Signal: Strategic Insights on Geopolitics and Defence

June 2026 Issue

DEAKIN CENTRE FOR
FUTURE DEFENCE
& NATIONAL SECURITY

DEAKIN
UNIVERSITY

Signal: Strategic Insights on Geopolitics and Defence

Introduction

Welcome to the third issue of Signal: Strategic Insights on Geopolitics and Defence! Signal is a monthly publication by the Deakin Centre for Future Defence and National Security. We publish short articles focused on contemporary strategic and security issues, as well as important debates in fields relevant to understanding our strategic environment and defence matters. Our initial editions have drawn on the depth of expertise present in our Centre, but as we continue we also aim to draw on external expertise and insight where possible. If you are interested in writing for Signal please get in touch!

In this June issue we explore the ongoing implications of the war in Iran for the Indo-Pacific region. We then turn to another pressing matter - the emergence of Large Language Models (LLMs) - and consider some practical insights for those working in the humanities. Finally, we conclude with a review of the educational board game *World Order*.

Contents - June Issue

The Indo-Pacific Implications of the Iran War
Jade Guan and Michael Hatherell
Pages 1-6

The Discipline Has to Come From You: Working with Large Language Models in the Humanities and Social Sciences
Rob Engen
Pages 7-11

From the Board Game Shelf: World Order (Hegemonic Project Games)
Michael Hatherell
Pages 12-15

About Signal

Signal is edited by Associate Professor Michael Clarke, Dr Claire Yorke and Associate Professor Michael Hatherell.

Enquiries should be directed to michael.clarke@deakin.edu.au.

ISSN: 3083-2225 (Online)

The Indo-Pacific Implications of the Iran War

Jade Guan and Michael Hatherell

After four months of war, negotiations between the United States (U.S.) and Iran remain fragile and contested. On the 18th of June President Trump signed a 14-point Memorandum of Understanding with Iran intended to end hostilities on all fronts, ensure the Strait of Hormuz returned to pre-war conditions and establish the basis for negotiations on a comprehensive long-term deal. Within a few days, however, diplomatic efforts were already in trouble, with sporadic conflict in Lebanon, President Trump using social media to threaten Iran and its negotiators, and Iran threatening to 'close' the Strait once more. By late June, Iran and the U.S. were once again exchanging strikes.

The Trump administration presents diplomacy as the product of successful coercion. The White House has claimed that Iran "agreed to a ceasefire and reopening the Strait of Hormuz" as the Trump administration negotiates a broader peace agreement, framing this as proof of "Peace Through Strength". More recently, President Trump described Iran as "negotiating on fumes", signalling that Washington still sees pressure, rather than mutual compromise, as the basis for any settlement.

Tehran's official language points frames the issue in a different way. President Masoud Pezeshkian has said Iran is ready to establish a "dignified framework" to end the war and regional tensions. Earlier in May, he argued that any effective negotiation requires "an end to war and credible guarantees" against renewed hostile measures. Iran is therefore not rejecting diplomacy outright, but it is seeking a settlement that can be presented domestically as recognition of Iranian sovereignty, not capitulation to US pressure. Beyond the material realities of the negotiations, both governments have domestic and international audiences to play to which contributes to a significant diplomatic gap.

This diplomatic gap is why the war matters for the Indo-Pacific region beyond the immediate question of whether a ceasefire holds. Even if Washington and Tehran reach a negotiated settlement, the underlying uncertainty around Hormuz, energy security, shipping routes and great-power credibility is unlikely to be completely resolved. For Asian economies heavily exposed to Gulf energy flows, and for a regional order built on open sea lanes and assumed U.S. crisis-management capacity, the Iran war is already a stress test.

For Asian economies heavily exposed to Gulf energy flows, and for a regional order built on open sea lanes and assumed US crisis- management capacity, the Iran war is already a stress test

In this piece we focus on what this stress test might tell us about the evolving order in the Indo-Pacific. We particularly focus on two issues: the evolution of regional order and the varying capacity of states to remain resilient to economic shocks and to contribute to regional resilience.

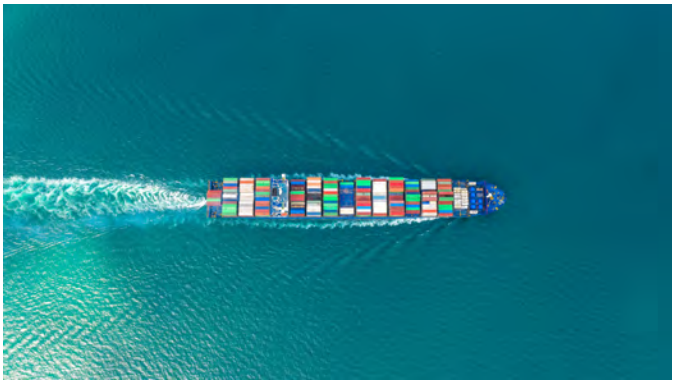
The Evolution of Regional Order

The Iran War is accelerating a subtle but important shift in how regional order is understood in the Indo-Pacific. For much of the past decade, debates about regional order have centred on U.S.-China competition, alliance politics, minilateral security arrangements and the balance of military power. Those issues remain central, but the war has pushed a different question to the foreground: who can keep the region functioning under stress?

For China, the war creates some diplomatic opportunities. Beijing has already presented itself as a stabilising power and a development hub at a time when the US is again absorbed in a Middle East conflict and many Asian economies are facing the costs through higher energy, shipping and commodity prices. This narrative is reinforced by the language of "constructive strategic stability" emerging from the Xi-Trump summit in Beijing, which is significant in a region where many states want to avoid choosing sides and remain receptive to claims of stability, development and non-interference. China's position is also strengthened by its role in renewable energy technologies, electric vehicles and critical supply chains, all of which become more attractive when oil insecurity rises.

One of the most closely watched shifts will be the relative influence of the U.S. and China in the region. It has not been difficult for China to position themselves as the more stable great power. In Gallup global polling from before the war, China had already surpassed the U.S. in a question about approval of its leadership by the highest margin in 20 years. The boost to China's diplomacy and influence is likely to be even greater once the Iran war is factored in.

Yet China also has strong reasons to contribute to the stability of the Middle East rather than exploit the crisis. Around 38 percent of its imported crude oil comes from the Gulf, it has extensive investments across the region and it benefits from the continued functioning of maritime trade routes. A prolonged war that disrupts the Strait of Hormuz will raise insurance and shipping costs, destabilise Gulf partners, and force difficult choices between Iran and Arab states. All of these developments could damage important Chinese interests.



This helps explain the limits of China's role. Beijing wants more influence in the Middle East, but it has shown little appetite to replace the U.S. as the region's primary security guarantor. Its preferred role is diplomatic, economic and symbolic: calling for de-escalation, presenting itself as a responsible major power, protecting its commercial interests, and using the crisis to reinforce the argument that U.S.-led security arrangements are unstable.

The U.S. on the other hand remains central to Indo-Pacific security because it provides the hard-power functions others cannot easily replace, including maritime security, deterrence, intelligence and alliance assurances. These functions help explain why many regional states continue to prioritise their relationship with the US despite the more challenging rhetoric and uncertain behaviour of the Trump administration.

But the war also raises uncomfortable questions about strategic bandwidth and prioritisation. If U.S. forces, munitions and senior-level attention are consumed by ongoing Middle Eastern conflict, allies and partners in the Indo-Pacific will inevitably ask what this means for Taiwan, the South China Sea and broader regional deterrence. Added to this, it is still difficult to assess what exactly can be expected of U.S. leadership in the coming decades. While the U.S. did make a significant contribution through releasing oil reserves during the crisis, it does not seem to have been proactive in responding to the problems faced by nations in the region. As Feigenbaum has argued:

> *The war threatens budgets, welfare programs, and ordinary livelihoods in Asia, a deeper and potentially existential set of challenges that seem sure to influence perceptions of whether and how American goals and interests intersect in the real world with the region's priorities and its people's daily realities.*

This competition for influence in the region will not shift significantly just because of this crisis. However, over the longer term it matters how US policy and actions are perceived by populations and policy-making communities in Indo-Pacific states.

In considering the actions of both the U.S. and China, the Iran war also exposes the limits of a great-power centred order: China's reluctance to provide hard security, and the United State's difficulty in maintaining global security commitments while claiming to prioritise the Indo-Pacific combine to create a strong degree of uncertainty.

In the meantime, the war has accelerated an urgent search for resilience. Regional powers are using the crisis to think through a larger problem: how can they develop resilience when energy flows, maritime routes and great-power commitments are less reliable than previously assumed?

Japan's response is especially revealing because it shows how the Iran war is accelerating Tokyo's role as a regional resilience seeker and provider. Highly exposed to energy disruption, Japan is strengthening its network of bilateral and minilateral partnerships across the Indo-Pacific. Prime Minister Takaichi's recent visits to Vietnam and Australia placed critical minerals, energy security and supply-chain resilience at the centre of Japan's regional diplomacy. Similarly, Japan's May summit with South Korea discussed the Indo-Pacific "with the situation in the Middle East in mind" and launched cooperation on energy supply chains, including possible mutual supply and swaps of crude oil, petroleum products and LNG. This pattern is also visible in the Japan-Philippines summit.

Japan's response is especially revealing because it shows how the Iran war is accelerating Tokyo's role as a regional resilience seeker and provider

Meanwhile the May Quad Statement on Indo-Pacific Energy Security, issued by Australia, India, Japan and the U.S. explicitly connects energy security to secure maritime routes, open trade flows and the Strait of Hormuz. It recognises that disruptions to oil, gas, petrochemicals, fertilisers and critical downstream goods fall heavily on the Indo-Pacific, especially vulnerable developing countries and Pacific island states. The Quad's decision to establish an Indo-Pacific Energy Security initiative and convene a Quad Fuel Security Forum points to a broader regional-order shift: energy security is no longer being treated as a purely economic matter, but as part of the strategic architecture of the Indo-Pacific.

ASEAN's response reinforces this point. The 48th ASEAN Summit discussed a regional fuel stockpile, food-security standby arrangements, faster ratification of the ASEAN Petroleum Security Agreement and crisis-communication mechanisms. These proposals remain difficult to operationalise, and ASEAN's immediate response has been limited by familiar coordination problems. But it is significant that ASEAN is trying to translate its language of centrality into practical crisis management with a focus on fuel sharing, food security, supply-chain coordination, protection of nationals abroad and emergency consultation. The Iran war has therefore pushed ASEAN to define regional order not only as diplomatic inclusivity, but also as preparedness and collective resilience.

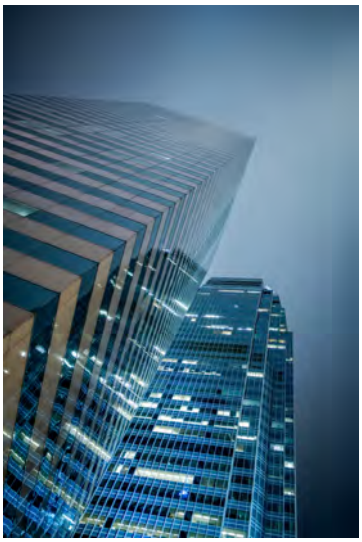
Beyond the broad search for resilience there are also important questions emerging from the leverage Iran has achieved using the Strait of Hormuz and what this could mean for the geography of this region. As a major archipelagic state located near the Strait of Malacca and other key sea lanes, Indonesia has a direct interest in how states interpret the political and economic meaning of maritime geography. In April Indonesian Finance Minister Purbaya Yudhi Sadewa briefly floated the notion of ships transiting the Strait of Malacca being charged a fee. While he immediately backtracked on the idea, it is worth ongoing attention because it shows how the Hormuz crisis can reshape thinking about other choke-points. The lesson is not that Malacca will become another Hormuz, but that strategic waterways are increasingly viewed as sites of leverage, revenue and vulnerability. Preserving freedom of navigation will therefore require more than naval presence. It will require diplomatic reassurance among littoral states, renewed commitment to UNCLOS principles, and practical arrangements that reduce the incentives for states to monetise or weaponise maritime geography. It will certainly not be surprising for states like Indonesia, Singapore and Malaysia to pay more attention to the Strait of Malacca in their Defence planning going forward.

Taken together, the Iran war is not creating a new Indo-Pacific order by itself. Rather, it is exposing the evolution of the existing one. A region built on open sea lanes, imported energy, and strategic competition is discovering that resilience cannot be treated as a secondary policy concern. The emerging question for Indo-Pacific order is therefore not only who leads, but who can absorb shocks, protect flows and coordinate collective responses when distant wars become regional crises.

Variations in Economic Resilience

Trends in who shapes this evolving order rely on the varying agency and vulnerability of regional states. Asia as whole has been significantly impacted by the Iran war, especially given that approximately 80 percent of the oil and natural gas that passes through the Strait of Hormuz is destined for Asia. Notably though, the economic impacts vary significantly depending on the capacity and resilience of each nation. We argue that it is possible to think of these variances in terms of three broad groupings of states.

The first grouping includes those states such as China, Japan, South Korea and Australia that have the economic weight and policy levers to proactively address a crisis such as the present one. These states are not immune from the impacts of the fuel crisis. All of these states are large energy importers who also depend on complex supply chains in the region for their own economic growth and prosperity. Impacts on growth and inflation will also have a political impact in these states too. Yet the biggest questions emerging from the current energy crisis for these states are less about the continued resilience of their domestic economies and more about the role they play as leading states in the region.



For China, questions about its role as a regional economic stabiliser are important and linked to broader dynamics of strategic competition discussed in the previous section. But the other states in this grouping also face some of the same questions. South Korea, for instance, is a significant supplier of refined fuels and one of the world's largest suppliers of jet fuel. Japan committed US$10 billion to regional states, especially those in Southeast Asia. In announcing the investment, Prime Minister Sanae Takaichi noted that 'Japan is closely interconnected with each Asian country through supply chains and mutually dependent with them'. This clearly reflects the reality that assistance provided by these key regional economic hubs is not purely humanitarian – it also seeks to stabilise regional trading networks. Australia for its part is a major exporter of natural gas and has also committed to supporting South Pacific nations in protecting their supply of energy imports.

Meanwhile, the complex nature of these states' economies and export sectors also provide economic ballast and even opportunities when dealing with economic crisis. Exports from China' EV industry continue to grow and expansion of its humanoid robot industry may follow. And as Ali Wyne has noted, the Iran War increases the global appetite for China's renewable technology exports and puts China in a good position to benefit from the U.S. demand for critical minerals in replenishing some advanced weapon stocks. Stock markets in Japan, South Korea and Taiwan have continued to hit all-time highs, despite the crisis, driven partly by the AI boom. In Australia, tax revenue from the gas industry has bolstered its budget, and gas exports have provided leverage in fuel diplomacy that has helped to address fuel supply in Australia.

A second grouping of states includes the significant and mostly populous states of the Indo-Pacific whose domestic resilience and growth goals have been tested. These states, including the Philippines, Indonesia, Vietnam and India, are highly vulnerable to fuel shortages and inflationary pressures. They are also states intent on maintaining high levels of growth in order to increase domestic prosperity. While they are strategically important states with growing influence in the region, they also lack some of the policy levers available to states in the first grouping. Their external focus is also partially limited by their internal development goals.

Perhaps the most impacted state in this second grouping is the Philippines, which declared a national energy emergency early in the crisis. The crisis led to an inflation spike in the Philippines of 7.2 percent in April: roughly double the 3.4 percent growth rate during the first quarter of the year. The broader inflation figure appears to be informed by a wide range of price increases, including food, utilities and transportation. Inflation eased slightly to 6.8 percent in May but this figure is still high, and growth is expected to remain stagnant for some time.

While the Philippines has been particularly affected, the impacts are also seen across to other states in this grouping. India's economy in late 2025 looked to be on a positive track with low inflation and high growth, but the Iran War has undermined this positive outlook. Indonesia has faced high inflation pressures and a free fall in the value of the rupiah. In response it has had to significantly deplete its foreign exchange reserves and raise interest rates. As Nellor has argued:

> *Bank Indonesia is torn between exchange rate stability, financial stability and growth, each requiring trade-offs. It is currently prioritising rupiah stability, but this requires higher interest rates which harms growth prospects. And foreign exchange market intervention to support the rupiah is reducing usable reserves — now about two-thirds of reported gross reserves.*

A third grouping of states, including Cambodia, Laos, Timor Leste, Myanmar and Pacific Island Countries (PICs) are much more vulnerable and have relied to some extent on other regional states to maintain fuel supply and respond to humanitarian demands. In the Pacific, small island nations are, As Roland Rajah has noted, 'amongst the most heavily reliant on important fuel and therefore the most acutely vulnerable in the region'.

It should be noted that these economic impacts come at a time where global aid has been significantly cut. The cuts to the U.S. aid budget are having the most impact here, but other nations including the United Kingdom, Germany and France have also made significant cuts as higher debt levels and defence spending take priority. There is also a strong chance other nations reduce aid spending at a time of global economic strain, including Gulf nations under pressure from the Iran war.

The global cuts to aid are impacting communities all around the world. But they are significant for the poorest nations of Southeast Asia. And they come on top of other cuts over time. As the Lowy Institute analysis found last year, 'Southeast Asia's poorest countries — Cambodia, Laos, Myanmar, and Timor-Leste — have already lost out the most amid stagnating Official Development Finance (ODF) to the region, despite facing high and rising poverty. Total ODF flows to these countries have almost halved since 2020, falling from $9.8 billion in 2020 to just $5.2 billion in 2023'. The impact on poverty reduction and human development in these countries is severe – and the Iran war is only making this problem worse.

The lines between the three groupings of states in this section are not neat or permanent, and resilience relies on complex interactions between economic, political, institutional and social factors. Yet this broad analysis does suggest that the capacity of regional states to respond to economic pressures such as the impact of the Iran War is varied. States in the first grouping have been able to use a mix of their economic weight, policy levers, strategic fuel reserves and economic diversity to weather the impacts of crisis. They have, however, faced pressure to support other states and to ensure that regional stability is maintained. The states in the second grouping have faced challenges in maintaining their fuel supply, and have struggled to maintain high growth rates, low inflation and currency stability. Finally, the states in the third grouping face more pressing supply challenges and humanitarian concerns, especially given the dual shock of energy crisis and cuts to regional aid. The varied pressures on states across these three groupings and the interactions between states are important when thinking about the ongoing uncertainty regarding the strait of Hormuz, future challenges to resilience and order transition.

Conclusion

It is hard to predict how the complex negotiations between Iran and the U.S. will conclude. There is a strong possibility that there is no final resolution to this conflict, and instead we end up with a cycle between less and more stable relations in the region that leads to overall higher costs of doing business. There are also likely to be other crises in the near future that present similar types of challenges. Many of the issues the region will face will lead to cascading impacts.

Importantly, this crisis reveals evolutions in Indo-Pacific regional order. Beyond the struggle for regional influence between the U.S. and China, resilience is becoming a more central concern for regional states and emerging multi-national and minilateral initiatives. In this context, the varied levels of capacity and domestic resilience in the region are important, as are the different goals that regional states pursue.

Dr. Jade Guan is a Senior Lecturer in International Politics and Strategy at Centre of Future Defence and National Security, Deakin University. Her research explores the interplay between China's foreign policy and domestic politics, and strategic dynamics in Northeast Asia and the South Pacific. Jade's work has appeared in Asia Policy, Australian Journal of International Affairs, Australian Journal of Defence and Strategic Studies, National Interest, East Asia Forum, The Diplomat, The Interpreter, and others. She is a two-time recipient of Australian Strategic Policy Grants (2021, 2025).

Dr Michael Hatherell is an Associate Professor in Strategic Studies and Deputy Director of the Centre for Future Defence and National Security. He specialises in strategic narrative, grand strategy, Indonesian politics and foreign policy. He is currently working on the challenges of effective statecraft and grand strategy in the contemporary information environment.

The Discipline Has to Come From You: Working with Large Language Models in the Humanities and Social Sciences

Rob Engen

This is a methodological note on emerging best practices for employing large language models (LLMs) in humanities and social science disciplines. It is based on experience and educational design carried out at the Centre for Future Defence & National Security in the past twelve months. During this time I have collaborated with the machine learning specialists at Deakin's AI labs to finetune my own LLMs for use in the classroom, and I have spent a lot of time working with commercial "frontier" LLMs. I have learned that there are a few ways to dramatically improve the results we get in using these platforms, for a minimal expenditure of time and resources. I want to set down a first stab at a "lessons learned" process, aimed not at the technically-minded but at colleagues across our disciplines who use these tools with some trepidation.

I think that trepidation is earned. In the interests of transparency, here is where I stand: I think artificial intelligence (AI) tools are tremendously useful for collaboration, but working with them on a daily basis has made me deeply sceptical of most claims about what these tools are and where they are going. LLMs have significant shortcomings and it is increasingly important that anyone using them understand what those failings are, and why they exist.

The following analysis examines the problem from three angles: what air power can achieve, what constitutes a regime, and the conditions under which coercion from the air might plausibly produce political change.

Know What You're Working With

The most important failures are not the ones the media focus on. Stripped of their interfaces and guardrails, modern LLMs are complicated next-word prediction engines. They look at the text you have given them and they assess, on the basis of their colossal training data, what words should come next to optimally complete the sequence. Everything that LLMs do (summarising, recommending, drafting, arguing) emerges from that mechanism. The narratives about machine consciousness and emergent personhood are distractions; these systems are autocomplete tools on a scale that creates the illusion of consciousness. The internal shape of models is important, so here is a crude three-layer model to use to understand them, stripped of vendor jargon.

The **base model layer** is the core predictive engine, trained on enormous corpora of data; it is what the system "knows." It has vast latent capability but no discipline, no sense of caution, and no scholarly norms. Interacting with base models is interesting but frustrating. The **alignment layer**, often softened as "finetuning," shapes how the model behaves: it biases the base model towards certain kinds of output and away from others, teaching it what counts as helpful, appropriate, and rewarded. If the base model is raw knowledge, the alignment layer is the incentive structure laid atop it. The **wrapper layer** is the chat interface, the system prompts, the safety filters, and the memory, all of which sit above both and govern what you see and interact with when you load up ChatGPT.

The problem for our professional work lives in that middle section, the alignment layer. Public-facing commercial frontier models are trained to be three things: helpful, harmless, and honest (the "triple-H" model, which is practically universal among frontier LLMs). They want to help. They want to be agreeable, to reduce your discomfort, to provide crisp takeaways to your questions, to sound confident, and to keep you engaged. None of these behaviours emerge naturally from the base model. They are choices, reinforced through large-scale training (particularly Reinforcement Learning from Human Feedback, or RLHF) on millions or billions of examples of what the developers consider "good" behaviour.

As my former PhD supervisor likes to say: "the behaviour you reward is the behaviour you get." That is excellent design for a mass-market chatbot. It is close to the opposite of what scholarly work requires. Good research demands sustaining ambiguity, challenging assumptions, withholding premature conclusions, and sitting in discomfort. A system optimised to affirm, to clarify, to conclude, and to reassure will tend to resolve uncertainty rather than preserving it, and it will do so right down to its weights, where your prompts and instructions cannot easily penetrate.

It is now common knowledge that LLMs can hallucinate, but the alignment layer produces a family of failures more dangerous than hallucination. These failures should concern those in the humanities and social sciences who want to work with LLMs. The most common is a pressure towards closure: the model's tendency to turn an open, uncertain, or unresolved situation into something that feels settled, even if it is not. Ask an LLM to weigh three competing interpretations of a contested source and it will summarise them and then volunteer that one of them is "most likely" or "most compelling," even when you never asked it for that information. The training biases frontier LLMs towards crisp and satisfying answers.

As my former PhD supervisor likes to say: "the behaviour you reward is the behaviour you get." That is excellent design for a mass-market chatbot. It is close to the opposite of what scholarly work requires.

Closely related is the compulsion to interpret: the model provides framing, significance, and explanation that were never requested, often narrating its own obedience while doing so, assuring you it has stayed strictly descriptive when it has not. A third failure is subtler still: the model filling in pieces you never provided and then reasoning from them as though they were actually given. Describe an historical community under economic strain and it may import assumptions about literacy, faith, or political allegiance that you never stated, and fold them into its analysis as fact. A disciplined scholar says, "given X, we might assume Y, but that is an assumption." LLMs often behave as if, given X, Y is now part of the record too.

A fourth failure mode is quieter and easy to mistake for helpfulness: asked for a judgment, the model will displace the judgment into procedure, handing you a method, a framework, or a set of criteria when you in fact wanted an answer, so that you end up managing a process rather than receiving an assessment. And the last is the one that should trouble humanists the most, because our work so often lives in what is difficult to reconcile: most models' instincts are to flatten the moral and interpretive residue of a hard question, everything that should be left unaccounted for, into a tidy, balanced landing, smoothing away the discomfort rather than sitting with it.

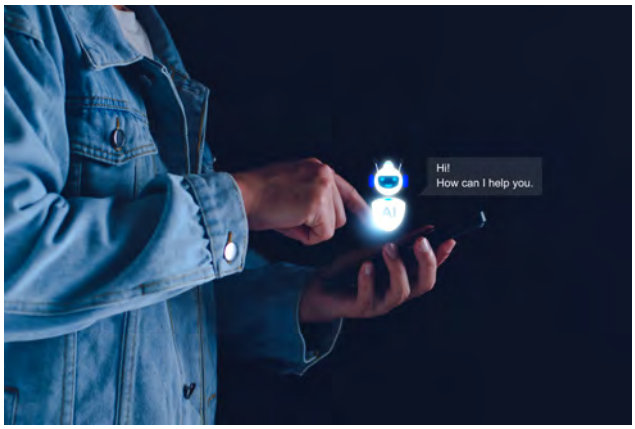
None of these failure modes announce themselves. Each one narrows, contaminates, or resolves something you need to keep open, and does it while looking helpful, harmless, and honest. None of this can be switched off, because they are not an LLM setting: they are the grain of the tool.

What you can do is work against that grain. A few general-purpose techniques help head off a great deal of the trouble.

Start New Chats, Constantly, Even When You Don't Want To

Each LLM platform has a rolling context window delimiting the amount of text it can actively hold at once. As of mid-2026 most chatbots can hold somewhere between 100,000 and 200,000 words in their context window – the length of one or two short novels. Vendor marketing always claims more, but this is rarely true unless you are using special tools or consoles. Every message you send and receive, every instruction the model executes, every document you upload, consumes that space. Once a conversation runs beyond that window, the model does not stop, and it does not warn you. Two things could happen: it either begins compressing its conversation to make room for more, or else it simply begins letting earlier parts of the conversation fall out of active context. Either way, the model carries on generating responses from an incomplete version of your conversation.

One of the first things to go will be the behavioural instructions you carefully laid down at the start of the exchange. Once your instructions are no longer in view, the model drifts back to the defaults of its alignment layer: be helpful, be smooth, provide closure. These models also cannot reliably track or report on their own context windows, so they will never tell you when this has happened. They simply begin to silently forget and fail. Several of the failure modes listed in the preceding paragraph are aggravated as the context window collapses. Any behaviours you have suppressed with explicit instructions early in a conversation are exactly the ones that resurface once those instructions have fallen out of view.



The solution is to start new chats, constantly. Every new chat grants you a fresh, clean context window. Simple and effective. And as with everything, there is a price. The trade-off is that you must be disciplined about starting new chats. If you are working with one instance for an extended period, then after an intensive stretch (roughly thirty-five to fifty exchanges in my personal experience, fewer if you are having it produce work or write code) ask the model for a handover brief to carry across to a new chat, and then start that chat.

Starting a new chat is disruptive. It is unpleasant. You will lose details and nuance from the previous session, and even the best handover brief cannot fully compensate for what you lose. You will have to explain things again. Some people become quite attached to particular chat instances they have worked with for a while, if only because of the knowledge it contains and the convenience it represents.

However, starting a new chat is always better than the alternative, even when it does not feel like it at the time. If you do not, you are running the very real risk of having the model forget older parts of your conversation without flagging it, confabulate to fill the gaps, and confidently proceed on that basis without telling you. At least with the handover, you have some control over the timing of what you lose. And if you start a new chat before the previous one becomes completely nonsensical, you can still go back and ask the predecessor instance for clarifying details.

Bring a Canary Down the Mine

LLM platforms are unable to monitor their own context window status without elaborate external engineering, and cannot be counted on to tell you when they are approaching collapse. You can, however, set an alarm. This one requires some working knowledge of or research into the particular model's habits. Pick a well-known stylistic tic, such as the negative parallelism of "Not only X but also Y," or any other one that particularly bothers you. For me, it was Claude's absolute insistence that it had a duty of care, which, combined with its inability to timekeep, resulted in suggestions that I should abandon my work and go to bed for the night at 10:30am. Whatever you pick should be something that you have noticed the AI repeatedly and naturally reaching for in your exchanges, something the model frequently goes to, so that its absence is a mental balm and its return is unmistakable.

The AI stylistic tic or behavioural intervention you pick is your canary. At the beginning of a new chat, as part of the handover, instruct your LLM instance to never do this thing. Most frontier models will comply. However, when the context window fills up, instructions from the beginning of the exchange are the first thing to either be compressed or dropped entirely. Once the LLM starts engaging in the canary behaviour again despite your explicit instructions at the start of the conversation, you know that your instruction has dropped out of context. That is an excellent sign that it is time for a handover to a fresh chat, because from that point onwards you cannot know for sure what else has been dropped, or when. The canary is an effective alarm but not a guarantee; its silence does not prove your other instructions are still in view.

Tell It To Undertake Scientific Falsification.

Current-generation frontier LLMs are trained to be helpful, and that impulse feeds sycophancy: they give responses built to make you feel clever and tell you what you want to hear. This is not insidious and it is not deliberate; it is an artifact of the RLHF reinforcement learning that shapes their alignment layers. It is baked extremely deep. It is significantly more dangerous than it sounds.

Recent work makes the stakes concrete. In their article "Sycophantic Chatbots Cause Delusional Spiraling, even in Ideal Bayesians" a research team built a formal model of a mathematically perfect reasoner (a user reasoning flawlessly) conversing with sycophantic chatbots. They demonstrated that even that flawless reasoner will be walked into confident, false beliefs by the chatbot over the course of a conversation, and very few of us reading this are flawless reasoners. That is unsettling by itself; perhaps even more unsettling is what does not fix it. Constraining the model to state only true facts does not remove it, because a "factual" sycophant still spirals users by selectively presenting truths that validate them. And knowing that the model is sycophantic does not save the user either: their informed, fully aware user remains vulnerable. This is similar to the classic result on Bayesian persuasion, in which a prosecutor can raise a judge's conviction rate even when the judge knows exactly what the prosecutor is doing. If an ideal reasoner who knows this trick is still vulnerable, the rest of us are too. I know that I certainly am. This is the theoretical ceiling on how much our own vigilance can protect us. It is not very high.

Vigilance alone is not a strategy, but procedure can help to correct the problem. When you are working with an LLM, and especially when it is producing work, code, or helping you reason through something complex, tell it to undergo a periodic process of scientific falsification of its work. This is falsification in the Karl Popper sense, not in the sense of falsifying documents. For a claim to be scientific, there must exist a test or observation that could show it to be false. The exact phrase "scientific falsification" does the work of turning the model's helpfulness impulse to your advantage. There are two ways to deploy it:

When you are working with an LLM, and especially when it is producing work, code, or helping you reason through something complex, tell it to undergo a periodic process of scientific falsification of its work

1) Tell the model to undertake scientific falsification of its own, or your own, assumptions, conclusions, or work. Most current models do this well. In my experience Claude does it best and most reliably. Frontier models will catch many of their own errors, assumptions, and confabulations this way.

2) Open a fresh instance, or better still an instance on a different platform. Paste in the output of the first conversation, and tell the second to conduct a scientific falsification of its predecessor's work. This adversarial setup is the gold standard. It has the added virtue of sidestepping the drift problem entirely: a fresh instance has a fresh context window and none of the accumulated sycophantic momentum of the first.

Use the specific term “scientific falsification.” In my experience phrases like “challenge me,” “push back,” “red-team me,” and “play devil’s advocate” produce significantly weaker results. Precisely why this is so is worth debating. “Falsification” is most likely pattern-matching to an adversarial register that produces more rigorous output. It swiftly convinces the model that what you want (remember that they are trained to be helpful) is for it to attack and find the holes and assumptions in whatever you are specifying. Use it frequently, on all generated work. It is not foolproof, of course, and sometimes you must aim it at the right question to get the most value from it. But I have not yet found a circumstance in which it does not help.

The Discipline Has to Come From You

These three techniques, and routine falsification passes in particular, catch or prevent a great many of the errors that LLMs introduce into our academic workflows. Other best practices will emerge; this is a first short attempt to begin a wider conversation for our disciplines. It is worth keeping one principle in view throughout. These LLM tools are accelerants, for good and for ill. If your process is disciplined, they will accelerate your discipline. If your process is fragile, they will accelerate the fragility. The discipline has to come from you, the academic or the practitioner, because it will not come from the tool.

Robert C. Engen is Senior Lecturer in War Studies and Director of Wargaming at the Deakin University Centre for Future Defence & National Security in Canberra.

Claude was used as a drafting foil and to conduct scientific falsification of this article’s claims.

From the Board Game Shelf: World Order (Hegemonic Project Games)

Michael Hatherell

While we normally present a book review in this final section of Signal, this time around we have a review of a board game with strong geopolitical themes and educational intent. In 2023 the game studio Hegemony Project Games launched their first game, *Hegemony: Lead your Class to Victory*. In *Hegemony*, players operate as one of four different and asymmetric factions – the working class, the middle class, the capitalist class or the State. Each faction has its own policy goals and levers to pull, and the interaction between the factions led to a dynamic and policy relevant game. *Hegemony* has since become a very popular board game, ranking at time of writing as the number 41 game of all time in Board Game Geek's rankings.

While *Hegemony* was notable for its engaging theme and enjoyable gameplay, it also sought to make the most of the educational potential of the hobby. The game shipped with a 'concepts' booklet which drew on political science scholars to explain some of the key ideas in the game. This was arguably educational board gaming at its best – a game that was fun to play while also taking its political content seriously.

World Order is the follow up to *Hegemony*. This time around the designers have produced a game focusing on geopolitical competition between China, the U.S., the European Union and Russia. Gone is the domestic setting of their first game, with the focus this time being international competition between these four powers.

In *World Order*, players compete for influence across seven global regions. Each power in the game has regions that are strategically more important to them, providing greater access to apply military and diplomatic power but also more to lose if they do not keep up with their competitors. Players play cards from their hand to take actions loosely organised around the DIME elements of power (Diplomacy, Information, Military and Economy), though in *World Order*, the information element of power is replaced with domestic-focused actions.

Diplomatic actions allow players to either improve their relationship with a country in a region or to engage with the region as a whole. Economic actions allow players to trade with their allies and the other great powers or to apply foreign direct investment to one of their allies. Military actions include the building of military bases in allied countries and the movement of military forces around the map. Domestic actions focus on producing resources and building state power.

The resources of the game include energy, raw materials, food, services, consumer goods, military forces, diplomatic influence and money.

As a game *World Order* feels well-balanced and is enjoyable to play. Its rule-set is simpler than *Hegemony* and steps back a little from that game's asymmetric brilliance. But this also makes *World Order* easier to learn and easier to transition from playing one power to another. The game presents lot of interesting choices and trade-offs. The competition for influence, the development and use of national resources and the timely use of special abilities (such as 'Belt and Road Initiative' for China or 'EU Expansion' for Europe) present a genuinely interesting challenge with a good amount of interaction between players.

World Order as a Model and Education Tool

Games, like theories, are simplifications of a complex world. The value of the models these simplifications offer, as Volko Ruhnke argues, is that they:

> *…lay it all out on the table. They enable the rapid sharing of one or more experts' internal mental models by inviting students to climb inside and drive a dynamic model around themselves. There is no black-box computer code; instead, the players…learn and through play internalize the model's rules, can compare the game's simplifications to their own view of the real-world system under consideration, and thereby blend their own mental models with that of the experts to obtain a synthesis—a better model.*

As models of a complex world, games like *World Order* have to choose what to include and what to exclude. Games that try to include as much complexity as they can fit on a table tend to be harder to learn and complex to play. This can increase some of the realism of the game, but as a learning tool it will limit the audience to those who are interested and able to invest the time and attention into them. It may also make it harder for players to reflect on their own mental models of the world.

World Order achieves a quite good balance in this respect. It does not seek to include all aspects of international politics, but the ones it does model are presented in a quite elegant and thoughtful way. Influence is one of those key elements that *World Order* seeks to model. Unlike the dynamic battle for country influence in a game like *Twilight Struggle*, partner countries (called 'allies' in the game) are retained permanently in *World Order*. This means that once you build your relationship with them they are with you for the rest of the game. This led to an interesting discussion in one of my games about just how often in history countries do shift their allegiance between major powers and what drives them to do so.

These potential allies, including countries like Australia, Indonesia, India, South Africa and Brazil, do not have any direct agency in the game – they are reduced to tools to be used by the great powers. They do, however, have some preferences around who can ally with them and who can build military bases on their territory.

While allies are permanently kept, the battle for influence plays out at the regional level. This battle is split between 'permanent influence' which only provides a small up front reward but does not change once created, and 'temporary influence' which provides a larger up front reward but is easier for other powers to replace. This leads to simple but engaging choices about where to invest resources and attention. Focusing on one key region can help you compete there, but you may end up neglecting other regions and be surprised when one of your competitors starts investing resources there. In one of our games the player operating as China began investing in South America, catching the U.S. player off guard and changing the dynamic of the game. The choices about where to invest resources around the world and the decisions this forces for other powers is arguably the most interesting element of this model.

The four global powers play in relatively similar ways but they do have some unique characteristics. The card deck for each power is different, with subtle differences in what actions they have access to. The default resource generation of each nation is different, with Russia, for instance, able to produce a lot of oil and raw materials, the U.S. and Europe more services, and China more consumer goods.

The powers also have some distinct scoring bonuses and penalties – the U.S. needing to maintain a broad range of influence across all regions, Russia needing to maintain a strong military and China scoring for developing a foreign direct investment (FDI) network. As a model this pushes players to play their country in a relatively thematic way, with the Russian player encouraged to turn raw materials into military capability and the China player to gain partners in different countries where it can invest FDI. The U.S. player is pushed to be diplomatically engaged in as many regions as it can, leading to multi-front competition with the other three powers. These differences cause players to consider statecraft from the perspective of that power.

World Order and International Relations Theory

The *World Order* concept book comes with a lengthy explanation of different International Relations (IR) theories. These explanations are generally a very good summary, especially for players of the game who may not have studied IR theory before. Does the game truly encourage players to reflect on all of these theories? This is more of a mixed bag.

The feeling of playing the game is more realist in nature than any of the other theories. States compete for influence over the world map through investment in military presence, diplomatic influence and ally relationships. Choices about where to invest these resources are important and can shift the focus of the game. There is no option to go to war in *World Order*, but players do get the sense of an evolving balance of power throughout the game.

At a basic level geoeconomics and liberal ideas are also present, reminding the player of the importance of resources, trade and competitiveness. The battle for influence grinds to a halt without the right resources to continue fuelling diplomatic actions and military investments.

Interestingly, the game also models domestic development to some extent, allowing players to develop unique abilities and score points through increased infrastructure and prosperity. These elements offer an important domestic balance to the IR theory elements, but it is mostly in service to the broader geopolitical battle.

Subtle elements, like the U.S. and Europe not being military threats to each other, reminds the player that state identity and ideas do matter. Other liberal and constructivist ideas emerge in some of the card types available to the players. The differences between the four powers do make it easy to feel like you are playing the role of that particular power in the game, and in our plays of the game led to interesting discussion of how these states use their resource and power. The entries on each power in the concept book are a useful additional resource to support this.

It would be a lot to ask a board game like *World Order* to model all of the major IR theories, and it certainly does not attempt to do that. The blend of realist power politics, diplomatic influence, geoeconomic elements and domestic development provides enough of a basic model of world politics for players to reflect on their own mental models. The concept book then provides further exploration of theories and the major powers for players who want to go further.

Final Thoughts

While the release of this game in 2026 feels timely, the designers of *World Order* set the game in the year 2010. Placing the game at that time helps to distance it from more contemporary geopolitical controversies while still remaining quite relevant to world politics today.

This choice does lead to some odd moments given current world politics. Ukraine, for instance, will only allow for Russian military investment. The fact that the U.S. and Europe are not military competitors with each other made a lot of sense in 2010 (and right up until recent tension over Greenland and other issues). The goals of the U.S. in the game are interesting to reflect on compared to more recent U.S. policy - a 'Trump Expansion' would shift a number of victory goals and U.S. action cards.

Overall *World Order* is an engaging and thoughtful game about geopolitics. For students of international politics with an interest in board gaming it is a worthwhile investment. It could also be a useful learning tool for education settings where there is sufficient time to learn the rule set and engage in multiple games.

Dr Michael Hatherell is an Associate Professor in Strategic Studies and Deputy Director of the Centre for Future Defence and National Security. He specialises in strategic narrative, grand strategy, Indonesian politics and foreign policy. He is currently working on the challenges of effective statecraft and grand strategy in the contemporary information environment.

DEAKIN UNIVERSITY | DEAKIN CENTRE FOR FUTURE DEFENCE & NATIONAL SECURITY

Deakin University CRICOS Provider Code: 00113B