding up, across all pathways, the number of times a particular package is assessed to be critical will, as noted in the section ‘Risk Evaluation’, lead to bias and sensitivity to the particulars of de-aggregation, that is, the choice and balance of pathways represented. A more justifiable approach involves aggregating a package’s contribution to risk reduction across all pathways. The difference between the two approaches is illustrated through Tables 1 and 2.

In Table 1, each pathway or scenario is risk-scored and each particular package receives an aggregate score across all scenarios from the product of the scenario risk and whether the package was judged to be critical (scoring a one) or not (a zero). Given cost estimates for each package, a form of benefit-cost ratio (BCR) is then generated in the final column. However, as discussed above in the introduction and ‘Risk Evaluation’, this form of scoring attaches importance to risk magnitude and does not assess overall risk reduction contribution. It provides only a weighted frequency of criticality for each package.

Table 1: Notional weighted scoring of capability packages due to binary view of criticality

<table>
<thead>
<tr>
<th>Risk (Expected Impact)</th>
<th>Scenario 1</th>
<th>Scenario 2</th>
<th>Scenario 3</th>
<th>Score</th>
<th>Cost ($m)</th>
<th>BCR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capability Package A</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>11</td>
<td>4</td>
<td>2.75</td>
</tr>
<tr>
<td>Capability Package B</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>9</td>
<td>4</td>
<td>2.25</td>
</tr>
<tr>
<td>Capability Package C</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>6</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>
A stronger basis for strategic investment choices\textsuperscript{50} is based on risk reduction assessment and true cost-benefit techniques. In Table 2 stakeholders assess the risk reduction provided by a capability package for each scenario, which can be elicited directly, or by subtracting estimates of residual risk from an initial risk. The total risk reduction of each package is then achieved through summation, leading to a more appropriate benefit-cost ratio which, in this illustration, implies different investment priorities.

It is in the problem of strategic aggregation that we believe the method described here is most useful. By breaking the complex interaction of interventions and threats into pathways, layers and functions in the manner that leads, through discussion, to constructs like Figure 11, stakeholders are able to greatly improve their subjective estimates of probability and consequences reductions.

Table 2: Notional benefit-cost of capability packages using a risk reduction view of criticality

<table>
<thead>
<tr>
<th>Scenario 1</th>
<th>Scenario 2</th>
<th>Scenario 3</th>
<th>Total Red.</th>
<th>Cost ($m)</th>
<th>BCR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected Impact before</td>
<td>4</td>
<td>2</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expected Impact after</td>
<td>2.5</td>
<td>2</td>
<td>4.5</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Capability Package A</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Capability Package B</td>
<td>1.5</td>
<td>1</td>
<td>6.5</td>
<td>4</td>
<td>3</td>
</tr>
</tbody>
</table>

Neither approach, however, will necessarily identify a capability package that has relatively small benefit in any particular scenario, but which aggregates to a high total risk reduction across all pathways. This flaw should be highlighted in decision-making discussions, and a separate check conducted in the aggregate analysis.

The benefit in the BCR in Table 2 is the expected reduction in impact using the single metric of discounted cost or, more accurately, de-utility measured by discounted cost, for impact.\textsuperscript{51} Strategic agencies may wish to retain several benefit-cost indices, (economic cost, lives lost, reputation, etc.), in which case techniques of multi-criteria optimality or ordering (through Pareto fronts and Pareto domination) can be used.\textsuperscript{52}


\textsuperscript{51} Attorney-General's Department, \textit{Submission to the Productivity Commission Inquiry into Natural Disaster Funding Arrangements} (Canberra: Attorney-General's Department, Commonwealth of Australia, 2014), p. 6.

\textsuperscript{52} Yacov Y. Haimes, \textit{Risk Modelling, Assessment, and Management} (Chichester, UK: Wiley, 2004); J. Figueira, S. Greco, and M. Erehott, \textit{Multiple Criteria Decision Analysis} (New York: Springer, 2005).
Uncertainty is a major issue in subjective estimates (of both benefit and cost) in the complex environment of national security. It is beyond the scope of this article, but there are existing statistical techniques that allow uncertainty in subjective opinion to be elicited and characterised, and to be retained in cost-benefit assessments. When dealing with uncertainty one should look at other possible future values apart from expected value: the industry standard in finance is to consider the worst 5 per cent of outcomes as well as the expected outcome.

As a final point, capability development and acquisition processes generally span several years or even decades, particularly in military capability acquisition. Therefore, discounting or inflating future benefit should consider indicators of trends. An obvious example here is the potential growth (in both frequency and severity) of cyber-attack scenarios. For this reason, it makes more sense to measure risk reduction and the ratio of risk reduction to cost for capability packages rather than simply risk assessing scenarios. A strategic threat may endure, but the performance of capabilities over time generally will not, and it is the latter which is relevant.

Summary

Despite significant and growing arguments in the literature about the flaws surrounding the use of risk (likelihood-consequence, probability impact) matrices in decision-making, national security agencies in western nations continue to use these devices presumably because of their relatively intuitive (but inappropriate) basis for prioritising capabilities and allocating resources according to risk magnitude. While it is reasonably understood and accepted that investment of resources should be governed by benefit-cost considerations, disproportionate allocation of resources to higher risks is often justified through (potentially flawed) assessment that identifies certain risks to be high, and because assessing risk reduction benefit is felt to be methodologically difficult. No nation has yet articulated a practical method for assessing the strategic risk reduction value of particular investments, or a means to compare alternative risk treatment strategies. The Dutch do note the importance of considering multiple factors, not just risk magnitude, but it is still common to see investment decision ratings as an overlay placed directly onto the risk matrix.53

Ultimately we believe that current approaches have limited defensibility according to risk and governmental Standards. The approach in this article addresses these flaws by providing a means to assess and evaluate the risk reduction contributions of treatments and capabilities. We depart from established methods in three fundamental ways. Firstly, there is no need to define the likelihood-consequence bands (the specific cell in a risk matrix)

that a particular threat or hazard belongs to and give it a specific or implied score in isolation from consideration of relevant treatment capabilities. It is far more useful to articulate the path by which a source of risk (e.g., a terrorist) can cause harmful consequences, because it is the steps along the risk pathway that provide opportunities for intervention and the application of capability. Development of risk pathways, by the policy and intelligence communities, working alongside the capability owners and operators ensures that the fidelity of each step in the pathway is tailored precisely to the process of gap analysis and capability-needs assessment.

The second departure from orthodoxy centres around the unit or quantum by which capability values are assessed. Through inter-dependency, any particular capability will play a role that is dependent on the scenario, the context and the presence or absence of other capabilities. This makes it untenable to ascribe an intrinsic value to any capability. Capabilities should instead be assessed in packages. Such packages play key roles in delivering security functions such as detection or response but, ultimately, the value of a capability package should be determined according to the ability of the security layer to which it belongs, to reduce risk. Using the SiD security layer construct makes it possible to gain meaningful assessments of the risk reduction generated by a package of capabilities in a given context (pathway), and thereby a means to assess benefit-cost.

Our third departure was to provide a robust means to identify which of the capabilities or packages should be considered critical to risk reduction efforts on a per-pathway basis, and then how those insights should be aggregated in order to determine strategic priorities for capability enhancement or acquisition. We argue that it is important to generate an appropriate balance across the spectrum of risk pathways representing threats and potential outcomes, because artefacts of an unintended emphasis, for example, on cyber scenarios relative to organised crime, can easily carry through to unbalance assessment of priorities and resource allocations.

We have applied, and subsequently enhanced the approach from our work supporting one of Australia’s strategic agencies. Despite its stronger and more defensible methodology, there remain limitations in our approach and significant barriers to its adoption. Not the least of the barriers is the inertia represented in the simplicity and wide adoption of current methods. More importantly, however, the complexity of national security issues dictates that their translation into useful pathways is likely to demand significant time. One advantage of the biannual Dutch process is its evolutionary refinement, based on accumulation of knowledge about the system. Risk pathways provide a similar mechanism for building and representing the knowledge, insights and opinions of agencies and experts. Mining ‘data’ of this kind and, if appropriate, commissioning new targeted research to understand and refine the pathways should be a requirement of each iteration of national
assessment. In the end, our understanding of the problem is our most important asset for effective risk-informed capability prioritisation.

Rick Nunes-Vaz is Head of the Strategic Security Risk Analysis (SSRA) group in the Joint & Operations Analysis Division of the Defence Science & Technology Organisation, where he leads research to inform strategic decisions in national security, for Defence and non-Defence clients. Rick.Nunes-Vaz@dsto.defence.gov.au

Steven Lord is a Research Scientist in the SSRA Group of DSTO. He uses quantitative robust risk analysis techniques to support operational and strategic decision making. Published security risk research includes analysis of airport front-of-house and the optimisation of controls for security-in-depth. Steven.Lord@dsto.defence.gov.au

Daniel Bilusich is a Research Scientist in the SSRA Group of DSTO. He works on the analysis and implications of strategic risks to Australia’s national security from a range of threats, supporting understanding of treatment option priorities. Daniel.Bilusich@dsto.defence.gov.au
After Afghanistan: A Small Army and the Strategic Employment of Land Power

Mick Ryan

The Australian Army commitments to East Timor, the Solomon Islands, Afghanistan and Iraq since 1999 comprise the longest if not the largest commitment by this small army since its formation. All of these commitments were the result of circumstances largely unforeseen by strategic planners, and demanded significant institutional adaptation to build, deploy and sustain land forces over a period of fifteen years.

The Army is now entering an era where it will be less involved in deploying forces for contemporary warfare, and more involved in transforming and posturing its forces for the contingencies of future warfare.\(^1\) Direction from government in the 2009 and 2013 Defence White Papers confirmed the strategic role and the expeditionary character of the Australian Army.\(^2\) However, this direction needs to be enriched with the lessons from the myriad of operations conducted since 1999. This requires the Army to undertake a form of ‘intellectual pivot’ to learn the lessons of the past fifteen years and think deeply to ensure it is prepared for conflict in the coming decades.

There are three types of lessons that might be gained from reviewing the Australian operational commitment in the past fifteen years. Strategic lessons examine the effectiveness of national armed forces in securing by force the objectives set by political leadership.\(^3\) Tactical lessons review the

\(^1\) However, as recent events in Iraq and Syria have demonstrated, the unpredictability of the global security environment may dictate a shorter inter-war period than we might have anticipated even 12 months ago.

\(^2\) The Defence White Paper in 2009 stated the requirement for “land forces to be undertake combat in our littoral environment and territory, are necessary to secure offshore territories and facilities, defeat incursions onto Australian territory and potentially deny adversaries access to staging bases from which they could attack us. They are also required to undertake amphibious manoeuvre, and stabilisation and reconstruction operations in our immediate neighbourhood, as well as operations further afield in support of our wider interests”. See Commonwealth of Australia, Defending Australia in the Asia Pacific Century: Force 2030 (Canberra: Department of Defence, 2009), p. 60. The 2013 Defence White Paper requires land forces “proficient in joint and multi-Agency operations for the security of Australia and the region”. Commonwealth of Australia, Defence White Paper 2013 (Canberra: Department of Defence, 2013), p. 85.

conduct of land operations to secure operational outcomes and are generally short term in nature.\(^4\) In between are the institutional lessons that will allow the Army’s current and future leadership to gain sufficient manoeuvre room in the policy and budget debates to structure the Army to meet the demands of future conflicts.\(^5\)

The aim of this article is to examine institutional lessons the Army might take from the East Timor, Solomon Islands, Afghanistan and Iraq commitments. As historians and military professionals understand, the immediate period following a war is a rich opportunity for learning the lessons of the recently concluded conflict in order to help inform the institution for future conflicts. The Australian Army must now address this challenge. It must take the lessons of the past fifteen years, use them in concert with the larger body of lessons from previous wars, and apply them to preparing for future military commitments.

**Learning Lessons**

In her book on the evolution of strategy, Heuser has noted that most works on war since antiquity have commenced with a firm assumption that one could learn lessons from past examples.\(^6\) Learning from operational experience, and encoding in the force the key elements of those lessons, is a core function for professional military organisations. Irrespective of the size of the organisation, it must be designed to ensure that it remains fit and effective for future conflicts.

The Australian Army has captured tactical lessons from all of its recent commitments. The Centre for Army Lessons and the Combat Training Centre have executed robust programs to collect lessons from Army task forces, elements and individuals who have returned from operations in Iraq and Afghanistan. This has provided a sound return on the Army’s investment in people, time and resources by ensuring that deploying individuals and units are prepared, based on the most recent lessons from both theatres. The large database of lessons and publications resulting from this collection activity\(^7\) has also resulted in changes to the Army’s certification processes for deploying forces, revised training courses and a range of new doctrinal publications.

\(^4\) For a description of tactical effectiveness, see, ibid., pp. 19-26.

\(^5\) This differentiation between strategic and institutional lessons is discussed in Robert H. Simpson and Mark C. Smith, *Army Adaptation from 1898 to the Present: How Army Leaders Balanced Strategic and Institutional Imperatives to Best Serve the Nation*, Land Warfare Papers no. 98 (Arlington, VA: The Institute of Land Warfare, September 2013).


\(^7\) For example the Centre for Army Lessons has published thirty-five editions of its ‘Smart Soldier’ handbook since 2001, as well as a series of handbooks on leadership from junior non-commissioned officers through to unit commanders, and has standing pre-deployment material for all theatres in which Australians have served since 1999.
As useful as the collection of tactical lessons is for force generation in the Army, it does not provide the full spectrum of lessons that might be taken from the past fifteen years of operations. There has been limited collection and analysis of lessons that might apply to the Army as an institution. These are the lessons of how the Army sees itself, how it executes the full spectrum of raise, train, sustain and adapt functions across time, how it commands and plans, and how the Army interacts with the various non-Army organisations and entities—be they government, other Services or families.

In the remainder of this article, I propose five institutional lessons from the last fifteen years. These form, in my view, the most important institutional lessons for the modernisation of the Army to ensure it retains its effectiveness for future operations. These lessons incorporate the changing character of land warfare, the preparation of land forces to adapt to these changes, and how land forces can work with a range of different actors to build tactical excellence and generate strategic impact.

Lesson 1: Evolution in the Character of Land Warfare

Over the past fifteen years, the Army has relearned that the nature of warfare—humans seeking to impose their will on other human beings—is enduring. This has been best described by Clausewitz, when he wrote that “war is thus an act of force to compel an enemy do our will … war is not the action of a living force upon a lifeless mass but always the collision of two living forces”. Despite this enduring nature of war, a key lesson for the Army from the past fifteen years has been that the character of land warfare has evolved—significantly. There have been several aspects to this evolving character. Some, like the need to achieve influence, support to populations, and sustenance to families, are old themes with new manifestations. Other aspects, such as the ubiquity of technical networking, the availability of low-cost highly lethal weapons, closer integration of conventional and special operations or the greater ambiguity in the strategic outcomes of war, are newer elements in land warfare. Together, they have provide the Army with insights in how much warfare has evolved since its last major conflict in Vietnam, and how it must assume ongoing changes in its approach modernisation.

Even with all of the developments in the technological systems that provide land forces with an unprecedented capacity to monitor the battlefield and share information, the human capacity to seek and achieve surprise remains and will continue to be an enduring aspect of military operations. Various adversaries over the past fifteen years have demonstrated a capacity to watch and learn from friendly forces, to ascertain weaknesses and adapt their tactics—and strategic messaging efforts—accordingly. In particular,

---

Afghanistan has provided a demonstration in how humans will outwit their opponents to achieve surprise, even in what has probably been the most densely surveilled area in the history of warfare.\(^9\)

This human adaptation is a key aspect of the changing character of the threat to land forces. Often described as hybrid\(^10\) threats, contemporary conflicts have often denied land forces a clearly defined enemy. Authors such as Frank Hoffman, Nathan Freier, and Christopher Bowers have proposed similar definitions for this type of threat, broadly defined as the ability to engage effectively in multiple forms of war, simultaneously.\(^11\) This is indicative of the types of threats faced by land forces in the past fifteen years. Multiple reports on trends in future conflict and national security have found that hybrid threats are also likely to remain a persistent element of the environment. Publications such as the UK Ministry of Defence’s *Future Character of Conflict*,\(^12\) the US National Intelligence Council’s *Global Trends 2030*,\(^13\) and the Australian Army’s 2014 *Future Land Warfare Report*\(^14\) have described the convergence of regular and non-state actors as a highly likely feature in future conflict.

Many reports focused on future trends in warfare have also identified the decreasing cost of increasingly lethal weapon systems. The proliferation of improvised explosive devices\(^15\) has demonstrated another ‘low cost’ way that

---

\(^9\) There are multiple examples over the recent years that prove this, including the September 2012 attack on the large Marine Corps base in Helmand province where fifteen insurgents penetrated the perimeter outfitted in American military uniforms, attacked across the north-east flight line, killing two American Marines, wounded nine others, destroyed six AV-8B Harrier jets and caused hundreds of millions of dollars in damage. Matthew Komatsu, ‘Responding to an Insurgent Attack on an Afghan Base’, *New York Times*, 29 October 2013. See also, International Security Assistance Force, ‘ISAF Provides Additional Details on Camp Bastion Attack’, 16 September 2012, <http://www.isaf.nato.int/article/isaf-releases/isaf-provides-additional-details-on-camp-bastion-attack.html> [Accessed 13 September 2014].

\(^10\) Bowers notes there is no commonly agreed definition for ‘hybrid threats’. The US military definition is: “the diverse and dynamic combination of regular forces, irregular forces, and/or criminal elements all unified to achieve mutually benefitting effects”. Christopher Bowers, ‘Identifying Emerging Hybrid Adversaries’, *Parameters*, vol. 42 (Spring 2012), p. 39.


\(^12\) See UK Ministry of Defence, *Global Strategic Trends—Out to 2045*, fifth edition, Strategic Trends Programme, 30 June 2014.


\(^15\) The impact of IEDs on contemporary warfare has been examined in multiple works. Some of these include Max Boot, *War Made New: Technology, Warfare and the Course of History 1500 to Today* (New York: Gotham Books, 2006); David Kilcullen, *The Accidental Guerrilla: Fighting
future adversaries might seek to influence or attempt to impose their will. The Army should expect that it will see these most future missions. Cheap, precise lethal weapons are becoming ubiquitous. Almost anyone can use Google Earth to gain high fidelity targeting information. Coupled with cell phones and cheap, man-portable weapons, the barriers have been significantly lowered for those who might wish to threaten friendly land forces now and in the future.

The Army has regained an appreciation of the need to influence populations as a core competency of military forces and senior leaders. Among other sources, this is examined in Emile Simpson’s recent book, War from the Ground Up (2012). Simpson proposes that this greater understanding of influence is a profound change that is being facilitated by the inter-connectedness provided by the information age, and a more ‘population-centric’ approach to war, and has consequent importance for strategic influence operations.

But lessons on influence have extended beyond the conduct of land operations. To generate strategic effect, armies must also have the capacity to influence their coalition partners and allies in developing strategic objectives as well as the many subordinate and supporting operational plans. This includes informing and influencing military and other agency plans that may impact on the conduct of land operations. Over the last decade, the placement of senior planners in such diverse locations as NATO headquarters in Brussels, within the Joint Staff in the Pentagon, and in senior coalition appointments in Baghdad and Kabul has become an essential complement to the deployment of land forces. These embedded officers build relationships and provide a variety of views within coalition headquarters. For a small army that can provide niche support to larger coalitions, the provision of high-quality staff officers in coalition headquarters is now business as usual and likely to remain a key requirement for a small army that seeks to influence the planning and conduct of wars.

---


16 This is not a new lesson. The works of David Galula (Counterinsurgency Warfare: Theory and Practice, Praeger Security International, 1964); Frank Kitson (Bunch of Five, Faber and Faber, 1977); and Andrew Krepinevich (The Army in Vietnam, 1988) are all examples of scholarship that have covered this issue prior to the current series of operational deployments.

The evolving character of war has seen an evolution in the kinds of missions armies undertake. The engagement with—and development of—indigenous forces has come to the fore in places as diverse as Timor Leste, Iraq, Afghanistan and the Solomon Islands. Population support and reconstruction operations, like the Australian experience in Vietnam, have again become something conducted concurrent with, instead of after, combat operations. This evolution in mission sets has been driven by the different types of threats and adversaries that present themselves. As a result armies have changed their structures and interactions, with both conventional and special forces adapting their organisational approaches. For conventional forces, battle grouping—building a force first around a mission—has become the norm. Units such as the engineer-led Reconstruction Task Force and the Mentoring Task Forces have been formed, certified and deployed since 2006. Special Forces have also learned about grouping different capabilities around their particular segments of the Afghanistan mission. The Special Forces Task Group has routinely combined Special Air Service, Commandos and other elements in their deployed task forces.

A closer integration of conventional and unconventional forces has also evolved over the past fifteen years. The Australian special and conventional units, co-located in Uruzgan province, developed a more closely synchronised approach to operations over several years of operations. This deepening integration of land conventional and special forces has been a defining feature of the broader coalition approach in Afghanistan. Reinforcing this in Army doctrine, and its individual and collective training regimes, will be essential to continue this collaboration in the future.

One of the enduring lessons reaffirmed in recent conflicts is that combined arms close combat remains a core capability of a professional land force.
An army proficient in close combat will possess the foundational skills (in broad terms these are to command, see, move, engage, and support) that can be adapted to changing conditions and applied to less taxing missions such as peacekeeping and low tempo stability missions. While this runs against the narrative of several prominent Australian commentators, it is not a new revelation. Tactically effective military organisations that employ a mix of capabilities (combined arms) to achieve realistic strategic objectives remain the foundation of military success in land operations.

A final element of the evolving character of land warfare—and warfare more generally—is that the past fifteen years has been an era of ambiguous outcomes. In East Timor, the Army largely departed in the early 2000s and then returned in force in 2006 to address continuing instability. Neither of the operations in Iraq or Afghanistan has resulted in clear strategic success, or victory as might have been understood in earlier conflicts. As Sir Hew Strachan has noted, “wars have become fuzzy at the edges: they have no clear end and army forces increasingly have to reject the appropriateness of classical definitions of military victory”.

The Australian Prime Minister noted in addressing troops during an October 2013 visit to Afghanistan, “Australia’s longest war is ending, not with victory, not with defeat, but with we hope an Afghanistan that’s better for our presence.” The following day, an editorial in the Australian Financial Review noted that “Afghanistan cannot be judged in conventional terms as win, loss or draw … but the country is no longer the failed state of 2001”. It is possible that land forces will be required to participate in operations in the future with similar ambiguous outcomes. Michael Howard describes this as warfighting and peacekeeping melting into one another, with the conduct of each determining the success of the other. The cognitive demands on leaders at all levels are extremely high in these types of situations. It

__References__

22 For example, see Brendan Nicholson, ‘Slim Down the Army and Think Again about New Subs and JSFs: Defence Analysts’, *The Australian*, 11 December 2013, p. 6.


requires mental agility and enhanced tolerance for ambiguity and chaos.\textsuperscript{28} ‘Train for certainty and educate for uncertainty’ is a common mantra in many professional military organisations.\textsuperscript{29} The Army’s training continuum and its approach to professional education (which must include the study of military history) must evolve, built on its sound existing foundation, to ensure it prepares its people appropriately for these circumstances.\textsuperscript{30}

\section*{Lesson 2: Sustaining an Expeditionary Army}

The 1976 Defence White Paper resulted in the Army being focused on continental defence. The Army would no longer be required to be sent overseas to fight as part of another nation’s force.\textsuperscript{31} By the time the Army commenced small-scale offshore operations—commencing with Namibia in 1989—it’s capacity to train, deploy and sustain land forces offshore had atrophied.\textsuperscript{32} As one former Chief of Army noted,

\begin{quote}

during the 1980s and for much of the 1990s, the strategic guidance given to the Army ultimately diminished land force capabilities. We gradually lost strategic agility; our units became hollow; and our ability to operate away from Australian support bases declined to a serious degree.\textsuperscript{33}
\end{quote}

But as Lieutenant General David Morrison also notes, while the Army did atrophy post-Vietnam, this was partially because the Army did not make the case to government to sustain its warfighting proficiency at unit and formation levels.\textsuperscript{34}

Since 1999 the Army has relearned the personnel, logistic and other difficulties of training and sustaining deployed land forces—in diverse


\textsuperscript{30} The study of military history within a professional military education continuum has been a core to the development of officers in many armies. See E. Keogh, ‘The Study of Military History’, \textit{Australian Army Journal}, vol. IV, no. 2 (1965); and Williamson Murray, \textit{The Past as Prologue: The Importance of History to the Military Profession} (Cambridge: Cambridge University Press, May 2006).


\textsuperscript{32} For more on this topic, among others, see Mick Evans, \textit{The Tyranny of Dissonance: Australia’s Strategic Culture and Way of War 1901-2005} (Canberra: Land Warfare Studies Centre, February 2005).


\textsuperscript{34} David Morrison, ‘The Australian Army for the Next Decades’, speech to the Lowy Institute, Sydney, 19 February 2014.
After Afghanistan: A Small Army and the Strategic Employment of Land Power

theatres—over a long period of time.\textsuperscript{35} It has been forced through necessity to re-examine logistic structures, planning processes and the training mechanisms to deploy and support land forces over long periods of time. This self-examination resulted in the 2008 Adaptive Army initiative that reorganised the Army. This reorganisation, the largest since the 1971 Hassett Report,\textsuperscript{36} aimed to ensure the Army was better at incorporating operational lessons into force preparation in the short term, better at assuring the performance of our forces through a very robust and transparent certification process and developing medium and longer term initiatives to enhance the Army’s combat weight and effectiveness through training, equipment and doctrine.\textsuperscript{37}

Recent operations have reinforced that effective expeditionary armies must be trained in a core set of military individual and collective skills, that can then be adapted before deployment by additional training and equipment relevant to the specific theatre of operations. The Army’s force structures have not always facilitated this—in particular the different structures of each manoeuvre Brigade have posed substantial force generation problems. While the Army’s current Plan Beersheba\textsuperscript{38} aims to address this challenge, in many respects it is likely to be the initial steps in moving from an analogue, twentieth-century organisation to one that is digitised (and all that means culturally and organisationally) and optimised for warfare in the information age.\textsuperscript{39}

Operations in Afghanistan and Iraq have demonstrated the need to rebalance land force capabilities towards many of the combat support functions—human intelligence, electronic warfare, and explosive ordnance disposal among others. These have become essential for the conduct of land operations and will likely be required in greater densities for future


operations. An important response should be a rebalancing of the combat and combat support (or enabling) functions of the land force. The Army that deployed to East Timor in 1999 was largely an infantry force, with limited integral intelligence, reconnaissance, surveillance and other combat support functions. The past fifteen years has demonstrated that a higher proportion of capabilities such as engineers, intelligence analysts, unmanned reconnaissance and logistics are needed.

The Army has also learned how to better employ a total force—its regular, reserve and civilian workforce—on expeditionary operations over the past decade. While this has involved large reserve deployments to East Timor and the Solomon Islands concurrent with the Afghanistan commitment, it has also involved the broad employment of reserves in specialist roles in deployed and force preparation functions. This represents the Army moving closer to achieving the aspirations for a ‘total force’ recommended in the 1974 Millar Report into the Army Reserve.40

A final and most crucial element of sustaining an expeditionary army is support to its soldiers. Since 1999, the Army has relearned the importance of supporting soldiers and their families better. The Army has launched ‘wounded warrior’ initiatives, including clinics and health centres, for the recovery of those with physical wounds. But the care of those with psychological wounds is just as vital. While the Army and wider Defence Department has instituted a range of programs for reporting and dealing with post-traumatic stress disorder, there remains room for improvement. Recent audits by the Australian National Audit Office41 and the Joint Standing Committee on Defence and Foreign Affairs42 have found weaknesses in the systems developed since 2001, particularly those for supporting members with psychological wounds. These require ongoing attention from the Army and the Department to redress.

There will be a temptation to reduce investment in this area as the Army draws down the number of soldiers deployed in Afghanistan. While this would align with a similar approach to that taken after Vietnam, it is a mistake. Not only will the Army continue to deploy soldiers in various exercises and deployments, there remains an ongoing need to support

---

40 The Millar Report investigated the efficiency of the then Citizens Military Forces (CMF), which after the report was renamed the Australian Army Reserve. Jeffrey Grey, The Australian Army (South Melbourne: Oxford University Press, 2001), p. 227.
soldiers and their families who suffer from the physical and psychological trauma of their service on operations. The Army had to relearn many of its lessons from Vietnam over the past decade. It should not have to relearn them again should the Army be called upon to undertake long-term commitments overseas in the future.

Lesson 3: Joint, Coalition, Networked, and Interagency Approaches

The Army has learned how to synchronise and collaborate more broadly over the past fifteen years. This more connected approach—with joint, coalition and interagency elements—has manifested a range of lessons in the conduct operations where networking, synchronisation and sharing information has become ‘business as usual’ and have been a key enabling aspects for operations where adversaries are also adept at exploiting information age global connectivity.

The deployment of Australian forces, under the Australia-led International Force for East Timor (INTERFET), represented a new era for the conduct of joint operations by the Australian Defence Force. The integration of land, air and maritime operations under a single operational headquarters represented a significant enhancement of the capabilities of the Australian Defence Force. The follow-on operations in Iraq and Afghanistan have served to strengthen the tight relationships between the Services, notwithstanding the different geographic circumstances involved.

Just as East Timor introduced the Army and broader Australian Defence Force to contemporary joint operations, it provided many experiences that primed the Army for learning about coalition operations in the more complex coalition environments of Iraq and Afghanistan. The former Defence Minister, Steven Smith, provided some insights into this in a speech in April 2013 where he noted that Australia’s operations had enhanced its relationship with the United States and NATO. Ryan has noted that the demands of the counter-terrorism and counterinsurgency operations since

---

43 Homer, ‘Deploying and Sustaining INTERFET in East Timor 1999’, p. 204-29.
2001 have provided further evidence that any contemporary multinational operation is likely to involve a wide variety of partners.46

This has been the case for the Army in operations in Iraq and Afghanistan. While a junior partner in both conflicts, Australia operated alongside the armies of multiple allies and coalition partners. In Iraq, Australia partnered with a Japanese land force while under the command of the British.47 In Afghanistan, Australia partnered with the Dutch while serving under Canadian, British and American commanders in southern Afghanistan.48 These coalition operations have highlighted some critical capabilities which underpin successful prosecution of operations with allies and other coalition partners. In a coalition, being able to provide intelligence, surveillance, reconnaissance, transport and logistic support to one’s forces, and to the broader coalition force, is valuable. Being able to shoulder risk and conduct combat missions is also a key contribution within a coalition force.49 Recent remarks by the President of the United States,50 and the 2014 Quadrennial Defense Review, highlight US expectations of increased burden sharing with allies in the future.51

Australia has led several regional security operations in the past fifteen years. Defence has learned much by doing so, including that as the lead nation, Australia becomes responsible for providing the framework of the coalition for such capabilities as communications and logistics. As experiences in East Timor and the Solomon Islands demonstrate, a key underpinning of Australia’s success has been the establishment of the coalition framework at both the strategic and operational level—the Regional Assistance Mission Solomon Islands architecture has been a useful model in this regard.

47 For more on these operations, see John Blaxland, *The Australian Army from Whitlam to Howard*, pp. 218-56.
48 In the later stages of the Afghanistan deployment, Australian forces were joined by small contributions from other nations such as Singapore.
49 For one useful study on this topic, see Dickens, ‘Can East Timor be a Blueprint for Burden Sharing?’, pp. 29-40.
50 During a speech at West Point in May 2014, President Obama noted that ‘when issues of global concern do not pose a direct threat to the United States, when such issues are at stake—when crises arise that stir our conscience or push the world in a more dangerous direction but do not directly threaten us—then the threshold for military action must be higher. In such circumstances, we should not go it alone. Instead, we must mobilize allies and partners to take collective action’. See The White House, ‘Remarks by the President at the United States Military Academy Commencement Ceremony’, U.S. Military Academy-West Point, New York, 28 May 2014, <http://www.whitehouse.gov/the-press-office/2014/05/28/remarks-president-west-point-academy-commencement-ceremony> [Accessed 16 September 2014].
51 The QDR noted that ‘with our allies and partners, we will make greater efforts to coordinate our planning to optimize their contributions to their own security and to our many combined activities’. US Department of Defense, *Quadrennial Defense Review*, 4 March 2014, p. vi.
Australian Defence Force operations over the last fifteen years have demanded close cooperation with other government agencies and a broad range of international and non-government organisations. Some missions have proven to be best led by these other agencies (for example, the Department of Foreign Affairs and Trade led the Regional Assistance Mission in Solomon Islands); while others are best led by the Australian Defence Force.

Interagency operations, while increasingly routine, are hardly seamless. While the Australian Defence Force is structured and prepared for overseas operations in violent and dangerous circumstances, most other government agencies are not. They are typically staffed with large numbers of people who are not available at short notice for long and arduous deployments. Nor do they typically have the theatre-level logistic, intelligence, and security support required to operate in an expeditionary mode.

Risk management has also been a difficult concept for different government departments to agree on, with no common method for determining risk within the non-military components. However, the establishment of the Australian Civil Military Centre in 2008 has provided a useful starting point for examining the lessons of interagency cooperation. While it is at the forefront domestically in developing better Defence, government and non-government agencies cooperation, there is more to be done within the Army and Defence to educate and train its personnel in this facet of operational capability.52

While many insights have been gained into human networking over the last fifteen years, there have also been major developments in technical connectivity and networking. The integration of technology with the political, economic, and military institutions of the state is accelerating at a tremendous rate.53 This has enabled a deepening of the capacity of the land force to synchronise and integrate its planning and operations within a networked, joint, coalition and interagency construct. By necessity, the Army had to improve its capacity to network its forces, within a joint force, during the East Timor operations in 1999. The Army commenced that operation as an industrial age organisation, with an immature and unsophisticated approach to networking. Through the mid-2000s, Defence followed the developed theory of Network Centric Warfare that emerged in the United

52 For additional information and a range of publications, see Australian Government, 'Australian Civil-Military Centre', <http://acmc.gov.au/> [Accessed 16 September 2014].
States. This led to the publication of joint doctrine in 2003, and a series of Network Centric Warfare Roadmaps.

Although the term Network Centric Warfare has fallen into disuse, the broader examination of the benefits of networking a military organisation has continued. In particular, the linkages between sensors, decision-makers and weapon systems have continued to become more tightly coupled. The broad availability of multiple layers of intelligence, surveillance and reconnaissance feeds, linked to commanders and thereon to weapon systems, has been a defining feature of Iraq and Afghanistan. The increasing availability of cheap, commercially available airborne and ground based surveillance capabilities and the enhancement this provides to the precision of engagement systems is likely to be a key element in future operating environments.

But in the future the Army in a joint force might not be able to guarantee the dominance in these capabilities as they have in the past fifteen years. Contemporary adversaries have seen the value in asymmetric approaches which use simple, inexpensive techniques. With the proliferation of cheap sensors, unmanned vehicles and electronic systems that can interfere with friendly communications, future land forces must also be prepared to operate where an adversary might deploy these almost ubiquitous capabilities to generate degraded information environment.

It is also probable that future adversaries, particularly non-state actors, will continue to exploit global connectivity in the pursuit of their tactical and strategic objectives. Over the past decade, a range of different non-state actors have used the internet and social media to share lessons, conduct information operations, raise funds and support radicalisation.

---

55 The first of these was endorsed by the Chiefs of Service Committee in November 2003. Subsequent editions were released in 2005, 2007 and 2009.
57 The degraded information environment, and the conduct of degraded operations, has been the topic of much investigation over the past five years particularly in the United States (for the impact on its systems) and in Russia and China (for how they might generate this degraded effect on others). See US Department of Defence, *Enhancing the Adaptability of US Military Forces*, Defense Science Board 2010 Summer Study, January 2011, pp. 75-121. On China’s capacity to degrade friendly systems in informationised warfare, see Office of the Secretary of Defense, *Annual Report to Congress, Military and Security Developments Involving the People’s Republic of China 2014*, 24 April 2014.
Whether working in a large, international coalition led by the United States or leading a smaller regional commitment, *joint, coalition and interagency* has been the experience of the Army for the past decade. The more tightly connected approaches of the past fifteen years of friendly and adversary organisations—in the human and technical areas—provide insights for training, education and force design for the Army and wider Defence organisation.

**Lesson 4: Strategic Warning, Mobilisation and the Readiness of Land Forces**

Former Prime Minister John Howard recently noted that East Timor “put an enormous strain on us and it brought home to me just how much of the post-Vietnam and post-Cold War peace dividend we had taken. Our military capabilities had been run down and we needed to do something about it”.  

Throughout the 1970s, 1980s and into the 1990s, an underlying narrative existed in the Defence establishment that the Army could be maintained in its traditional role of foundation for an expanded land force should the need arise.

The result of this was a hollow, barely capable force that was just able to cope with the deployment to East Timor in 1999. In a government inquiry in the wake of East Timor, it was found that the funding of the Army reflected a lack of appreciation of the funds required to generate and maintain ground combat capability in a useable state; and a failure to determine minimum acceptable levels of capability and then provide consistent resourcing to meet these levels. It possessed a force structure ‘hollowness’ that had been a persistent feature of the Army organisation, which consumed resources while not delivering capability in meaningful time frames. This situation was the outcome of two strategic concepts.

---


61 Ibid., p. 105-6.
The first is that Australia would have the necessary strategic warning to undertake the required expansion of the land force. As the eminent author Colin Gray has reminded us, the peril of strategic surprise is a condition of international and national security.\(^{62}\) Strategic shocks occur—and the past fifteen years in particular have routinely demanded the deployment of land forces on much shorter deployment timelines than anticipated by intelligence agencies. Examples such as the 1999 East Timor deployment,\(^{63}\) commitments to Iraq and Afghanistan, the return to East Timor in 2006, as well as several large-scale disasters testify to this. Australian land power theorist Mike Evans perhaps described this situation best when he stated that:

between 1999 and 2003 Australia entered the new age of globalised security, in which it became apparent that there was no longer any such phenomenon as convenient warning time or preparation for a protracted mobilisation of manpower and economic resources.\(^{64}\)

The second concept posits that the Army can be reduced to small cores for rapid regeneration in an emergency. But the mobilisation model adhered to for the First and Second World Wars is not relevant to the twenty-first century for several reasons. Most notably the character of warfare has changed significantly. It has resulted in the contemporary Combat Brigade, an information age organisation comprising more than 3,000 troops and hundreds of vehicles all linked in a digital network, which is arguably more complex than an Air Warfare Destroyer or Joint Strike Fighter. This is far removed from giving citizens a rifle, a tin hat and minimal training, as occurred in the early twentieth century. Additionally, with the exception of a major conflict, the globalised nature of the Australian economy is unlikely to support rapid industrialised mobilisation, either economically or socio-culturally.

The lesson from this is that the twentieth-century industrial age approach to sustaining a small army as the basis for expansion no longer holds. An army, especially a small one, must be prepared when called upon to deploy within short periods of time. It must be an organisation that is prepared to ‘deploy as is’. The first three lessons in this article provide sufficient insights into potential future conflict, and the type of land forces required, to challenge old concepts of mobilisation and highlight the likelihood of surprise. The Army must therefore retain sufficient breadth of capabilities,


\(^{63}\) The 2000 Defence White Paper described in some detail the key lessons of the East Timor deployment, which included the acknowledgement that warning times for regional crises, and the need to respond, were shorter than Defence planning had generally allowed for. Commonwealth of Australia, *Defence 2000: Our Future Defence Force* (Canberra: Department of Defence, 2000), p. 49-50.

\(^{64}\) Evans, *The Tyranny of Dissonance*, p. 95.
combat weight and mass to be able to respond quickly to government requirements for unanticipated needs.  

Lesson 5: The Strategic Relevance of Land Forces

Recent debate has questioned the relevance of sustaining the current size and capabilities of the Australian Army after its Afghanistan commitment draws down. Hugh White has argued that in future, the biggest task “may well be the defence of the continent itself”. An alternative view, expressed by Alan Dupont, is that the Army is likely to continue to bear the brunt of future deployments. This debate on the future strategic relevance of land forces is an important one to have. But it is one that the Army must also contribute to.

Stephen Metz recently noted that contemporary armies have been so heavily engaged on operations over the past fifteen years that they have possessed limited time for reflection, strategic thinking, and writing. However, this is changing. Foreign military organisations have commenced reviewing the lessons of the past fifteen years, with the US Joint and Coalition Operational Analysis publication A Decade of War being one exemplar of this process. The ongoing collaborative effort between the United States’ military instruments of land power in reviewing the future of its application is another manifestation of this desire to learn from contemporary operations. The 2013 Strategic Land Power paper, developed jointly by the US Army, US Marine Corps and US Special Operations Command, describes the aspirations for the employment of land forces over the coming decades.

The term ‘land power’ reflects the dynamism of the strategic environment over the past fifteen years. Land power encompasses the employment of an array of land capability to achieve specified strategic and tactical objectives. It is a ‘multi-dimensional’ approach: land power may include the employment

---

65 An effective army is one that retains sufficient breadth of capabilities, combat weight and mass to be relevant to the achievement of national strategic objectives in the coming decades. To do so requires ongoing and iterative modernisation. This must be founded upon better knowledge about trends in the environment in which land forces will likely be employed and demands a sound understanding of the convergence of developments in conflict, society and technology. Exploiting this understanding allows an army to remain relevant for future contingencies through modernisation, acquisition of new capabilities or disinvestment in obsolete systems.


70 US Department of Defence, Strategic Landpower: Winning the Clash of Wills, 2013, p. 3.
of capabilities from multiple operational environments (land, sea, air, space, the electromagnetic spectrum and cyberspace) to achieve results on land. While the growing literature on the strategic role of land power is mainly generated from the United States, it is relevant as well to the future employment of the Australian Army.

The differences in scale and outlook between the US and Australian armies have not precluded the Army thinking and writing about its role in achieving Australia’s national security objectives. In particular, the landmark 1998 edition of the Australian Army’s *The Fundamentals of Land Warfare* examined the strategic effectiveness of land forces generally, and in the Australian context. It noted that armies generate strategic effect through four key functions: maintenance of a war fighting capability (although the definition of this was generic); sustaining a force in being at certain preparedness levels, shaping operations (through engagement activities and exercises); and, military support operations.  

In 2002, a subsequent edition of this publication defined the Australian conception of land power. It described the strategic application of land power in its description of the Army as an expeditionary organisation, focused on securing national sovereignty and interests, and a force, the commitment of which provides the ultimate demonstration of national commitment. A later 2008 version eschewed the term ‘land power’ and substituted a concept called ‘strategic manoeuvre’. Notwithstanding some of the weaknesses in this definition, it is indicative of an Army that aspires to operate beyond the realm of tactics and be capable of generating strategic effects in support of national objectives. The past fifteen years have provided several examples where this is the case. The commitment of strategically deployable land forces into East Timor resulted in tactical security being achieved in Dili and its surrounds, through the rapid build-up of land power. It was also a demonstration of political and national will, and employed land forces to achieve the strategic outcomes of Australia in concert with the international outcomes adopted by the United Nations Security Council in 1999. Similar commitments to Afghanistan in 2001, Iraq in 2003 and Afghanistan in 2005 demonstrated Australia’s adroit use of land forces to achieve strategic outcomes.

---

The 2005 re-commitment of land forces into Afghanistan is another example of this approach. Initially a small, tailored Special Operations deployment in 2005, the commitment was significantly expanded in 2006 into the larger Reconstruction Task Force. While the commitment to securing Afghanistan from international terrorist organisations was a part of the rationale for this, the deployment of Australian forces to Uruzgan province also played a role in the strategic aim to secure the commitment of a non-ABCA (American, British, Canadian, Australian and New Zealand) NATO partner (The Netherlands) into southern Afghanistan.

The operations in East Timor, Solomon Islands, Iraq and Afghanistan have demonstrated that armies operate beyond the exclusive domain of tactics. As Metz notes, land power is particularly important in the human domain, largely because it puts forces in direct contact with those they seek to influence; whether by deterring enemies or convincing them to stop what they are doing, or by convincing civilian policymakers and populations that they share objectives and priorities. In the contemporary security environment, strategic success requires an ability to understand, influence, or control the human domain.\(^{75}\)

It is in the human domain where the capability of land forces—large or small—converge. The capacity of a land force of any size to understand, influence, or control the human environment in which they operate provides a unique, strategic capability. Notwithstanding the potential need for sea and air forces as well as cyber and space operations, the capacity to execute operations on land is an important instrument of national policy for any sovereign nation, regardless of size. In doing so, these land forces provide their national governments with the capacity to generate tactical and strategic impacts.\(^{76}\)

**Conclusion: Preparing for the Next War**

While the Army’s current operational commitments are on the wane, the long-term strategic uncertainty of maintaining stability in the Indo-Pacific and regions such as the Middle East remains high. As Michael Wesley has recently noted, our strategic environment is changing far more quickly than Australians appreciate, with complexities and challenges greater than any the nation has had to deal with in its history.\(^{77}\) The Australian military may therefore face a range of daunting challenges in its future. As the draw down in Afghanistan takes place, it is time to turn more fully to preparations for the next conflict. A smaller institutional commitment to operations provides

---

\(^{75}\) Metz, *Strategic Landpower Task Force Research Report*.

\(^{76}\) This theme is examined in the 2014 re-write of the Army's capstone *Land Warfare Doctrine 1*, which has been re-written to incorporate lessons from the past fifteen years. It will be published in late 2014.

increased intellectual opportunity to learn from those operations and produce innovative technical, organisational and training solutions for future conflict.

Recent studies have also highlighted the perils of ignoring or paying lip service to this post-conflict endeavour. In his 1986 study of institutional learning in the post-Vietnam era, Andrew Krepinevich proposed that in spite of its anguish in Vietnam, the US Army learned little of value.78 The collective institutional avoidance of counterinsurgency between Vietnam and Iraq was to have a profound impact on recognising and responding to an unconventional threat after the initial operation in 2003 that resulted in the US seizure of Baghdad. Small armies, with their limited manpower and resources, cannot afford such an approach.

But learning lessons from past operations is only part of the pathway to success in future conflict. As Cohen and Gooch have proposed, the flipside is that failing to learn the lessons of the recent past is but one of several pathways to military failure.79 As Rosen finds, preparing for the future demands the capacity to take lessons and apply them within a strategic approach to innovation, led by senior military leaders, that has both intellectual and organisational components.80 It demands that the Army has a vision of the future, and its place within that future be connected to operational realities drawn from the lessons of recent operations.81 The Army post-Afghanistan must also be one that is highly energetic in the intellectual investment in preparing for future conflict.

All military organisations to some degree get the next conflict wrong.82 It is simply not realistic to expect the Army to "get it right" from the outset of all complex missions in the future. But, it must be able to improve performance over time, and do that more quickly than an adversary.83 As a component of a small military, the Australian Army also has less capacity than larger nations to absorb the consequences of getting it wrong. A robust approach to learning the institutional lessons of recent operations will assist in the Army husbanding and applying its scarce resources in order to minimise the chances of this occurring.

Mick Ryan is an Australian Army officer who is a graduate of the Royal Military College Duntroon. He has served in a range of command and staff appointments, and has operational experience in East Timor, Iraq and Afghanistan. He is a distinguished graduate of the Marine Corps Command and Staff College, and a distinguished graduate from the Johns Hopkins School of Advanced International Studies. He assumed the command of the 1st Brigade, one of Army’s three combat brigades, on 13 October 2014. email address
Old but Gold: The Continued Relevance of Naval Gunfire Support for the Royal Australian Navy

Steven Paget

Despite an extensive history of providing naval gunfire support (NGS) during the twentieth century, the relevance of the capability to the Royal Australian Navy (RAN) came under scrutiny at the turn of the twenty-first century. However, the provision of NGS during the 2003 Iraq War indicated that the capability remained useful in the modern era. Furthermore, lessons learned during past RAN operations and those of the Royal Navy and United States Navy demonstrate the enduring relevance of NGS. Australia’s development of an amphibious capability will ensure that NGS remains useful in the future for a number of important reasons, including utility, cost-effectiveness and interoperability.

In the Autumn 2012 edition of Security Challenges, Peter Dean pertinently observed that, “the history of the Australian military can provide both relevant historical examples for future directions as well as a critically important cultural heritage platform in amphibious operations”. Australia’s intention to develop an amphibious capability has raised multiple issues and sparked a number of debates, including the requirement for such a force, its necessary size and the range of missions that it will need to fulfil. Australia’s amphibious assets will need to be deployed for an assortment of essential missions, including humanitarian purposes. Whilst the Australian Defence Force (ADF) will need to be configured to conduct a number of different missions, power projection has continually proven to be an important task for the Royal Australian Navy (RAN). In order to project power, the capacity of the RAN to deliver firepower ashore is an essential issue.

If Australia is serious about developing an extensive amphibious capability, lessons can be learned not only from past Australian operations, but also from those of her closest allies. The potential importance of naval gunfire support (NGS) was borne out during RAN operations alongside allied navies.

---

1 The views expressed in this article are entirely those of the author and do not necessarily reflect those of the New Zealand Defence Force.
during the Korean War, the Vietnam War and the Iraq War. In addition, the United Kingdom’s involvement in the Falklands War and the American intervention in Grenada demonstrated that NGS remains an important element of amphibious operations. Time and time again, the utility of NGS has been demonstrated, but it is a lesson that has had to be continually re-learned.

The development of precision guided munitions and sophisticated aircraft capable of providing accurate air support created the perception that NGS would be confined to being an issue of historical interest rather than practical relevance. However, conflicts throughout the twentieth century, and particularly those in the late-twentieth and early twenty-first centuries, have indicated that NGS can make a useful contribution to amphibious operations. At the very least, NGS can be used to supplement other forms of firepower. More importantly, NGS can be used to replace other forms of fire support and in some instances, may prove to be the only available asset. Land-based artillery, air support, and precision guided munitions have all proved to be unobtainable on occasion during operations, which can lead to a vast increase in the significance of the availability of NGS. Under certain circumstances, the nature of NGS and its capacity to spread fire can make it the most appropriate option to further the aims of an amphibious force. Aside from the practical applications, NGS has the appeal of being cost-effective, as rounds are significantly cheaper than precision guided munitions. All of the aforementioned factors have served to ensure that NGS should remain an important consideration in the conduct of amphibious operations. As the ADF’s primary operating environment is maritime and archipelagic in nature, it is especially conducive to the use of NGS under the right conditions.

The value of NGS has, once again, been confirmed during the current decade. Throughout Operation Ellamy—the United Kingdom’s designation for the 2011 intervention in Libya—personnel from 148 (Meiktila) Commando Forward Observation Battery, Royal Artillery, a dedicated spotting unit, provided observation for over 500 rounds of NGS from British and French ships across the course of fifty missions. The requirement for NGS was outlined in a North Atlantic Treaty Organisation (NATO) fragmentation order, which noted that sustained NGS should be employed as a “show of force”. Scott Bishop has contended that the requirement to minimise the risk to civilians and infrastructure “relegated” surface warships to “very limited roles in strikes against targets ashore” due to the absence of precision strike capability. Nevertheless, he has also acknowledged that as the conflict progressed, there was increased opportunity to provide NGS. Indeed, air strikes and NGS throughout Libya, in places such as Brega, Sirte,

---

Tawurgha, Tripoli and Zlitan, helped to prevent forces operating in support of Colonel Gaddafi from utilising their “combat assets”. One of the most notable employments of NGS occurred off Zlitan in August, when HMS *Sutherland*, HMS *Liverpool* and the French ship *La Vallee* rotated on the gunline in order to disrupt pro-Gaddafi forces. In one engagement, HMS *Liverpool* destroyed a pro-Gaddafi convoy with fifty-five rounds. Star shells were also used on a number of occasions, including at Misrata in June, in order to aid rebel forces and provide a demonstration of force, whilst at the same time preclude the risk of collateral damage. The versatility of the naval gun that resulted from the range of shells available meant that NGS was a useful tool under varying circumstances.

The mobility of naval forces was seen as being central to operations in Libya. Naval power provided a “boots off the ground” option, which allowed NATO forces to influence the land battle whilst maintaining a limited commitment. Prohibitive rules of engagement and the occurrence of fighting in residential and built-up areas ensured that precision guided munitions were favoured over NGS, but not exclusively. In reflecting on the campaign, Lieutenant Colonel Tim Wood, an NGS support officer (LNO), concluded: “NGS demonstrated yet again its versatility in the littoral”. The fact that Libyan operations “reinforced British interest in the development of more sophisticated means of naval gunfire support” was testament to the ongoing utility of the capability.

Despite some commentators indicating that the use of NGS in Iraq in 2003 was an anomaly, the limited use of the capability had not signified its impending extinction, but instead represented a transformation in the manner in which it was employed. The availability of other forms of fire support has ensured that NGS will likely be used more sparingly in the future, but it still has an essential part to play. For instance, in Iraq, NGS compensated for the temporary absence of tactical air support and the lack of artillery, whilst in Libya, the appearance of warships provided a dramatic demonstration of force. The lessons of history have shown that when the capability has been written off, NGS has proven to be essential.

### The RAN’s Experience of NGS since the Second World War

The reputation of NGS has waxed and waned since 1945, and, ultimately, the capability has never obtained the level of reverence that was achieved.

---


during the Second World War. Nevertheless, in the post-Second World War period, the RAN was called on to deliver NGS during the Korean, Vietnam and Iraq (2003) wars, as well as during the Malayan Emergency and Indonesian Confrontation, albeit on a much more limited basis. During the course of the Korean War, between 1950 and 1953, the RAN made a significant contribution to the provision of NGS by the United Nations (UN) coalition. RAN ships steamed over 674,000 kilometres and fired: 9,614 4.7", 5,576 4.5", 9,983 4", 50,417 40mm and 9,921 2-pounder rounds.\(^{11}\) As naval historian Eric Grove has aptly surmised, Korea was “a major British naval war”, but it is equally true that it was a major Australian naval war, particularly when the relative resources that were available to each navy are taken into account.\(^{12}\)

NGS was also utilised by the RAN in the 1950s, but only on a much more limited scale. During the Malayan Emergency, which was declared in response to Communist terrorist actions, and continued from 1948 and 1960, the RAN provided a small commitment to the Commonwealth Strategic Reserve from 1955 onwards. The nature of the Malayan Emergency ensured that the use of NGS was constrained. In January 1956, the Director of Operations, Malaya, recorded: “Bombardment of the terrorists by ships is a most helpful form of support for ground operations, but unfortunately there are not many opportunities for this type of attack.”\(^{13}\) In fact, RAN ships provided NGS on just four occasions.\(^{14}\) Although the performance of the RAN ships was commendable, it is difficult to contend with the opinion of the official historians that the results “demonstrated that naval power was inappropriate to the circumstances of the Malayan Emergency”.\(^{15}\)

The results of the Indonesian Confrontation (Konfrontasi) replicated the evidence garnered from the Malayan Emergency. Whilst NGS was used in


the harassment role and naval guns, including Bofors, were used as a deterrent to infiltration, assessments of bombardments were inconclusive, despite some indications of limited success.\(^{16}\) Therefore, whilst the manner in which NGS was used was not always justified, it retained some merit under the right circumstances.

It was not long before the significance of NGS was tested once again. As was the case in Korea, the United States Navy (USN) was at the forefront of the provision of fire support during the Vietnam War. The RAN’s contribution was numerically smaller, both in terms of ships and rounds fired, but the Australians were able to commendably support their American allies. The RAN deployed a single destroyer on a rotational basis between March 1967 and September 1971. Despite the limited number of RAN ships deployed to Vietnam, Australian destroyers made a useful contribution to the war effort by patrolling off the coast of North Vietnam, participating in interdiction missions, delivering NGS, occasionally taking part in in-country river operations and sporadically forming part of the aircraft carrier screen at Yankee Station.\(^{17}\)

Whilst the RAN favoured deploying the Charles F. Adams class guided missile destroyers (DDGs) HMA Ships Brisbane, Hobart and Perth, HMAS Vendetta, a Daring class destroyer, did undertake one tour. As well as being in service with the USN, the RAN’s DDGs were fitted with the latest technology and possessed “very good … AAW [anti-aircraft warfare] and surface gunnery capabilities”, which ensured that they were well suited to NGS operations in Vietnam.\(^{18}\) Although HMAS Vendetta was an older generation of ship, the capability of her six 4.5” guns to fire up to one hundred rounds per minute in good conditions made her an extremely useful asset when NGS was required.\(^{19}\) In total, during the RAN commitment to Vietnam, the destroyers fired approximately 80,000 5”/54-calibre and 6,800 4.5” rounds at a daily average of 115.\(^{20}\)

In the wake of the Vietnam War, the importance of NGS appeared to be in continual decline. Indeed, the initial design of the Anzac class frigate was


\(^{17}\) ‘Joint Royal Australian Navy–USN Activities, 1911-1971’, Undated, Naval History and Heritage Command, Contemporary History Branch.

\(^{18}\) Department of Defence, ‘Minute by the Joint Planning Committee at a Meeting on 10th January, 1968’, Undated, Australian War Memorial, AWM 122, 68/6001 Part 1; Ian Pfenningwerth, Missing Pieces: The Intelligence Jigsaw and RAN Operations from 1939-71, Papers in Australian Maritime Naval Affairs No. 25 (Canberra: Sea Power Centre-Australia, August 2008), pp. 244-5.


notable for its absence of a 4.5”/5” gun, with the 76mm Otto Malero gun favoured instead. Although the ship was eventually fitted with a 5” gun, the initial decision to forego the traditional naval gun was a barometer of the thinking in regard to NGS in Australia.²¹ Events during the ensuing decades indicated that the final decision was fortuitous at the very least, and, arguably, judicious.

Although the RAN had no cause to provide NGS during the 1980s and the 1990s, unlike her allies who found use for it in places as diverse as the Falkland Islands (Royal Navy (RN)), Grenada and Lebanon (both USN), the capability was unexpectedly required during the Iraq War in 2003. As part of the British drive to capture Basra, the Royal Marines launched an amphibious assault on the Al Faw peninsula, which commenced in the early hours of 20 March 2003, and involved the landing of 3 Commando Brigade, Royal Marines (as well as the 15th Marine Expeditionary Unit (US) that was under their tactical control) by helicopter. NGS, which was provided by HMAS Anzac and HM Ships Chatham, Marlborough and Richmond, proved to be an important factor in the assault.²² All four ships delivered fire support against defensive positions and fixed installations up to eighteen kilometres inland, but from an Australian perspective, it is notable that Anzac spent the longest period on the gunline and fired the most missions.²³ In total, the ships fired 155 rounds across the course of seventeen missions, with Anzac delivering forty-six 5” rounds.²⁴ NGS was not decisive, but it was a critical enabler for the Royal Marines, and contributed to the success of the assault, which was accomplished without any combat casualties being incurred.²⁵ Ultimately, the Iraq War provided the RAN with confirmation of a lesson that the RN learned during the Falklands War and that the USN discovered throughout operations in Grenada—NGS had sustained relevance in the modern era.

²¹ Professor Hugh White, written interview with author, 15 October 2013.
The Falklands and Grenada Experience

In the United Kingdom, by the end of the 1970s, the death knell for NGS had seemingly been sounded. It was believed that missiles would be more effective than the naval gun for both air defence and surface action and, therefore, when combined with the availability of close air support (CAS), NGS would be redundant in the future. Consequently, the RN began to develop ships without a 4.5" gun, such as the first and second batches of the Type 22 frigate. An interrelated decision resulted in a proposal to eliminate 148 Battery, the dedicated spotting unit, as part of a significant cull of amphibious forces following the 1981 UK defence review.

Thinking about amphibious operations was reversed by Argentina’s invasion of the Falkland Islands on 2 April 1982, which prompted the deployment of a British task force to recover the islands. The Falklands War, officially designated as Operation Corporate, led to British forces being deployed to reclaim the islands some 13,000 kilometres away. NGS was required for a range of reasons during Operation Corporate, including to support the brigade attacks and raids, for the purpose of harassment and interdiction, and to deceive and distract the enemy. The long-distance nature of Operation Corporate ensured that NGS was always going to be a valuable capability, but in the absence of medium and heavy artillery, it proved to be invaluable. The official review of Operation Corporate concluded that: “The infantry would not have been able to carry their objectives without the support they received from artillery and naval bombardment”. The 8,000 4.5" rounds fired by RN ships effectively provided a substitute for the absence of medium artillery.

NGS was a significant contributor to the success of British forces during the Falklands War. Captain Chris Brown, a naval gunfire forward observer (NGFO) during the Falklands War, remarked: “I think the conflict saved 148 [Battery]. It saved NGS.” Tellingly, the design of future RN ships was reversed in the wake of the Falklands War. The third batch of Type 22 frigates was fitted with a 4.5" gun, as was the Type 23, in direct contrast to

---


27 148 Battery had even received a farewell visit from the director of the Royal Artillery prior to the Falklands War.

28 Although the RN deployed two aircraft carriers to the South Atlantic (HMS Hermes and HMS Invincible), air support to land operations was only one of many roles required of the aircraft. Ultimately, the defence of the Task Force was inevitably a priority. Royal Air Force (RAF) reinforcements were committed to support ground operations, but were also required for the same roles as their RN counterparts. Whilst RAF Vulcan bombers conducted five ‘Black Buck’ bombing missions, they were not used in a CAS role. In addition, the expeditionary nature of the operation and the competing demands on transportation, limited the amount of artillery available in theatre.


30 Lieutenant General Chris Brown, telephone interview with author, 31 August 2012.
the first two batches of the Type 22. ³¹ This was despite the fact that 148 Battery had been informed prior to Operation Corporate that new generations of RN ships would not incorporate a 4.5” gun. In the end, the level of NGS provided was sufficient and the performance of the ships was highly commendable, although both could have been enhanced through a greater focus on the capability in the decades preceding the Falklands War.

The British were not the only ally of Australia to be reminded of the requirement for NGS during the 1980s. By the time the United States commenced Operation Urgent Fury—the invasion of Grenada—on 23 October 1983, NGS had become so outmoded that even establishing radio contact between ships and troops ashore for the purpose of obtaining fire support proved to be difficult. The crux of the problem was the procurement and utilisation of incompatible communications equipment. The Vinson secure radio equipment that was in use with the US Army was incompatible with the systems operated by the USN. ³² As a result of the rapid deployment of forces for the operation, there was insufficient time to ensure that units had compatible equipment, which meant that it was only after the 82nd Airborne arrived in Grenada that their officer-in-command, Major General Edward Trobaugh, discovered the problem. ³³

The stymied communications were a direct result of stovepiping. Incompatible equipment was not just a source of frustration, it was a hindrance to combat operations and led to dangerous situations. Most notably, in the process of reconnoitring and subsequently assaulting the runway at Port Salines, US Army Rangers were faced with a Cuban fortified machine gun position on high ground, which required indirect fire support to be attacked. The Rangers had their own mortars and hoped to call on AC-130 Spectre helicopter gunships. Communications with the AC-130s on the air-to-ground radio was problematic at first, but the Rangers were eventually able to summon their assistance. Nevertheless, as it was broad daylight and there was intense anti-aircraft fire, it was difficult for the gunships to provide effective fire support.

Captain John Abizaid, in command of A Company, 1st/75th Rangers, acknowledged that the Rangers also encountered problems in obtaining NGS:

Normally we have ANGLICO [Air Naval Gunfire Liaison Company] and there were plans for ANGLICO but because of the speed of the deployment, the ANGLICO never joined us. So here’s this perfect platform [naval 5” guns] sitting out on the coast unable to be used by us because the initial planning

---

³³ Major General Edward Trobaugh, written interview with author, 24 September 2012.
didn’t envision the need for it and because the right people didn’t get to the right place in the pre-deployment phase.34

Without the essential communications bridge that ANGLICO provided, the Rangers were unable to overcome the equipment compatibility issues that hindered their operations, which left them unable to establish contact with the ships for the provision of fire support. As a result, the Rangers were forced to launch an unsupported infantry assault up the hill, using a “commandeered” bulldozer in place of a tank.35 The attack on Port Salines was ultimately successful, but it resulted in a number of casualties.

The 82nd Airborne were limited in obtaining the support of helicopter gunships due to difficulties with the runway at Port Salines and, just like the Rangers, they had no direct means of calling for NGS. In the absence of direct contact with the Independence Battle Group, the paratroopers were required to relay fire support requests back to Fort Bragg in the United States, who, in turn, would provide directions to the ship.36 It was an incredibly inefficient process and one which was so convoluted that the 82nd Airborne were able to make little use of it. Without sufficient fire support, the 82nd Airborne’s progress was hindered and their advance inevitably slowed. Deficiencies in communications inhibited all aspects of Operation Urgent Fury, but no issue was more adversely affected than the provision of fire support. The value of NGS had been underestimated and the impact of that attitude trickled down and negatively affected US troops on the ground.

The lessons provided by British operations in the Falklands and American operations in Grenada should be heeded by the ADF. Both of the conflicts were amphibious and expeditionary in nature, which is a contingency for which Australia must be prepared. If Australia is serious about developing a comprehensive capability, the RAN, and for that matter, all service branches, can ill-afford to discount the importance of NGS, as demonstrated by the experience of her allies. Although the RAN was ultimately not required to provide NGS between the Vietnam War and the 2003 Iraq War, the potential for such an eventuality did exist.

Whilst great caution should be applied to the consideration of counterfactuals, Australia’s contribution to the International Force in East Timor (INTERFET), which was established by a United Nations Security Council resolution, provides a pertinent example. INTERFET forces were initially deployed in September 1999, in order to restore peace and security

34 ANGLICO is the dedicated observation force of the United States Marine Corps. General John Abizaid, interview with author, 23 June 2012.
35 The Rangers were also unable to utilise Marine helicopters because their maps contained different grid systems. Richard Connaughton, A Brief History of Modern Warfare: The Changing Face of Conflict, From the Falklands to Afghanistan (London: Robinson, 2008), pp. 105-6.
in the face of escalating violence following the independence vote. From an Australian perspective, it was fortunate that the 2nd Battalion Royal Australian Regiment’s (2RAR) entry into Dili was not opposed, as they were operating at the end of a long air bridge and without heavy equipment. As a result, 2RAR would have been reliant on close air support from Royal Australian Air Force (RAAF) Base Tindal or NGS. It was fortuitous that neither CAS nor NGS was required, but the potential for power projection was important. Then-Captain James Goldrick opined that, “maritime forces created a protective umbrella, within which the land component could operate, confident that it could concentrate on the job to be done in East Timor itself, without the possibility of external interference”.

NGS: There When You Need It

A range of arguments have been put forward to justify the demise of NGS, but the capabilities that were believed to have superseded NGS have not always been reliable, and in some instances, have not been available at all. The benefits of CAS and, more broadly, tactical air support, are obvious and those capabilities have often been essential, but they have not always been available to those conducting operations on the ground. Operations in the Falklands, Grenada and Iraq all provided a stern reminder that the availability of air support could be constrained by a number of factors, including: the absence of air superiority, the number of aircraft in theatre, a lack of coordination and weather.

CAS was a limited commodity during Operation Corporate due to the dearth of available aircraft. The limited number of Sea Harriers in theatre—and the world for that matter—and their primarily defensive role in protecting the fleet meant that their employment in offensive action was constrained. Following an attack on the airstrip at Goose Green by three Sea Harriers from HMS Hermes on 25 May 1982, resulting in the loss of one aircraft and the death of Lieutenant Nick Taylor, Rear Admiral Sandy Woodward, Commander of the Carrier Battle Task Group, made the conscious decision to preserve the aircraft. Woodward concluded: “I resolved that from now on, if they were to bomb at all, it would only be against specific high-value targets or from high level—less accurate, I knew, but less expensive in Sea Harriers.” Although the Sea Harriers and the Royal Air Force Harrier GR3s made a significant contribution to the land battles through the provision of CAS, the lack of air superiority and inclement weather limited the capacity of these extremely capable platforms.

The Official History of the Falklands Campaign recorded that: “[Brigadier Julian] Thompson found air support too weather dependent” and “so sparse as to be a bonus and nothing more”. Concerns over the potential availability of air support prompted Major General Jeremy Moore, who had assumed command of the land forces on 12 May 1982, to plan for the eventuality of it not arriving during the brigade attacks, which led to a focus on ensuring that there was sufficient NGS to provide the required support.

During Operation Urgent Fury, AC-130 gunships provided invaluable support and the availability of more aircraft would undoubtedly have been beneficial. Although operational fires were provided by a combination of A-7 fighter-bombers, AC-130 Spectre gunships and naval destroyers, it was the AC-130s that provided the distinguished support. Major General Trobaugh reportedly informed his staff that: “He would give up his offshore naval gunfire support, his landbased artillery, and his helicopters before he would release the gunships for redeployment”. Nevertheless, the involvement of the AC-130s was not unrestricted and they were not always capable of providing the requisite support, as was the case at Port Salines.

During Operation Iraqi Freedom (OIF), some elements of the coalition perceived that the USN viewed NGS as “outdated” because of the increasing propensity to rely on tactical air support. The nature of OIF and the overwhelming force advantage held by the coalition fostered the notion that CAS would make NGS redundant. The coalition had achieved air superiority during Operations Northern Watch and Southern Watch, maintained an overwhelming air power advantage over Iraq and had five carrier battle groups on station. Consequently, it was believed that aircraft using precision guided munitions could satisfy any fire support requirements during the assault on the Al Faw peninsula.

The weather and its potential knock-on effect on operations changed the equation and led to an increased need for NGS. The coalition was aware that there are periods when there is extremely poor visibility over Iraq at that time of year, which could preclude the availability of air support. As it turned out, a colossal sandstorm grounded a large proportion of the aircraft, which meant that the ships ended up providing fire support for forty-eight rather than the planned twenty-four hours. In the end, none of the sorties that were put in the air to provide preparatory fires on the landing sites made

it to the targets, because they were all re-directed while in the air to high-priority targets that had been called in by the United States Marine Corps (USMC). While the system of allocating air support on a priority basis was equitable, it meant that the availability of other forms of fire support was essential.

In the case of small-to-medium navies, which are often operating without aircraft carriers, the potential lack of air support should always be a major consideration. For the ADF, the absence of aircraft carriers and the limitations on the range of land-based aircraft are especially significant. As Andrew Davies warned in 2009: “If we wanted the ADF to be able to provide air cover for the entire immediate neighbourhood, then we would need to use bases located in other countries.” The restricted range of aircraft conducting operations from Australia increases the potential for NGS to be employed within the ADF’s primary operating environment.

Although ground forces often tend to favour their own organic artillery support due to the level of control that can be exerted over it, it has regularly proved to be insufficient to satisfy all requirements without additional firepower. During Operation Corporate, land-based artillery in the form of 105mm light guns, which were largely transported by helicopter, were essential to the success of the campaign. The official lessons of the Falklands campaign concluded that the performance of the artillery, which fired around 17,500 shells, was “excellent.” Despite the commendable performance of the artillery, there can be no doubt that British forces would have benefitted from the availability of medium and heavy artillery. NGS was effectively used as a substitute for medium artillery and proved to be invaluable. The absence of heavier artillery meant that the importance of NGS increased continually as the campaign progressed. The sinking of Atlantic Conveyor on 25 May 1982, resulting in the loss of a number of helicopters, dramatically reduced the lift capability of British forces and, ultimately, made it difficult to rapidly transport artillery to the required locations.

Similarly, during OIF, the transportation of artillery and the range of the guns limited the availability of firepower. Prior to the commencement of the Al Faw peninsula operation, it was intended that artillery and light armour would be landed on the beach by Landing Craft Air Cushion class hovercraft. The

existence of beach mines, however, led to the plan being deemed infeasible by the Americans, which resulted in the withdrawal of the hovercraft.\footnote{Dutton and Waldhauser, ‘Operation Iraqi Freedom’, p. 10; Rear Admiral S. R. Gilmore, ‘The Royal Australian Navy Fleet—Past, Present and Future’, \textit{Institute Proceedings}, The Royal United Services Institute of New South Wales, 16 February 2010; Rear Admiral David Snelson, interview with author, 20 June 2011.} Although the British had positioned 155mm artillery on Bubyian Island, it could not reach the tip of the Al Faw peninsula.\footnote{Lieutenant General Jim Dutton, interview with author, 22 July 2011.} As a result, NGS played a significant, if not entirely expected, role in supporting the assault. NGS has proven to be an extremely important aspect of fire support and whilst it may not always be the most effective, its availability can be critical for forces fighting ashore.

**Guided Missiles: The End of NGS?**

Land-attack cruise missiles have been a relatively topical issue due to the much-anticipated introduction of the Hobart class Air Warfare Destroyers (AWDs) to the RAN in 2016. The ships will be equipped with SM-6 anti-air missiles and have the capacity to fire long-range land-attack cruise missiles, as well as being fitted with advanced sensors to enhance target selection.\footnote{Sea Power Centre-Australia, ‘The Navy’s New Aegis’, \textit{Semaphore}, Issue 7, June 2009; ‘Aussie Rules: Air Warfare Destroyers Push Boundaries’, \textit{Jane’s Navy International}, vol. 116, no. 8 (2011); Australian Association for Maritime Affairs, \textit{Australian Maritime Digest}, vol. 187, no. 5 (1 February 2010), p. 5; Royal Australian Navy, ‘Air Warfare Destroyer (AWD)’, <http://www.navy.gov.au/ship/boats/avd/> [Accessed 28 September 2013]; Henry S. Kenyon, ‘New Ships for 21st Century Missions’, \textit{Signal}, vol. 64, no. 2 (October 2009), p. 30.} In 2012, the RAN declared that their “high end war fighting capability” had been “given a major boost” following the successful firing of a Harpoon Block II Missile at a land target range in the United States by HMAS \textit{Perth}.\footnote{Royal Australian Navy, ‘Harpoon Block II Firing a First for Navy Capability’, 22 August 2012, <http://www.navy.gov.au/news/harpoon-block-ii-firing-first-navy-capability> [Accessed 28 September 2013].} The GPS-guided missile provides a 500lb blast and can be used against a variety of land targets. The capacity to conduct long-range attacks against both land and waterborne targets was viewed as a significant advancement for the RAN.

The RAN’s commitment to missiles has mirrored developments in partner nations. Admiral Robert Natter, who occupied the dual role of Commander-in-Chief, US Atlantic Fleet and Commander of US Fleet Forces Command during OIF, has conceded that whilst the naval gun is “a very capable asset” and “still a good asset to have”, by 2003 it was no longer central to American naval thinking: “The short answer is that it is not as critical as it was in the past because in the past it was our main battery and then we developed an air plane, and then we developed missiles and then we developed guided missiles.”\footnote{Admiral Robert Natter, interview with author, 16 February 2012.} NGS is clearly not as critical for the USN as it once was, and, in truth, it will never be as important as it was during the Second World War.
However, NGS does still have a useful role in the twenty-first century, particularly for small-to-medium size maritime powers.

For smaller navies, such as the RAN, cost will be an important and ever-present consideration, especially in an era of austerity. During the Korean War, Rear Admiral J. E. Gingrich (USN), who was in command of the blockade and escort forces as Commander, Task Force 95, reasoned in mid-1952 that: “The cost of a 5-inch shell at the end of the Korean pipeline was approximately $200 … Unless it did that much damage, we were hurting ourselves more than the enemy.”

The logic behind Gingrich’s reasoning still holds firm and there is no doubt that financial considerations will be a fundamental factor in the selection of ordnance to be used.

With the cost of a single Tomahawk missile being reported as $US 1.5 million, a critique of US spending on ammunition in The Economist concluded: “Money, as Cicero observed more than two millennia ago, is the sinews of war. That is still true today. But lately, from the American point of view, things have got ridiculous. Guided missiles, in particular, are ludicrously expensive.”

Despite advancements in the development of naval shells, the cost remains significantly lower. Even improvements in accuracy are doing little to reduce the disparity in cost. For example, the low cost munitions project in the United States was carried out with the intention of developing a guidance and control unit for under $3,000 per round. If the required effect can be achieved using a cheaper projectile, it is only logical that navies should seek to utilise NGS, when the circumstances lend themselves to such a decision.

The number of missiles maintained in national inventories will be determined by a range of factors, such as the amount of vessels capable of utilising the projectile, production lines and cost. It is highly unlikely that the level of stockpiled land-attack missiles would exceed the number of shells in a navy’s inventory. Therefore, although the intended purpose is not always identical, when there is an overlap, the recourse to land-attack missiles is far from an inevitable decision. The example of the RN during OIF provides a pertinent example for the RAN. Only two of the RN’s vessels were capable of firing Tomahawk Land Attack Missiles (TLAMs) and both of those vessels were submarines. As none of the surface vessels were capable of firing TLAMs, the flexibility of British forces was reduced. Nonetheless, as Rear Admiral J. E. Gingrich (USN) pointed out, cost was a crucial consideration.

---

Admiral David Snelson, Commander, UK Maritime Force and Deputy Coalition Force Maritime Component Commander during OIF, explained, TLAMs “lend a particular capability to the UK expeditionary effort, providing precision strike in all weathers, without endangering aircrew, and are available at short notice”.58

Information on the quantity of TLAMs that are maintained by the RN is of limited availability in the public domain, but it was revealed that sixty-five were originally purchased, one of which was fired on the test range.59 During the conflict in Libya, a “defence industry source” reportedly declared: “At this rate we are using up five or ten per cent of our stock [of TLAMs] per day and soon it could become unsustainable … What if the strikes go beyond a second week? We will simply run out of ammunition.”60 As a result of the finite resources possessed by the RN, the use of TLAMs and the impact on stock levels needed to be taken into consideration. Under such circumstances, it was hardly surprising that British and French naval forces found a role for NGS in Libya. Even with the addition of precision guided land-attack missiles to the service’s inventory, the selection of projectiles during any amphibious operation will need to be seriously deliberated by the RAN.

Whilst there are a great many benefits to using precision guided munitions over ‘dumb’ projectiles, there are some advantages to the use of shells. In certain scenarios, the spread of fire provided by NGS is exactly what land forces require. It was estimated that during Operation Corporate, a barrage of twenty-five rounds of 4.5" naval gunfire was dispersed within an area that was smaller than a tennis court.61 RN ships carried two different types of 4.5" guns during Operation Corporate. The newer Mark 8 was deemed to be more accurate and provided a greater rate of fire, whilst the older Mark 6 was considered to be more reliable. The divergences in accuracy were somewhat of a double-edged sword. For precision targets and missions in support of friendly troops in contact, any increase in accuracy was obviously advantageous. However, in the harassment role, which was designed to restrict the enemy's movement and create the feeling that they could be killed at any given moment, a small dispersion of shells could be beneficial. Lieutenant Commander Ian Inskip, the Navigating Officer of HMS Glamorgan, noted: “Whereas the Mark 8 would place shells in a very tight pattern, the Mark 6 tended to spread them about. When firing five or ten

58 House of Commons: Defence-Minutes of Evidence, Wednesday 3 December 2003, Examination of Witnesses, Rear Admiral David Snelson and Brigadier James Dutton CBE ADC.
59 Ibid.
rounds, one wanted to blanket an area rather than land shells in the same place.”\textsuperscript{62}

The psychological strain inflicted on Argentine troops by NGS was also significant. One Argentine soldier noted: “We were very demoralized at that time because we felt so helpless, we couldn’t do anything. The English were firing at us from their frigates and we couldn’t respond.”\textsuperscript{63} The demoralising capability of NGS was never more apparent than during the capture of South Georgia. The extensive bombardment by HMS \textit{Plymouth} and HMS \textit{Antrim} was recorded as the main factor in inducing the Argentinian troops to surrender.\textsuperscript{64}

During OIF, whilst accuracy was obviously of paramount importance, the ability to spread fire was one of the great advantages of NGS. The intention, where possible, was to clear Iraqi troops and force them to surrender, rather than kill them.\textsuperscript{65} In order to achieve that objective, the first round fired by the ships was normally offset by one hundred yards, causing the round to fall close to the target without hitting it. The Royal Marines then opened fire with personal weapons, whilst at the same time using loudspeakers to coax Iraqi forces into surrendering. In most instances, Iraqi forces capitulated by that stage, but if they did not then two or three rounds were fired on top of the target.

One of the benefits of the naval gun is its versatility in terms of the types of rounds that can be used. Aside from high explosive rounds; star shells for illumination, white phosphorous rounds for smoke, armour piercing and anti-aircraft rounds, amongst others, can be fired. In addition, different fuses can be utilised to allow for explosion on impact or after a set time. In some circumstances, such as Libya, star shells can achieve the required effect, whilst at the same time reducing the risk of collateral damage. Although it can be essential under some circumstances, precision accuracy is not always required and, in some instances, it is not even desired. Each situation needs to be evaluated on its merits and it is important that the NGS capability forms part of the options available to land forces.

\textbf{NGS and the Australian Amphibious Force}

The intended development of an effective amphibious force has necessitated a process of restructuring. As part of Plan Beersheba, 2RAR will form the backbone of the Australian Army’s contribution as a specialist amphibious
infantry unit. Perhaps inevitably, the public focus has been drawn to the decision to purchase a range of amphibious vessels, particularly the keynote acquisition of two Canberra class amphibious assault ships (LHD). Amongst the significant transformations and key acquisitions being pursued by the ADF, the humble surface combatant has maintained an important role. As the RAN outlined in 2005:

> The highly capable surface combatant is well placed to provide a flexible and rapid response to the broad range of threats that may arise in Australia’s large sovereign area, as well as supporting our nation’s global interests further afield.

In *The Navy Contribution to Australian Maritime Operations*, the RAN acknowledged that: “Surface combatants provide unique capabilities that complement other elements of the ADF force structure to ensure a layered defence or concentration of effects against a wide range of threats.” Sea combatants are expected to engage in anti-submarine warfare (ASW), AAW and anti-surface warfare to protect advance forces (hydrographic, mine warfare and clearance divers) and the amphibious force itself, as well as other maritime power projection forces.

While the surface combatant will be required to undertake a range of tasks, *Australian Maritime Doctrine* (2010) made it clear they offered “considerable potential to contribute to combat operations on land and in the air”:

> Medium calibre guns in surface combatants can be used for fire support or bombardment operations, while air warfare sensors and weapons contribute to counter air operations. These capabilities will be particularly useful when networked with AEW&C and fighter aircraft, or with land-based sensors and weapons. Army battlefield helicopters (embarked in the Amphibious Ready Group (ARG)) and naval combat and utility helicopters can also provide extensive support to operations on land.

Whilst the introduction of another capability, especially an additional form of fire support, necessitates consideration, not least because of the

---


69 Ibid., p. 88.

70 Ibid., pp. 91-2.

requirement for deconfliction, the surface combatant is still a critical component of an amphibious force.\textsuperscript{72}

RAN doctrine encapsulates the idea that surface combatants can play a crucial role during amphibious operations. In the area of maritime power projection, which is considered to include: "the landing of amphibious or special forces, the delivery of land forces by sea, and the provision of bombardment by guided or unguided weapons from seaborne platforms", surface combatants have the potential to be useful contributors.\textsuperscript{73} In some circumstances, surface combatants can provide an essential increase in firepower, whilst in others, they can be utilised to reduce the amount of land-based artillery required and, consequently, lessen the logistical demands of the amphibious force. Of the five types of amphibious operations outlined in Australian joint doctrine (raid, demonstration, assault, withdrawal and military support operations), NGS has the potential to be of great utility to all but military support operations, which incorporate tasks such as relief efforts, peacekeeping and humanitarian assistance.\textsuperscript{74}

The structure of Australia’s amphibious force and the ADF in general not only leaves room for the continuation of the NGS capability, but actually warrants it. With the Anzac class still in the service with the RAN, the amphibious force can continue to rely on the support of the 5" gun and its capacity to fire twenty rounds per minute. The Hobart class AWDs, in addition to anti-air and anti-ship missiles, will also be equipped with a 5" gun. The indication is that future frigates will also possess the naval gun. The very fact that the United Kingdom and Australia are considering working jointly on future frigates, including the RN’s Type 26 Global Combat Ship, at a time when interest in NGS has been rejuvenated in the United Kingdom suggests that the naval gun will remain an important tool in the RAN’s kit bag.\textsuperscript{75}

Although some tasks are likely to occupy Australian forces more than others, all contingencies must be prepared for in advance. As John Blaxland noted, “while we hope the amphibious capability will never be called upon for warfighting, it must be prepared to do so”.\textsuperscript{76} Australian amphibious forces will obviously be able to utilise other forms of fire support than NGS during

\begin{footnotesize}
\begin{enumerate}
\item\textsuperscript{73} Commonwealth of Australia, \textit{The Navy Contribution to Australian Maritime Operations}, p. 92.
\item\textsuperscript{76} John Blaxland, ‘In War and Peace an Amphibious Capability is Apt’, \textit{The Canberra Times}, 22 October 2012.
\end{enumerate}
\end{footnotesize}
operations, but the naval gun still has a role to play. The capacity of the LHD to carry helicopters means that Tiger attack helicopters can be used to support amphibious operations, either by providing direct fire support or delivering precise target information.\textsuperscript{77} Tiger helicopters offer an impressive array of firepower, including a 30mm gun, 70mm rockets and Hellfire Air to Ground missiles (AGMs).\textsuperscript{78} The 30mm gun is capable of firing 750 rounds a minute at air or ground targets. The Hellfire AGMs, which provide a heavy anti-armour capability, are laser guided and offer the capability to designate the target either before or after firing.\textsuperscript{79}

Nevertheless, the availability of Tiger helicopters to support amphibious forces will not diminish the value of NGS. Although the Australian Army considers the Tiger to be “one of the most advanced armed reconnaissance helicopters in the world”, and there can be doubt that the aircraft will offer the ADF huge advantages in power projection capabilities in comparison to their predecessors, they will still be subject to the same risks.\textsuperscript{80} The Tiger may have far greater survivability and be capable of minimising risks, but they are still vulnerable in a hostile air environment or in the case of adverse weather. The two capabilities should be seen as complementary because as well as offering a range of fire support, the targeting technology fitted in the Tigers could enhance the provision of NGS. Whilst Australia continues to deploy ships armed with naval guns and those weapons can be usefully employed to support forces operating ashore or to achieve a desired effect, NGS will remain a valuable capability.

**NGS: Ready Aye Ready**

The conduct of integrated operational training will be essential to the development of an effective amphibious force. It was noted in *Future Maritime Operating Concept—2025: Maritime Force Projection and Control* that:

> Real time mission simulation, detailed operations analysis, distance education systems and opportunities to learn from the experiences of other forces will enhance the capability of maritime force personnel. It should be noted that while simulation will permit enhanced training, it will not replace field training and live fire activities under realistic conditions.\textsuperscript{81}

\textsuperscript{77} Commonwealth of Australia, *Defending Australia in the Asia-Pacific Century*, p. 77.


That the surface combatant’s capacity to deliver NGS remains highly valued by the RAN is demonstrated by the training undertaken by Australian ships. RAN ships practise a number of different serials, including communication exercises, system trials and live firings. Non-firing drills are conducted at least monthly and live firings on multiple occasions throughout the year. Post-maintenance trials, exercises and some mission work-ups also include live firings.82

The RAN’s commitment to preserve the skill set suggests that NGS will remain relevant for the foreseeable future and will arguably increase in importance as the focus on an amphibious capability intensifies and developments in fire support are enacted. Indeed, in 2007, the RAN declared: “The development of sea based land attack weapons, such as land attack missiles and extended range gun munitions, will allow maritime forces to responsively support manoeuvre ashore at considerable distances.”83

**NGS: A Current Capability**

The intent to use a capability (if required) is an important determinant in the assessing its usefulness, but so is the potential requirement to do so. The 2013 Defence White Paper outlined four principal tasks for the ADF:

- Deter and defeat attacks on Australia
- Contribute to stability and security in the South Pacific and Timor-Leste
- Contribute to military contingencies in the Indo-Pacific
- Contribute to military contingencies in support of global security.84

Whilst a case could be made for the naval gun during all of the aforementioned contingencies, NGS is far more likely to be required under certain circumstances. In the event of an attack on Australia, the naval gun could foreseeably be used in ship-to-ship engagements. However, from an NGS perspective, it is particularly worthwhile to consider the potential requirement for the capability during contingencies other than an attack on Australia.

The maritime expanse of the South Pacific with its extensive littoral environments creates the possibility that NGS could be useful in the future, as it has been in the past. Hugh White has noted:

82 Written interview with RAN officer, conducted by the author, 4 May 2013.
Australia will face continuing demands to undertake military interventions in small island states in our neighbourhood. Indeed, Australia’s strategic stake in our small neighbours will only increase if, as seems likely, the wider Asian scene becomes more contested between the US and China.\textsuperscript{85} Australia’s experience during INTERFET demonstrated that stability and security operations could potentially lead to the requirement for NGS, even though it was not actually required in East Timor. Restrictive rules of engagement may well circumscribe the scope for employing fire support in many contingencies.

The maritime and archipelagic nature of Australia’s primary operating environments lends itself to the utilisation of NGS in the event of participation in military contingencies in the Indo-Pacific. Although frigates and the AWDs, which will both possess 5” naval guns, will primarily be used for other purposes, their presence during operations ensures that the NGS capability can be quickly utilised as and when required. In the event that artillery cannot be landed or that it is insufficient to address the requirement of the forces ashore, NGS can fill the gap.

Events in Iraq and Libya have demonstrated that NGS can and will be used in contemporary military contingencies in support of global security. There is no reason why the usefulness of NGS should diminish in future operations outside of Australia’s primary operating environment. If anything, the development of Australia’s amphibious force will increase the relevance of NGS in the future. Australia’s history of conducting expeditionary operations provides a platform for future involvement in military contingencies in support of global security. Furthermore, as Australia has frequently acted as part of a coalition and those forces have continued to employ NGS, it is an area where a valid contribution can be made to multinational efforts, often at minimal cost.

**Here Today, Here Tomorrow**

The golden era of NGS has long since passed. The importance of NGS has been diminished by the continual developments in firepower since the Second World War. Despite a resurgence of interest in the capability after the Falklands War, which was followed by the battleships USS *Missouri* and USS *Wisconsin* firing 1,185 shells at enemy shore defences during Operation Desert Storm, the appeal of NGS waned as the twentieth century came to a close.\textsuperscript{86} However, the important, if unexpected, utilisation of NGS during OIF and the conflict in Libya has indicated that whilst navies have had to adapt the way in which the capability is used, it is still relevant in the twenty-first century.

\textsuperscript{85} Hugh White, ’Army Should be Careful with its Aim’, *Sydney Morning Herald*, 1 May 2012.

For the RAN, HMAS Anzac’s involvement in NGS during OIF was a milestone, as it was the first time that an Australian ship had fired in anger since the Vietnam War. Quite simply, there had been no requirement for the RAN to deliver NGS in the intervening period. However, this does not mean it will not be required again, either in support of coalition operations like OIF or in direct support of operations in the South Pacific, South East Asia or the broader Indo-Pacific. The RAN’s limited utilisation of NGS should not be allowed to cloud its continued relevance. Whilst there are many variations of amphibious operations, offensive capabilities must be seriously considered, even though they may only be used infrequently (if at all in the near future). Even if land-attack missiles are freely available to the RAN, there will still be a place for the more modest naval shell, which is not only cheaper and more prevalent, but can, in some instances, be more useful to those fighting on land. NGS came perilously close to being written off in Australia once before, when doubts over the merits of the capability led to the initial design of the Anzac class including the incorporation of the Otto Malero 76mm gun. In that instance, advocates of NGS won the day and it would be wise for that lesson to be heeded into the future.

That is especially true since the decision was taken to transform the Australian Army into an amphibious force. Indeed, as Hugh White has explained: “Amphib [amphibious] ops were not seen as a big priority in the [19]80’s, but today they are seen as central to Army’s future, and in that context NGS makes much more sense”. Other forms of firepower have superseded NGS, but they have not invalidated it. Although NGS is likely to be used relatively infrequently and more sparingly than in the past, it is a capability that should be maintained and practised to provide forces ashore with an additional source of firepower. There will undoubtedly always be people that deem NGS an irrelevance in the modern era, but history has proven otherwise.

Dr Steven Paget is a Lecturer in Strategy and Operations with the New Zealand Defence Force and a Teaching Fellow at Massey University. Prior to that, he was a Visiting Scholar in the History Department at Dartmouth College. He completed an MA and an MLitt in Modern History at the University of St Andrews and undertook a postgraduate fellowship at Emory University. He completed his PhD, which was an examination of interoperability between the Royal Australian Navy (RAN), Royal Navy (RN) and the United States Navy (USN) during coalition operations since the Second World War, at the Australian National University.

---

87 Although capable of providing NGS if required, there was no call for RAN ships to do so during Operation Desert Storm. Rear Admiral Ken Doolan, interview with author, 18 April 2011.
88 Professor Hugh White, written interview with author, 15 October 2013.
Power Flows: Hydro-hegemony and Water Conflicts in South Asia

Paula Hanasz

Water is a point of friction on the Indian subcontinent, which is home to more than 21 per cent of the world's population yet must make do with barely 8.3 per cent of global water resources.¹ Now burgeoning populations—a result of rapid development of the region—are increasing demand for water at an unsustainable rate. Climatic changes affecting glacial melt in the Himalayas exacerbate the problem of water availability.² Scarce in itself, however, is not the only trigger of conflict here; major controversies also exist in the region as to the location and construction of dams.³ Moreover, mutual suspicions and reluctance to cooperate between riparians may impair timely approaches to the collective action problems of non-traditional security threats such as water conflict.⁴ Indeed, fears are rising about the possibility of ‘water wars’.

Proponents of the water wars thesis are vociferous, and too often simplistic in their understanding of what constitutes and contributes to water disputes. Popular commentary and influential books such as Brahma Chellaney’s *Water: Asia’s Next Battleground*⁵ perpetuate the fear that state-versus-state zero-sum games are the new norm for hydropolitics. This unsophisticated view skews assessments of regional security and stability. The reality of

---

water cooperation and conflict is complex and does not manifest itself in the simplistic binary proposition of peace and war.

Nonetheless, the fear of water wars in South Asia persists and is based largely on the following four factors:

- There is intense competition over water resources in South Asia.
- Water scarcity due to climate change and increasing demand from growing populations is expected to exacerbate this competition to a level of outright conflict.
- The region is already politically and socially unstable. Such volatility combined with increasing environmental stresses may create a situation particularly vulnerable to conflict.
- India is perceived as a regional bully.

It is this last point that is the focus of this article, which argues that a significant factor preventing war over water is the very presence of a hydro-hegemon. This is because India elicits consent from its co-riparians for the order it has established, and India’s superior power position effectively discourages any violent resistance against the order.6

Australia is increasingly interested in the challenges of inter-state water governance in South Asia and the potential of significant conflict that may arise out of unresolved issues. Australia is a major partner in the South Asia Water Initiative, led by the World Bank, which aims to increase regional cooperation on shared water issues.7 In October 2012, Australian Prime Minister Gillard visited India to launch the India-Australia Water Science and Technology Partnership to enhance cooperation on water management in the region.8 These are two of the most prominent initiatives that Australia is involved in, but there are many more examples of lower-level collaborations between academic institutions and businesses.

Yet despite the growing interest of Australia in these issues, little research has been conducted into the multi-level dynamics of water conflict and water cooperation—together known as water interactions—in South Asia, nor into how India’s role as the regional hegemon affects these interactions and the prospect of outright conflict. This article attempts to fill this knowledge gap.

This article begins by establishing the concept of hydro-hegemony and how it applies to water conflict and cooperation. The second half of the article systematically addresses each of India’s bilateral relationships with its co-riparians Pakistan, Nepal, Bhutan and Bangladesh. It illustrates that the power asymmetry in the region creates an equilibrium of conflict and cooperation that is highly unlikely to tip over into a ‘water war’. The lengthy and fraught negotiation process between India and Pakistan over the Indus Waters Treaty was ultimately a confidence-building measure that resulted in an agreement that has withstood numerous other conflicts. So too the agreements India has with Nepal and Bangladesh, though problematic, were entered into through consent not coercion. As for Bhutan, this Himalayan kingdom has on the whole benefited tremendously from its cooperative relationship with the hydro-hegemon.

This article takes the realist view that states are predominant actors in international relations and that they act as coherent units. Of course water interactions occur within a complex network of numerous types of actors on multiple levels, but it is not the purpose here to explore these. Similarly, it is not in the scope of this article to assess the numerous inequalities and injustices that arise out of situations of great power asymmetry. While hydro-hegemony may produce unfair outcomes, it also produces stability that mitigates the prospect of water war, and it is this point that is of primary concern here.

**Hegemony and Water Conflicts**

Power relations between riparians largely determine the control over water resources that each riparian wields. Even the upstream/downstream dynamic is predicated on power; those upstream use water to get more power, and those downstream use power to get more water. Unsurprisingly, it is the more powerful side in a water interaction—especially when the competition is over scarce water resources—that achieves and maintains the upper hand. This is not to say that the stronger party will always use brute strength in the attainment of self-serving objectives. Creating compliance through the use of soft power (i.e., having a self-serving ideology freely adhered to by the weaker actor) is a mechanism most frequently utilised by the stronger party.

Power asymmetry of a lesser or greater degree is usually present in water basins, and affects the water interactions that are established, thereby also

---

11 Ibid., p. 442.
12 Ibid., p. 443.
influencing the outcomes of and approaches to water conflicts.\textsuperscript{13} In contexts of a relative balance of power, riparians can decide together whether or not to collaborate on water projects or governance issues; but in an ‘anarchic’ basin, each state acts unilaterally and opts against communicating with its neighbours beyond the bare minimum.\textsuperscript{14} This is the case in South Asia currently, where, as is discussed below, a culture of secrecy prevails.

Hydro-hegemony occurs when one state within a shared river basin asserts its power over other riparian states—even upstream ones.\textsuperscript{15} This is because hydro-hegemony rests not on riparian position alone, but on three pillars:

- **Pillar 1: Power**, consisting of:
  - Political power
  - Economic power
  - Military power

- **Pillar 2**: Riparian **position**

- **Pillar 3**: **Potential** for water resource exploitation.\textsuperscript{16}

For every riparian in any given transboundary water scenario, the strength of each of these pillars is relative to the others, and to the pillars of all other riparians. Needless to say, power asymmetry is highly dynamic because the only constant pillar is riparian position. Furthermore, power is amorphous, relative, subjectively perceived, and therefore impossible to quantify.

‘Hegemony’ connotes domination and coercion, but its real-world applications are nuanced. It can bring stability to a region, provide patronage to weaker states and provide a platform for leadership on issues of mutual concern. Indeed, hydro-hegemony is a significant factor preventing war over water because non-hegemonic states usually comply with the order preferred by the hegemon, whose superior power position effectively discourages any violent resistance against the order.\textsuperscript{17}


\textsuperscript{16} Zeitoun and Warner, ‘Hydro-Hegemony’, pp. 451-452. These pillars of hydro-hegemony have been originally developed by T. Naff and R. Matson in 1984 in relation to the Middle East, and expanded on by A. Medzini in 2001 in relation to the River Jordan.

\textsuperscript{17} Zeitoun and Warner, ‘Hydro-Hegemony’, p. 437.
Hegemony is “leadership buttressed by authority” and is thus distinguished from dominance, which is “leadership buttressed by coercion”.\textsuperscript{18} Certainly, in the case of transboundary water interactions in South Asia, India’s hydro-hegemony has not been an oppressive or destructive force, though India is yet to mature into a true leader on the regional water front.

Whether a hydro-hegemon chooses leadership or domination, and when they choose to do so, is ultimately governed by the broader political context.\textsuperscript{19} Likewise, a weaker party to the interaction may choose to comply with the hydro-hegemon’s direction for reasons other than its apparent or immediate best interest in the matter at hand. Control over water resources, then, is not achieved through violence but rather consent, which is gained through a suite of often subtle mechanisms.\textsuperscript{20}

There are four mechanisms used to produce compliance. The mechanism of coercion can utilise military force, covert action, or direct pressure. Military force is rarely used in water conflicts, and usually then only as a last resort. Covert action could include undercover operations aimed at weakening the political, military or hydraulic apparatus of its competitor, or make a pact with those who will. Direct pressure is arguably the most commonly used tactic in hydropolitics, and includes trade embargoes, diplomatic isolation, threat of military action, espionage and propaganda.\textsuperscript{21}

The utilitarian mechanism employs incentives as a ‘carrot’ to the ‘stick’ of coercive measures. Incentives for compliance with a hydro-hegemon’s preferred state of affairs can include trade incentives, diplomatic recognitions, military protection, and the promise of cooperation on mutually beneficial water projects. These tactics can create stability in the water interactions between the involved parties.\textsuperscript{22}

Normative mechanisms are those that result in a formalised agreement, such as a treaty. A treaty is generally considered an instance of cooperation, however when viewed through the lens of hydro-hegemony this is not necessarily the case. The signing of an agreement to institutionalise the status quo may be to the hydro-hegemon’s advantage, even if the weaker riparian will benefit enough to justify signing.\textsuperscript{23} If an agreement is bilateral, it precludes participation of non-signatory riparians, thereby pre-empting their rights and ultimately leaving the issue unresolved.\textsuperscript{24}

\textsuperscript{18} Ibid., p. 438.
\textsuperscript{19} Ibid., p. 455.
\textsuperscript{22} Ibid., p. 447.
\textsuperscript{23} Ibid., pp. 447-8.
\textsuperscript{24} Ibid.
The ideological mechanism can take many guises—soft power, securitisation, and the creation of a sanctioned discourse. The use of soft power allows issues to be framed by one riparian in a way that their portrayal is accepted without question; somewhere between fully conscious bargaining within an established order and unacknowledged acceptance of that order, the weaker side’s implicit (or explicit) compliance with the outcome is assured.\(^{25}\) An important component of soft power is how the hegemon is perceived by those it aims to influence.\(^{26}\)

The mechanism/s employed by a hydro-hegemon depend on its capacity to persuade subordinate actors to accept not just the hegemon’s authority, but also to adopt and internalise its values and norms intended to impose one solution over others. In other words, hydro-hegemons will avoid coercive tactics if the non-hegemons would comply unwittingly.\(^{27}\) This is the situation in South Asia currently, and we now explore how India as hydro-hegemon has been able to create consent.

**India as Hydro-hegemon**

India scores highly on the three pillars of hydro-hegemony. It wields the most power in South Asia in terms of political, economic and military might. Because of this, it also has significant capacity for water resource exploitation such as the construction of dams for hydropower generation. India certainly has the largest political and geographic stake in the transboundary waters of South Asia. But most rivers in eastern and north-eastern India cut across a number of countries, complicating the number of stakeholders in dispute resolution and treaty negotiations.

Neighbours view India with suspicion, making it difficult to conduct discussion on common-interest issues in good faith.\(^{28}\) India’s hegemony is assumed to be in some way nefarious, undermining regional stability, or necessarily contrary to the interests of India’s co-riparians. But while India’s relations with its co-riparians (except Bhutan) are far from cordial,\(^{29}\) violent conflict is unlikely because ultimately India’s hydro-hegemony is predicated on consent, not coercion. India is able to use its soft power as well as normative and utilitarian mechanisms of hydro-hegemony to overcome the fact that it is not the upper riparian on all the rivers that pass through its territory. India has been able to gain access to upstream water resources


\(^{26}\) M. Bhasin 2008, ‘India's Role in South Asia—Perceived Hegemony or Reluctant Leadership?’, *Indian Foreign Affairs Journal*, vol. 3, no. 4 (October-December 2008), p. 3.


\(^{29}\) Ibid., p. 88.
and influence downstream riparians to overlook transboundary water arrangements that may adversely affect them, all without the use of violence.

The primary way in which India deals with its neighbours is bilateral, and tensions remain about whether transboundary water disputes should continue to be handled bilaterally (the hydro-hegemon’s prerogative) or internationalised. India’s persistence in establishing purely bilateral arrangements and not involving the international community in matters of transboundary water governance in South Asia certainly has been decisive in shaping the politics of water sharing in the region. Even though all the rivers flowing through India are international and pass more than one country, all the treaties on these rivers are bilateral. India did not even acknowledge that usage of the Ganges is an international issue until 1970.

Tellingly, India’s latest National Water Policy devotes only two paragraphs to transboundary waters, and these emphasise bilateralism as the sole approach. Of course this should not be surprising; hegemonic powers benefit through bilateral arrangements, while small and medium powers enjoy greater leverage within multilateral institutions. Yet the lack of multilateralism in South Asia regarding transboundary water management is not solely a manifestation of India’s self-centred ambitions. The large number of riparian countries in the region complicates the processes of multilateral diplomacy and negotiation—of finding common ground for reaching agreement on collective action, norms or rules. In the interest of avoiding conflict, bilateralism is a faster and surer approach than multilateralism.

International agreements, be they bilateral or multilateral, are rarely instances of pure cooperation; more often than not they belie conflict and asymmetry of outcome. An example of a hegemon enshrining inequalities in international agreement is the Treaties of Friendship that India instigated in the region. The treaties of friendship arose at a time when India gained independence from the British and wanted to prevent a communist influence spreading south from China. Consequently, India concluded three lopsided treaties with its small neighbours; Nepal, Bhutan and Sikkim to bring them

---

31 Ibid., p. 60.
35 National Research Council of the National Academies, Himalayan Glaciers, p. 91.
36 Ray, ‘Global Conventions and Regulations on International Rivers’, p. 84.
into its sphere of influence in order to make the Himalaya a strong natural border with China. A similar treaty was signed with Bangladesh in 1972. Unsurprisingly, India has since been reluctant to update any of these treaties. But while this is a source of some discontent (especially in Nepal), there has been no real momentum from these smaller states to agitate for change. One tactic available to the non-hegemon is that of issue linkage wherein non-water-related issues are included in negotiations over water as a point of leverage. This, however, has not been a tactic utilised by India’s co-riparians to alter the status quo of the Treaties of Friendship.

India’s insistence on secrecy regarding hydrological data is contributing to the sense of distrust within the region, and increases tensions about transboundary water management. A striking feature of many transboundary hydropower projects in South Asia is that they were not known through government-to-government communication, but through public media. Timely and adequate information is never easily or fully given. This has been Pakistan’s complaint over the Baglihar proposal, and Bangladesh’s complaint over the Tipaimukh and the National River Linking Project in India. India’s 2012 National Water Policy hints at declassifying more hydro data, but as the balance of power is already tipped in India’s favour, there is arguably little political imperative to do so. Moreover, a culture of secrecy and suspicion prevails across all governments in South Asia, thus stifling any inclination to declassify or share data.

As we can see, on the whole India’s hydro-hegemony has been a factor in preventing violent conflict because it creates compliance among the co-riparians rather than relying on coercion to maintain the status quo. The rest of this article will discuss how this occurs in each of the bilateral co-riparian relationships India has.

39 Bhasin, ‘India’s Role in South Asia’, p. 16.
40 Subedi, *Dynamics of Foreign Policy and Law*, p. 5.
42 Singh, *Trans-boundary Water Politics and Conflicts in South Asia*, p. 16.
43 Government of India, ‘National Water Policy (2012)’.
India and Pakistan

Pakistan is downstream of India on the Indus River system, and at the same time is increasingly water scarce. Between 1947 and 2011, Pakistan's population swelled almost twelvefold. Now Pakistan has more than 180 million people, and is likely to have 335 million citizens by 2050. While the population has ballooned, the quantity of water in the single river system on which the country is dependent has remained the same. With per capita availability of freshwater declining at an alarming rate, Pakistan has gone from being a water-surplus country to a water-distressed one. Moreover, the issue of water quality is as important to Pakistan as water quantity. An estimated 40-55 million Pakistanis do not have access to safe drinking water, yet the government spends forty-seven times as much on the military budget as on water and sanitation.

The confluence of water-related problems and military might in Pakistan is troubling for the region. There are grave concerns about what might happen in this water-distressed, nuclear-armed, terrorist-besieged, overpopulated, heavily irrigation dependent and already politically unstable country if its single water lifeline, the Indus River, continues to be depleted at the current rate. This reduction in the Indus River flow could be due to unsustainable irrigation practices by Pakistan, climate change affecting Himalayan glacial melt patterns, or diversions upstream in India. Indeed, there is a widespread perception in Pakistan that Indian control of the Indus water head can be misused to block water to Pakistan and devastate its economy. This perception is a manifestation of the soft power that India wields as hydro-hegemon. In such a context, disputes over water will likely continue to undermine the prospect of a stable and sustainable peace between India and Pakistan.

The Indus Waters Treaty (IWT) is the object of most water-related contention between India and Pakistan. It is both a symbol of cooperation between the two countries (because it has withstood armed conflict) and discord (because it continues to foster resentment). It thus illustrates that conflict and cooperation coexist, and is a case in point of how the absence of war is not equivalent to an efficient and equitable solution. Although it is often hailed as a great example of bilateral cooperation at a time of conflict between the riparian states, this argument ignores the historical contingencies that were so important in constraining Pakistan's initial scope of action.

47 Renner, 'Water Challenges in Central-South Asia', p. 6.
49 Singh, Trans-boundary Water Politics and Conflicts in South Asia, p. 10.
52 Singh, Trans-boundary Water Politics and Conflicts in South Asia, p. 32.
The agreement between India and Pakistan was largely brokered with the assistance of the World Bank and, although there was the appearance of concession, India was able to exercise its power as a hydro-hegemon and achieve its objectives. Yet the eight-year process of negotiating the IWT was an important exercise in compromise and confidence building between the two conflicted states. Both countries wanted the eastern basin of the Indus River system, which is better for agriculture, but Pakistan relinquished it and agreed instead to develop the western basin. India kept control of the upstream areas but refrained for a decade from developing canals and agricultural infrastructure on the eastern basin to allow Pakistan time for its own agricultural development.

The IWT divides the six major rivers of the Indus River system between the two countries. In 1960, India was allocated complete use of the three eastern rivers (Sutlej, Beas and Ravi), while Pakistan was allocated nearly unfettered use of the three western rivers (Indus, Jhelum and Chenab). Both countries are allowed under certain, narrowly defined circumstances, to use each other’s rivers. In practice, Pakistan has little to gain from this provision, as no major rivers originate within its political borders. The Indian Government, however, can significantly limit the flow of water into Pakistan. Nonetheless, India argues that it is actually Pakistan that has the better end of the deal: India contributes more to the Indus than does Pakistan, yet Pakistan is allowed to take proportionally far more than India.

Another example of the fraught relationship between India and Pakistan regarding shared water resources is the imbroglio over the Baglihar Dam on the Chenab River in the eastern basin of the Indus River system. Pakistan has opposed the Baglihar Dam on the grounds that it violates the IWT because of its potential use by India to store or divert waters destined for Pakistan. The IWT does not permit India to build retention or diversion projects on the Chenab, Indus or Jhelum rivers. Construction of the dam began in 1999 and, following Pakistani objections, the World Bank adjudicated in 2005 that the dam would only be filled between 21 June and 31 August, with Pakistan’s prior consent and with specified minimum river flows. Yet, in 2008, India continued to fill the dam well into September, considerably reducing the Chenab’s flow and causing crop damage in

58 Ibid., p. 38.
Pakistan. A World Bank tribunal subsequently asked India to lower the height of the dam.\footnote{Renner, ‘Water Challenges in Central-South Asia’, p. 7.}

India claims the Baglihar project is a fully legal scheme as it involves no water storage and therefore does not violate the IWT.\footnote{Singh, Trans-boundary Water Politics and Conflicts in South Asia, p. 23.} India has accused Pakistan of trying to prevent it from addressing the grievances of the people of Jammu and Kashmir.\footnote{Ibid., p. 23.} This claim is not entirely unreasonable considering that the people of Jammu and Kashmir are in dire need of power and have not been taken into account by the IWT. They believe Pakistan wants to deny them the right to use the state’s own rivers.\footnote{Condon et al., ‘Resource Disputes in South Asia’, p. 38.} Nonetheless, India’s insistence on the legality of the dam is putting strain on the institutions of the treaty by bending the rules of the agreement and violating its spirit of fairness and equality.\footnote{Madhuressh Kumar and Mark Furlong, ‘Securing the Right to Water in India: Perspectives and Challenges’, Our Right to Water (Ottawa: Blue Planet Project, 2012), p. 8.} While this may be unfair, it is significant that the grievances between the hydro-hegemon and its co-riparian are addressed through the established legal framework, to which Pakistan has consented, rather than through outright violence or war.

**India and Nepal**

It is said that water and energy pose the biggest constraint on India's growth because demand for both is increasing at a rate faster than current capacity can provide.\footnote{Chellaney, Water: Asia’s New Battleground, p. 178.} Part of India's energy shortfall can be addressed through domestic hydropower generation potential,\footnote{Anshu Bharadwaj, Rahul Tongia and V.S. Arunachalam, ‘Wither Nuclear Power?’, Economic and Political Weekly, vol. 41, no. 12 (25-31 March 2006), p. 1206.} though there is strong civil society opposition to dam building in India.\footnote{Iswer R. Ohta, ‘Harnessing the Himalayan Waters of Nepal: A Case for Partnership on the Ganges Basin’, in Asit K. Biswas and Juha I. Uitto (eds), Sustainable Development of the Ganges-Brahmaputra-Meghna Basins (New Delhi: Oxford University Press, 2001), p. 110.} Nepal’s enormous hydropower potential could provide a convenient and significant supply of ‘clean’ energy for India’s growing needs. By developing its major rivers, Nepal can provide large hydropower and storage projects to augment the low flows during the lean season and, to a large extent, mitigate India’s power shortfall, particularly in the north of the country.\footnote{Ibid., p. 107.}

Nepal's theoretical hydropower potential is enormous, at 83,000 MW (identified power potential is 42,000 MW).\footnote{Ibid., p. 23.} Revenues from hydropower sales could multiply the growth rate in several Nepalese sectors, including
industry, agriculture and tourism. If properly planned and managed, development of Nepal's storage potential could yield tremendous benefits not just in terms of hydropower generation, but flood control during the monsoon, flow augmentation for downstream irrigation and navigation and water supply. Nepal's lack of capacity for river diversion or water storage is at the root of disputes with India in relation to both hydropower generation and flood control.

Yet hydropower production remains largely undeveloped for domestic uses and export. Nepal produces only 714 MW of electricity for its 30 million citizens from all sources of energy, with the result that power outages for several hours a day throughout the year are common. Only 15 per cent of the population has access to electricity and per capita electricity consumption is among the lowest in South Asia. Indeed, a country that could be a major exporter of electricity actually imports power from India.

Nepal and India (then under British rule) began cooperating over water resources in 1920. Since then, there have been a significant number of positive and mutually beneficial instances of cooperation regarding water resources. To promote high-level coordination in implementing various agreements and understandings, a Nepal-India joint Committee on Water Resources, headed by each country's water resources secretary, has been set up as the umbrella mechanism covering all water related committees and groups. Indian-aided projects, besides establishing modest hydropower-generating capacity, have helped bring 300,000 hectares of farmland in Nepal under irrigation. Nepal and India have also engaged in issues-linkage by bundling together projects related to irrigation, hydropower, navigation, fishing and forestry.

Treaties, however, are not the embodiment of unequivocal cooperation. There are usually power asymmetries at play that can lead to less than equitable outcomes for the weaker party. The hydro-hegemon may structure a treaty to reflect existing inequalities and use utilitarian mechanisms or soft

---

69 Onta, 'Harnessing the Himalayan Waters of Nepal', p. 106.
70 Condon et al., 'Resource Disputes in South Asia', p. 13.
76 Ibid., p. 284.
77 Ibid.
power to attain the weaker riparian’s compliance.79 This can be said to be the case in the transboundary water management treaties and agreements between India and Nepal. There are several such treaties regarding the development of shared hydropower and irrigation projects, but most, though considered reasonable from India’s viewpoint, are seen as a sell out by many in Nepal.80 The Koshi and Gandak river treaties of the 1950s did not deliver the expected benefits to Nepal (even following amendments in the 1960s) and have left a bad legacy that, to this day, weighs down bilateral cooperation.81

Even the implementation of the more recent Mahakali Treaty (1996) has been impeded by different interpretations of its terms.82 Indeed, it has been argued that the saga of the Mahakali Treaty presents itself as a case in point of Nepali polity and its relationship with its big southern neighbour.83 It certainly provides an interesting illustration of how India has used its hydro-hegemony to securitise water issues between the two countries, especially in the Tanakpur Barrage controversy. India began construction on the barrage on its own side of the Mahakali River then used discursive power to create a narrative of ‘panic politics’ that put pressure on the Nepalese Government to allow construction on the Nepali side.84

Considering how easy it is for India to securitise water—to make it seem a threat in need of emergency action—it is understandable that Nepal is troubled by the reversal of the usual upstream-downstream dynamic. Resentment lingers, for example, over Nepal's obligation to inform India of its proposed non-consumptive developments (including the construction of small and medium hydropower plants). Meanwhile, India does not have to inform Nepal of the same, even when it developed large-scale irrigation schemes in Bihar and Uttar Pradesh. Because of an ambiguity in a bilateral agreement India claims it has prior right to consumptive water use in watercourses shared by the two countries.85

At the same time, there is enormous political pressure within Nepal to harness water resource vis-a-vis India, particularly through mega projects, which are believed to bring more economic benefit and political kudos than more moderately sized undertakings.86 Moreover, Nepali party functionaries

82 Ibid., p. 284.
fear political allegations that any project is not moving forward due to their action or inaction, irrespective of technical, economic or developmental demerits of such projects. They prefer to be seen supporting projects rather than questioning them.87

Several commentators have suggested that India has not been sensitive to Nepal's concerns, nor has it handled wisely the difficult and complex transboundary water negotiations.88 As hydro-hegemon, however, India retains the prerogative to establish the agenda for water interactions, and has no obligation to extend courtesies to its co-riparians above and beyond that which is stipulated in international water law and any other relevant treaties to which India is signatory.

India and Bhutan

In stark contrast to India’s occasionally troubled relationship with Nepal is India’s apparently symbiotic one with Bhutan. Bhutan's water abundance and topography contribute to the country’s propitious energy situation, namely hydropower production. Bhutan's dams have been developed with foreign aid, primarily from India, and it is India that is the largest customer of Bhutanese hydropower.89 India is connected to the Bhutanese hydropower through the 336 MW Chukha project, as well as the Kurichu, Chukha Stage II projects, and the 1,000 MW Tala Dam.90

The collaborative and seemingly friendly nature of the relationship between Bhutan and India in regards to transboundary water resource management can be attributed more to Bhutan's far-sightedness and political adeptness than to India’s attempts are creating more equitable regional relations. India, of course, can use the utilitarian mechanism and elicit compliance from Bhutan (and, indeed, Nepal) because Bhutan is heavily dependent on India for trade and almost entirely reliant on India for navigation and transport routes.91 But Bhutan has been able to turn that to its own advantage.

For Bhutan, the assistance gained from India in developing its hydropower capacity has been crucial in the socio-economic development of the country. Indeed, Bhutan has the distinction of achieving the highest per capita income in South Asia by exploiting its hydropower reserves through environmentally sound projects, mostly small in scale and based on run-of-the-river

87 Ibid., pp. 559-60.
technology, which it then exports to India.\textsuperscript{92} Export of hydropower brings in more than half of Bhutan’s total revenue.\textsuperscript{93} India has not begrudged Bhutan these benefits, and perhaps its willingness to facilitate mutually beneficial outcomes with a cooperative partner could serve as a model for similar cooperation with Nepal in due course. This illustrates that the presence of a hydro-hegemon need not lead to zero-sum games. Bhutan has been able, despite its relative weakness, to manage transboundary water interactions in a way that creates positive sum outcomes.

**India and Bangladesh**

Of all the South Asian states, Bangladesh is in the weakest negotiating position with India. It is also the most likely to suffer the most dire consequences of river diversions upstream and transboundary water resource mismanagement. Bangladesh’s external water dependency (the percentage of water that originates outside a country’s political borders) is 91.33 per cent and thus one of the highest in the world. Yet its per capita freshwater availability (7,569 cubic metres per year) is almost five times higher than India’s.\textsuperscript{94} This, however, is as much a curse as it is a blessing. Bangladesh’s low elevation makes it prone to flooding during the monsoon season, and also prone to drought during the dry season between January and May. Management of shared river resources is therefore particularly critical for Bangladesh.\textsuperscript{95}

Bangladesh is criss-crossed by 230 major rivers, with fifty-four of them (especially the largest ones) flowing in from India. Watercourses cover 7 per cent of the country’s total land area. The Brahmaputra is the most important river of Bangladesh, but the Ganges and Meghna are also significant.\textsuperscript{96} The Ganges-Brahmaputra-Meghna basin sustains approximately 10 per cent of the world’s population, but is also one of the poorest regions in the world.\textsuperscript{97} This poverty makes Bangladesh particularly vulnerable to the deleterious effects of upstream water engineering, such as large hydropower dams in the Himalayas and numerous diversions and water-storage dams in the middle and lower portions of the Ganges. These upstream hydro-engineering projects have caused massive side effects for Bangladesh, including dislocation of human communities, loss of resources such as fishery stocks, and increased hazards from flooding, to saltwater incursions and erosion of the Ganges delta.\textsuperscript{98} Bangladesh is always on the receiving end of these disasters, and so it is all the more important for it to reach

\textsuperscript{93} Nexant SARI/Energy, *Regional Hydro-power Resources*, p. xiii.
\textsuperscript{95} Condon et al., ‘Resource Disputes in South Asia’, p. 8.
agreements that affect water flow upstream. It is also a high priority for Bangladesh in developing a constructive domestic water policy and attaining development goals associated with this.\textsuperscript{99}

There is much criticism of the water-sharing arrangements between India and Bangladesh, which are seen as inequitable and symptomatic of the broader relationship between the two countries. India has aggressively asserted its own interest to the detriment of Bangladesh. The Indian perspective, in contrast, asserts that Bangladesh has been unwilling to compromise and has expected that its share of water resources will always remain undiminished.\textsuperscript{100} There is also a perception within Bangladesh that India secretly diverts a portion of the Ganges upstream during dry months, causing acute water stress and environmental damage to Bangladesh. The Indian External Ministry counter-claims that it releases more water than is Bangladesh’s genuine requirement and that Bangladesh exaggerates its needs.\textsuperscript{101}

Bangladesh is also frustrated by exclusion from Indian and Nepalese negotiations over the Sapta-Koshi barrage despite the project having implications for Bangladesh. Bangladesh has proposed all three countries should work together because the diversion of the Ganges will affect water flow to Bangladesh during the dry season. Yet to date all discussions over the dam, which will likely produce 3,500 MW, have remained bilateral between India and Nepal.\textsuperscript{102}

An example of the hydro-hegemon instigating a treaty that institutionalises an inequitable status quo is India’s signing of the double bilateral Ganges River treaties with Nepal and Bangladesh. The treaties helped maintain the status quo in favour of India and may be viewed as exclusionary devices within an inclusive process, which did confer some benefits onto the other, weaker riparians.\textsuperscript{103} India is likely to continue its pattern of resolving its water disputes with Bangladesh through bilateral negotiation rather than a regional approach.\textsuperscript{104} As hydro-hegemon, it is able to set the terms of cooperation and Bangladesh is likely to continue consenting to this stable arrangement.

A point of major discord between India and Bangladesh is the Farraka barrage. The barrage diverts water from the Bhagirati-Hoogli river system to flush out sediment from Calcutta, but Bangladesh argues that this harms the agro-ecological and economic wellbeing of southern Bangladesh.\textsuperscript{105} The

\textsuperscript{100} Hill, ‘Boundaries, Scale and Power in South Asia’, p. 92.
\textsuperscript{101} Singh, Trans-boundary Water Politics and Conflicts in South Asia, p. 35.
\textsuperscript{102} Ibid., p. 34.
\textsuperscript{103} Zeitoun, Mirumachi and Warner, ‘Transboundary Water Interaction II’, p. 163.
\textsuperscript{105} Ibid., p. 10.
Farraka barrage embodies all that is wrong in the water relationship between Bangladesh and India. It disregards Bangladesh’s ecology, water needs and the survival of its people. It is therefore seen as unfair treatment meted out and the disregard shown to a smaller country by a ‘big brother’.\textsuperscript{106}

A second significant dispute between India and Bangladesh is over the proposed National River Linking Project, which Bangladesh claims will lead to flooding in Bangladesh and intensify the country’s dry season.\textsuperscript{107} Conceived in 1982, the megaproject would transfer 178 billion cubic metres of water from India’s northern rivers including the Ganges and the Brahmaputra to the drier southern rivers through thirty diversion projects, including 3,000 storage reservoirs and 14,900 km of canals. The plan had been dormant for two decades but was revived in 2002 when India ordered that the entire project be completed by 2016, a timetable that has no chance of being met.

One reason that India gives for the urgency of the project is the need to increase the country’s irrigation capacity in order to meet the growing demand for grain, which, by 2050, would have to be sufficient to feed an estimated 1.5 billion people.\textsuperscript{108} India is also an irrigation intensive country,\textsuperscript{109} and the National River Linking Project might provide enough water to increase irrigated farmland by more than 50 per cent and to power 34,000 MW of hydropower capacity.\textsuperscript{110}

If this project comes to pass, it may increase hydropolitical tensions in South Asia. Like Bangladesh, Pakistan and Nepal have voiced their opposition.\textsuperscript{111} The National River Linking Project has also been met with widespread criticism by environmentalists and protests within India.\textsuperscript{112} It remains unclear, however, to what extent the regional fears about the projects are justified, and to what extent they are a manifestation of anti-Indian sentiment pervasive in South Asia.

\textbf{Conclusion}

India is the hydro-hegemon in various guises to Pakistan, Nepal, Bhutan and Bangladesh. It is often perceived as a bully, but when analysed closely, its actions are conducive to stability rather than conflict. This is not to say they are always equitable for all stakeholders, but they are unlikely to incite a water war. India does not dominate water interactions in the region, nor

\begin{itemize}
\item \textsuperscript{106} Singh, \textit{Trans-boundary Water Politics and Conflicts in South Asia}, p. 20.
\item \textsuperscript{107} Condon et al., ‘Resource Disputes in South Asia’, p. 11.
\item \textsuperscript{109} Chellaney, \textit{Water: Asia’s New Battleground}, p. 211.
\item \textsuperscript{110} Pearce, \textit{When the Rivers Run Dry}, p. 223.
\item \textsuperscript{111} Hangzo, ‘Transboundary Rivers in the Hindu-Kush-Himalaya (HKH) Region’, pp. 4-5.
\item \textsuperscript{112} Singh, \textit{Trans-boundary Water Politics and Conflicts in South Asia}, p. 21.
\end{itemize}
does it achieve its hydro-hegemony through coercion. It does have a strong and long-standing preference for bilateralism, which generally favours the hydro-hegemon. Its insistence on maintaining strictly dyadic relations has not been challenged by sub-regional groupings or multilateral institutions in South Asia, though capacity arguably exists for the development of such bulwarks to India’s power. Instead, India has used soft power, and normative and utilitarian mechanisms to elicit compliance from its co-riparians while its superior power position effectively discourages any violent resistance against the status quo.\(^\text{113}\)

In short, India’s hydro-hegemony has created consent and stability in the transboundary water interactions in South Asia thus making water wars unlikely. Tensions certainly exist about the management and development of shared rivers, but declared war or violent skirmishes are likely to undermine the complex system of mutually beneficial arrangements that currently exist. Such a state of affairs can be said to epitomise the ambiguous nature of transboundary water interactions; conflict and cooperation always coexist, but an equilibrium is possible that precludes water war.

Understanding these nuances of transboundary water interactions in South Asia will assist Australia in developing a conflict assessment that is neither unnecessarily alarmist nor simplistic. Australia is strongly placed to bring this sophisticated view to the regional initiatives for transboundary water cooperation it has engaged in, namely the South Asia Water Initiative and the India-Australia Water Science and Technology Partnership. There is an opportunity for Australia to provide a counterpoint to the proponents of the alarmist and counterproductive water wars thesis, and it should be taken. Australia must appreciate India’s powerful position in the transboundary water management in South Asia—and understand that while hydro-hegemony may act as a safeguard against the prospect of water war, it does not necessarily create equitable outcomes for the weaker parties.

\textit{Paula Hanasz is completing a PhD at the Crawford School of Public Policy at the Australian National University. Her thesis concerns hydro-hegemony and the securitisation of water in South Asia}. \url{ paula.hanasz@anu.edu.au}.

Notes for Contributors

*Security Challenges* contributes to innovative and practical thinking about security challenges of major importance for Australia as well as the Asia-Pacific and Indian Ocean regions. The journal’s website can be found at [www.securitychallenges.org.au](http://www.securitychallenges.org.au).

Possible topics of interest include but are not limited to: emerging security threats and challenges in the Asia-Pacific and Indian Ocean; the security role of the major powers; the management of Australia’s security relationship with the United States and other allies; strategies for Australia’s relationships with its neighbours; Australia’s and the region’s resource and economic security, the challenge of defence transformation in Australia and other countries; and strategies for managing and combating international terrorism.

*Security Challenges* welcomes submissions from any source. Early career scholars and new strategic thinkers are particularly encouraged to submit. Authors are strongly encouraged to submit manuscripts via email to editor@kokodafoundation.org preferably in MS Word format. The receipt of manuscripts will be acknowledged within 7 days.

*Security Challenges* contains comments as well as regular articles. Recommended length for comments and opinions is 2000-4000 words, for articles 5000-7000 words. Articles exceeding 8000 words are unlikely to be published. An abstract of no more than 100 words and an ‘about the author’ note of no more than 50 words should accompany the submission.

Each manuscript must be accompanied by a statement that it has not been published elsewhere and that it has not been submitted simultaneously for publication elsewhere. Authors are responsible for obtaining permission to reproduce copyrighted material from other sources.

The refereeing policy for articles requires that the anonymity of the author of the article is preserved. The anonymity of referees, whose comments may be forwarded to the authors, is likewise preserved. The review process normally takes about 4-8 weeks. The editor is responsible for the selection and acceptance of articles; the opinions expressed in articles published and the accuracy of statements made therein are solely the responsibility of the individual authors. The editors disclaim responsibility for statements, either of fact or opinion, made by the contributors. The editors retain the right to condense articles.

Authors receive three free copies of the issue in which their article/comment/opinion appears as well as an electronic version of the issue in PDF-format.

All parts of the manuscript should be type-written and double-spaced. The manuscript pages should be numbered consecutively throughout the paper. Authors should follow the style used in this issue. A detailed style guide can be found on the journal’s website at [http://www.securitychallenges.org.au/SCStyleGuide.pdf](http://www.securitychallenges.org.au/SCStyleGuide.pdf). It is the author’s responsibility to ensure that the submitted manuscript complies with the style guide. The editors reserve the right to reject manuscripts which do not accurately follow form and style requirements.
About the Kokoda Foundation

The Kokoda Foundation has been established as an independent, not-for-profit think tank to research and foster innovative thinking on Australia’s future security challenges. The foundation’s priorities are:

- To conduct quality research on security issues commissioned by public and private sector organisations.
- To foster innovative thinking on Australia’s future security challenges.
- To publish quality papers (The Kokoda Papers) on issues relevant to Australia’s security challenges.
- To develop Security Challenges as the leading refereed journal in the field.
- To encourage and, where appropriate, mentor a new generation of advanced strategic thinkers.
- Encourage research contributions by current and retired senior officials, academics, business people and others with relevant expertise.

Membership

The Kokoda Foundation offers corporate, full and student memberships to those with an interest in Australia’s future security challenges. Membership provides first-release access to the Kokoda Papers and the refereed journal, Security Challenges, and invitations to Foundation events. Membership applications can be obtained by calling +61 (0)2 6295 1555 or downloaded from http://www.kokodafoundation.org/join/join.htm.