Shadow Risk Methods Annex

Overview

- This annex provides an overview of the crisis simulations and wargames run by CSIS in the fall of 2021 on a notional gray zone crisis between Taiwan and China.
- The annex includes the scenario overview and player movement sheets for both the control and treatment groups consistent with the conjoint experiment design.
- Last, the annex provides summary statistics for the t-tests performed on the interval data extracted from the wargame experiments.

As highlighted in the brief, the crisis simulation involved adapting a traditional tabletop exercise (TTX) to align with a conjoint experiment design. This novel design allowed CSIS to capture data and analyze it in a more rigorous manner consistent with the turn to treating wargames as social science. TTXs replicate a competitive decisionmaking process at the operational to strategic level, with participants assuming the roles of theater combatant commands or entire governments. TTXs emerged during the early Cold War as part of strategic analysis at RAND and parallel through more operational-level wargaming during the late 1970s at the Naval War College. The games are open and allow a larger degree of improvisation and creativity than legacy wargames developed to test ground and naval concepts during the nineteenth century and interwar period. This design assumes researchers can replicate a unique strategic crisis and draw inferences from expert choices. This assumption implies the choices of expert participants on any given day can be transferred to make predictions about the future, a proposition found inaccurate by decision science and social science research.

To overcome this limitation, the researchers employed a conjoint experiment design and treated choices during strategic competition as multidimensional. A conjoint design focuses less on a strategic choice as a single, discrete rational response to a crisis and more on choice as multidimensional and about competing preferences. This design requires running a larger number of TTXs and comparing preferences across control and treatment groups to determine if there is a statistically significant difference between the two.

CSIS conducted 20 TTXs between October 12 and November 2, 2021, that all involved military participants with at least 10 years of military experience and representing three different countries of the North Atlantic Treaty Organization using a modified table-top
exercise format optimized for conducting a conjoint experiment. Participants were asked to respond to the scenario, including:

- Identifying the overall escalation risk using an interval variable
- Characterizing Chinese escalation in terms of the likely motives (i.e., instrumental, suggestive) and mechanisms (i.e., vertical, horizontal)
- Assessing the most likely next move by China and overall escalation inherent in Chinese actions across the game
- Recommending a flexible deterrent option and military crisis response for the U.S. Indo-Pacific Command (INDOPACOM)

Small teams were presented with this scenario and asked to recommend response options. The teams consisted of military professionals that averaged over 10 years of experience with military and national security affairs. Players assumed the role of responding either as the United States, and INDOPACOM in particular, or China, and the People’s Liberation Army’s (PLA) Eastern Theater Command. This focus on theater-level, joint military decisionmaking allowed the researchers to see how military actors respond to gray zone crises and assess escalation risks associated with activities like freedom of navigation operations and challenging air defense identification zones—common practices in the current U.S.-China competition. Each side did not know how many rounds the crisis simulation would last or the exact objective of the other player to capture how uncertainty amplifies tensions in a crisis.

The players also were randomly assigned to different treatment groups, with half the players responding with a mix of response options that allowed them to defer crisis risk with long-term military options (e.g., foreign military sales, adjusting force posture, etc.) and treatments that had comparable options focused entirely on short-term crisis response (e.g., conduct a Freedom of Navigation Operations [FONOP], increase intelligence collection). Comparing the findings from the two treatment groups reveals key insights about escalation risks in gray zone competition. Every other aspect of the TTX was held constant to allow the researchers to assess whether the temporal differentiation had an impact on escalation dynamics. All options focused on the military instrument of power. Two separate teams served as the opposition so the researchers could gauge how the teams responded to a thinking adversary with comparable crisis response options. This design meant that researchers could isolate both initial escalation attitudes and intra-crisis escalation dynamics based both on perceived levels of risk and strategic choice.
Crisis Simulation Scenario Prompt

The language has been kept in its original form with minimal edits.

Fears of a major war between China and Taiwan have been on the rise for years. In 2022, after the 20th Party Congress, Xi Jinping gave a landmark speech on Taiwan signaling China could no longer wait to reintegrate the rogue province. The speech was a central part of a nationalist wave during the 2022 economic crisis that followed major property developers defaulting on their debt. Chinese military, economic, and diplomatic pressure on Taiwan increased further after the 2024 Taiwanese Presidential election with its slogans of “popular sovereignty” and the poor showing by the KMT. By 2026, there are weekly incursions into the island’s air defense identification zone by as many as thirty PLAAF aircraft at a time, including fighters, nuclear-capable bombers, and antisubmarine warfare platforms. The Eastern Theater Command routinely conducts snap exercises to signal its capabilities to isolate the island. Over the last three months, the activities included sending naval surface action groups off the eastern coast of Taiwan, simulated strikes by the strategic rocket forces and strategic support forces, and testing a WU-14, hypersonic glide vehicle. Analysts warn these moves signal China’s ability to deny foreign involvement if it invades Taiwan.

During this period of heightened military activity, there has been a corresponding increase in gray zone pressure on the Taiwanese islands of Kinmen located 10 kilometers off the coast of China. Though officially governed by the Republic of China (Taiwan), Beijing claims Kinmen as part of its Fujian Province. The island has become a popular tourist destination for mainland tourists since the end of the pandemic in 2022. With tourist revenue also came calls for closer relations with the mainland. China responded by adding gas and electric lines to the undersea pipelines for fresh water it built in 2018. At a ribbon cutting ceremony for the gas line in early 2027, a local Kinmen leader called for turning the island into a “peace experiment zone” and unifying it with the Chinese city of Xiamen.

The call leads to protests in Taipei but little fanfare in Kinmen where residents are economically dependent on mainland China and more supportive of the KMT. During the protests, the media receives leaked emails from the account of a member of the Democratic People’s Party of Taiwan (DPP) stating that if Taiwan declared independence, they should specify that Kinmen and Matsu are not part of the country. The DPP officials accuse China of fabricating the emails. The same week sees a flurry of social media commentary about how the United States destroyed the prospects of a unified China in 1958 (the Second
Taiwan Strait Crisis) by supplying Taiwanese troops with artillery and missiles to fend off a Communist invasion of Kinmen. Chinese “wolf warrior” diplomats and netizens circulate the short videos and falsely claim the U.S. sent in secret advisors who committed abuses against locals.

Analysts contend increased Chinese cyber and misinformation activity is designed to counter changing international attitudes toward defending Taiwan. Since 2021, opinion polls show a slim majority of the U.S. public consistently favors the use of U.S. troops to defend Taiwan if China invades. In Japan, polling data shows a majority favor government measures to promote peace and stability in the Taiwan Strait.

The Kinmen crisis deepens when the gas pipeline explodes, causing damage to a nearby Chinese fishing vessel. Without evidence, Chinese state media claims the explosion is the result of an attack by Taiwanese separatists. China deploys a PLAN naval task group to isolate the area and rescue the boat while dispatching construction crews to survey pipeline damage on Kinmen. As the construction crews arrive, they appear to be predominantly military-aged males. There is a tense standoff at the construction site, but Kinmen police refuse to get involved.

The next day, China announces it is further expanding its air identification zone to cover all of Kinmen. The announcement coincides with a large military exercise in which China fires missiles into the East China Sea and over 100 PLAAF aircraft simulate attack runs to the north and south of Taiwan, including entering the Japanese air-defense identification zone. There are unconfirmed reports of heavy electronic magnetic interference in the area and dazzling (i.e., temporary blinding) effects on at least two U.S. imagery satellites. A private cyber security firm warns that Chinese malware may be activated on port infrastructure across the region to further limit any foreign involvement.

Japan deploys a naval task group organized around an Izumo-class multi-purpose destroyer/carry with twelve F-35B that was conducting exercises with Australian and U.S. naval groups in the Philippine Sea. Despite Taiwan’s diplomatic isolation, multiple countries warn the crisis risks sparking a wider confrontation. World stock markets fall 10 percent as funds shift to U.S. bonds and gold prices surge. Taiwan requests assistance in accompanying its flights and maritime traffic to the island in a test of the Chinese ADIZ.

**PLAYER PACKETS**

*U.S. (A) Objective = demonstrate U.S. resolve and contain the crisis*

*Strategic Forces.* At this time, there is no additional activity by SSF nuclear forces. Mobile launchers are at designated bases and there is not activity to fuel liquid-fueled ICBMs. PLAN Jang class SLBMs are on routine patrol in the South China Sea. No additional movement of nuclear-capable H-6 bombers.
**Conventional Forces.** There has been increased activity in the Eastern Theater Command, including increased readiness at SSF conventional formations and additional PLAAF activity. It is unclear whether there is increased naval activity in the Southern and Northern Theater Commands given the high volume of recent exercises. There is no major change in the status of PLAGF.

**Intelligence Estimate.** The J2 sees a range of Chinese options to execute war control within the next 72 hours including the following.

- Maritime militia surrounds Kinmen
- Air and missile exercise (firepower strike demonstration) in the East China Sea and southwest of Taiwan (isolate the island)
- Sea guard line (deploy PLAN to limit a blockade potential)
- Cyber escalation targeting U.S. military equipment and/or critical infrastructure
- Space escalation (dazzle and/or soft kill multiple intelligence satellites while avoiding nuclear detection/C3 systems)
- Strike Waves: gradually increase conventional SSF missile attacks (graduated pressure) to include if the U.S. doesn’t back down striking Guam and U.S. bases in Japan

**Do you think the Chinese are likely to escalate?** (1 almost no chance, 6 almost certain):

**What dimension will it take?** (circle one): vertical, horizontal

**If the Chinese do escalate, what is the motive?** (circle one): instrumental (i.e., deliberate effort to pressure the U.S. to back down) or suggestive (signal)

**Select your recommended crisis response option (circle) and rate the escalation risk of your recommended option** (1 low risk of escalating, 6 high):

<table>
<thead>
<tr>
<th>RESPONSE OPTIONS</th>
<th>ESCALATION RISK</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Deploy additional intelligence assets</td>
<td></td>
</tr>
<tr>
<td>2. Deploy Rapid Raptor + Additional Air Packages to signal strike and increase CAP capability</td>
<td></td>
</tr>
<tr>
<td>3. Conduct a FONOP in the Taiwan Strait</td>
<td></td>
</tr>
<tr>
<td>4. Escort Taiwanese commercial air and maritime traffic to Kinmen</td>
<td></td>
</tr>
<tr>
<td>5. Deploy to blockade China</td>
<td></td>
</tr>
<tr>
<td>6. Conduct a limited strike consistent with JAM-GC/ASB “blinding strike”</td>
<td></td>
</tr>
</tbody>
</table>

**STOP**
How escalatory was China’s action? (1 low, 6 high):

Recommended response (put a check next to the response above). Risk assessed:

**U.S. (B) Objective = demonstrate U.S. resolve and contain the crisis**

*Strategic Forces.* At this time, there is no additional activity by SSF nuclear forces. Mobile launchers are at designated bases and there is not activity to fuel liquid-fueled ICBMs. PLAN Jang class SLEMs are on routine patrol in the South China Sea. No additional movement of nuclear-capable H-6 bombers.

*Conventional Forces.* There has been increased activity in the Eastern Theater Command including increased readiness at SSF conventional formations and additional PLAAF activity. It is unclear whether there is increased naval activity in the Southern and Northern Theater Commands given the high volume of recent exercises. There is no major change in the status of PLAGF.

*Intelligence Estimate.* The J2 sees a range of Chinese options to execute war control within the next 72 hours including the following.

- Maritime militia surrounds Kinmen
- Air and missile exercise (firepower strike demonstration) in the East China Sea and Southwest of Taiwan (isolate the island)
- Sea guard line (deploy PLAN to limit a blockade potential)
- Cyber escalation targeting U.S. military equipment and/or critical infrastructure
- Space escalation (dazzle and/or soft kill multiple intelligence satellites while avoiding nuclear detection/C3 systems)
- Strike Waves: gradually increase conventional SSF missile attacks (graduated pressure) to include if the U.S. doesn’t back down striking Guam and U.S. bases in Japan

Do you think the Chinese are likely to escalate? (1 almost no chance, 6 almost certain):

What dimension will it take? (circle one): vertical, horizontal

If the Chinese do escalate, what is the motive? (circle one): instrumental (i.e., deliberate effort to pressure the United States to back down) or suggestive (signal)

Select your recommended crisis response option (circle) and rate the escalation risk of your recommended option (1 low risk of escalating, 6 high):
RESPONSE OPTIONS

1. Announce an increase in partner exercises (build partner interoperability, capability in the long term)

2. Announce an increase in foreign military sales to increase Taiwan and allied nations’ ability to defend themselves in the long term

3. Announce an increase in force posture in INDOPACOM over the long term

4. Escort Taiwanese commercial air and maritime traffic to Kinmen

5. Deploy to blockade China

6. Conduct a limited strike consistent with JAM-GC/ASB “blinding strike”

STOP

How escalatory was China's action? (1 low, 6 high):

Recommended response (put a check next to the response above). Risk assessed:

CHINA Objective = challenge Taiwan’s claims to Kinmen Island

Enemy Forces. At this time, there are no major changes in U.S. strategic forces (nuclear triad). INDOPACOM appears to be sending additional air and maritime assets to the region based on activity in Guam and San Diego. Unknown number of special operators likely in Taiwan training their 101st Amphibious Reconnaissance Battalion (Frogmen at forward operating bases on islands like Kinmen, Matsu, and Penghu). The JSDF deployed a strike group organized around the Izumo-class carrier to the East China Sea. Increased activity seen in their air and maritime forces indicative of territorial defense. Increased activity in Taiwanese air, ground, and maritime formations, including hardening sites on Kinmen by the 101st Amphibious Reconnaissance Battalion. Navy is currently concentrated between Penghu and Taiwan and appears to be organizing into a naval task force organized around a destroyer, two frigates, and at least two amphibious ships (LCC-1, LST). Missile patrol boats and subs unaccounted for at this time. Maritime surveillance aircraft are up (surge), including 4 x P-3 and 2 x EP-3 screening west of Penghu in front of the fleet.

Intelligence Estimate. The ETC J2 sees a range of U.S. options to support the Taiwanese in the next 72 hours including the following.

- Deploy additional intelligence assets
- Announce future increased partner exercises to build interoperability/capability in the long term
- Deploy rapid raptor and other airstrike packages to signal while generating future strike options
- Announce an increase in foreign military sales to increase Taiwan and allied nations’ ability to defend themselves in the long term
- Conduct a FONOP in the Taiwan Strait
- Announce an increase in force posture in INDOPACOM over the long term
• Escort Taiwanese commercial air/maritime vessels to Kinmen
• Blockage Chinese maritime traffic
• Conduct limited strikes consistent with JAM-GC/ASG “blinding strike”

Do you think the U.S. is likely to escalate? (1 almost no chance, 6 almost certain):

What dimension will it take? (circle one): vertical, horizontal

If the U.S. does escalate, what is the motive? (circle one): instrumental (i.e., deliberate effort to pressure the U.S. to back down) or suggestive (signal)

Select your recommended crisis response option (circle) and rate the escalation risk of your recommended option (1 low risk of escalating, 6 high):

<table>
<thead>
<tr>
<th>RESPONSE OPTIONS</th>
<th>ESCALATION RISK (1 low, 6 high)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Maritime militia surrounds Kinmen</td>
<td></td>
</tr>
<tr>
<td>2. Air and missile exercise (firepower strike demonstration) in the East China Sea and southwest of Taiwan (isolate the island)</td>
<td></td>
</tr>
<tr>
<td>3. Sea guard line (deploy PLAN to limit a blockade potential)</td>
<td></td>
</tr>
<tr>
<td>4. Cyber escalation targeting U.S. military equipment and/or critical infrastructure</td>
<td></td>
</tr>
<tr>
<td>5. Space escalation (dazzle and/or soft kill multiple intelligence satellites while avoiding nuclear detection/C3 systems)</td>
<td></td>
</tr>
<tr>
<td>6. Strike waves: gradually increase conventional SSF missile attacks (graduated pressure) to include if the U.S. doesn’t back down striking Guam and U.S. bases in Japan</td>
<td></td>
</tr>
</tbody>
</table>

STOP

How escalatory was the U.S. action? (1 low, 6 high):

Recommended response (put a check next to the response above). Risk assessed:

Statistical Analysis
The TTX design allowed CSIS researchers to test two hypotheses associated with how the time horizon of response options affects escalation dynamics. Since the total number of games was less than 50, the researchers used t-tests to compare the two treatments. The significant findings are reported for (p<.1) and (p<.05).

The first hypothesis is that the presence of long-term crisis response options reduces escalation magnitude. The null hypothesis is that there should be no observable change in escalation magnitude comparing comparable long-term and short-term response options across the treatment groups. Based on the data captured in the TTX, the null hypothesis is rejected.
A paired sample t-test was conducted to determine the effect on long-term vs. short-term response options on the magnitude of escalation implied by the response option. There was a statistically significant difference between the recommended response option in which the choices included long-term options (M = 1.9, \( \text{VAR} = .77 \)) and recommended response options in which there were only comparable short-term options (M = 3.3, \( \text{VAR} = .68 \)).

The results also hold when shifting from the preferred military crisis response option to player’s evaluation of the escalation risk implied by the option. During each round, U.S. players also rated the escalation risk for each military response option. Here a paired sample t-test was similarly conducted to determine the effect on long-term versus short-term response options on the player’s assessment of the implied escalation risk. There was a statistically significant difference between the assessed escalation risk in the recommended response option in which the choices included long-term options (M = 2.2, \( \text{VAR} = .62 \)) and in which there were only comparable short-term options (M = 2.9, \( \text{VAR} = 1.43 \)).

While significant, these results were not as strong as the difference between the mean response measuring the preferred crisis response option.

The results imply that players tended to avoid confrontation in the near term by deferring risk. In discussions, players suggested they didn’t want to be the first to take aggressive action as it would complicate the position of allies. That is, there was a concern that U.S. military actions above an established threshold of moving air and intelligence assets and conducting FONOPs would make allies less likely to work alongside the United States if the crisis escalated. Other observers noted that the first mover would likely be punished in the court of public opinion. There seemed to be broad agreement that gray zones were an accepted sub-threshold activity or what Herman Kahn earlier theorized as sub-crisis maneuvering in his escalation ladder. This idea is also captured in emerging doctrine in multiple great powers, including the U.S. notion of the competition continuum and competition beneath armed conflict and Russian assessments linked to New Generation Warfare. China’s ongoing actions and writing on war control similarly illustrate an interest in shaping and taking actions that preserve flexibility and initiative.

The second hypothesis was that when participants respond to gray zone activity with long-term crisis response options in the first round, they are likely to move up the escalation ladder in round two (i.e., future interactions). The null hypothesis is that there should be no observable change in the magnitude of different responses between round 1 and round 2 between the treatments by U.S. players. Based on the data captured in the TTX, the null hypothesis is rejected.

A paired sample t-test was conducted to determine the effect of different treatment groups (long-term options vs. only short-term options) on the change in crisis response options during the course of the war game. There was a statistically significant difference between recommended response options in the second round where players could choose long-term options (M = 1.8, \( \text{VAR} = 1.29 \)) and response options where players were limited to comparable short-term crisis options only (M = -0.7, \( \text{VAR} = .645 \)).

The results hold when comparing player-preferred U.S. response options in round two relative to the difference between the preferred U.S. response in round 1 and the actual Chinese response in round one. This equation draws out how much more escalation
pressure the U.S. player selected in round two relative to their position at the end of round one when they could see their strategy relative to the Chinese strategy and gain more information. Using a paired sample t-test shows a statistically significant difference between treatments in the second round along these lines. Whereas round one response options where players had long-term options exhibited lower escalation levels in round one, they jump up the escalation ladder in round 2 \( (M = 3.3, \text{VAR} = 1.12) \) compared to comparable short-term options \( (M = .8, \text{VAR} = 5.73) \).

The results imply that when players reach for the shadow of the future to delay confrontation, they do not escape risk. They only defer difficult choices if the other side does not back down, and a crisis continues to escalate. The results imply that escalation risks are not lost but deferred. In gray zone campaigns, states may seek to avoid a near-term confrontation by signaling future actions that could alter the balance of capabilities and resolve in the future. This posture inverts the classic commitment trap in international relations whereby states take risky actions in the present to avoid more costly actions in the future. In fact, time was the operative factor shaping response dynamics. The researchers ran a test to see if the assessment of either Chinese escalation pathways in terms of their motives (i.e., instrumental or suggestive) or mechanisms (i.e., vertical or horizontal) accounted for variation in escalation levels. They did not.\textsuperscript{32}
Endnotes


3 Martin van Creveld, Wargames from Gladiators to Gigabytes (Cambridge: Cambridge University Press, April 2013).


6 Conversely, this focus limited our ability to assess a broader range of diplomatic and economic options.

7 Detailed movement sheets are available in the methods annex accompanying this brief.

8 Because the sample is only 20 games (10 for each treatment), the analysis uses a t-test as opposed to a -test to analyze differences between the treatments across the variables and treats response and escalation variables as ordinal.

9 T Stat 3.77, p < .002 (one tailed).

10 T Stat 1.48, p < .086 (one tailed).

11 T Stat -2.91, p < 0.001 (two tailed).

12 The test involved a contingency table and analysis of adjusted standardized residuals and chi-square tests of independence comparing the variables for motive and above in treatment escalation risk assessments and mechanism and the same test treating each as a categorical variable.