



The Hague Centre
for Strategic Studies

Know Thy Enemy, Know Thyself

Simulations of Political Leaders' Crisis Behaviour after a Ceasefire in Ukraine

Jesse Kommandeur, Chiara Barontini and Tim Sweijts

October 2026





Know Thy Enemy, Know Thyself

Simulations of Political Leaders' Crisis Behaviour after a Ceasefire in Ukraine

October 2026

Authors:

Jesse Kommandeur, Chiara Barontini and Tim Sweijts

Contributors:

Alessandro Dell'Orto, Clarice van der Paardt,
Nicole Eichstaedt and Veerle Tielemans

Quality Assurance:

Frank Bekkers

The cover image was AI-generated.

The research conducted in preparation of this publication has been made possible in part by financial support from the Research Division of NATO's Defence College.

This publication is [AI-supported](#). The authors take full responsibility for the integrity of the work.

© *The Hague* Centre for Strategic Studies. All rights reserved. No part of this report may be reproduced and/or published in any form by print, photo print, microfilm or any other means without prior written permission from HCSS. All images are subject to the licenses of their respective owners.

Table of Contents

	Executive Summary	IV
1.	Introduction	1
1.1	Ceasefire breakdown scenarios	1
1.2	The ARENA simulation environment	2
1.3	This paper	2
2.	Setup, Method and Data	3
2.1	An agent-based simulation	3
2.2	The crisis simulation	4
2.3	Building psychological profiles	6
2.4	Evaluating psychological profiles	9
3.	Seven Leaders, Seven Signatures	11
3.1	Vladimir Putin	12
3.2	Volodymyr Zelensky	13
3.3	Donald Trump	14
3.4	Emmanuel Macron	15
3.5	Friedrich Merz	15
3.6	Keir Starmer	16
3.7	Mark Rutte	17
3.8	The Collective	17
4.	Pathways of Escalation	18
4.1	How the crises unfold	18
4.2	What drives the outcome	20
5.	Conclusion	23
5.1	Key findings	23
5.2	Recommendations	24
6.	Annex: Leaders' Profiles	27
6.1	Vladimir Putin	27
6.2	Volodymyr Zelensky	28
6.3	Donald Trump	29
6.4	Emmanuel Macron	30
6.5	Friedrich Merz	31
6.6	Keir Starmer	32
6.7	Mark Rutte	33

Executive Summary

At some point in the future, the Russia-Ukraine war will be followed by a ceasefire, agreed to by the belligerents. It is in the interest of the Coalition of the Willing that supports Ukraine to uphold such a ceasefire and prevent escalation. Ceasefires may be the runup to more comprehensive peace agreements, but they are also fragile. Historical evidence shows that they are prone to breaking down quickly, followed by escalation into crisis and potentially renewed war.

Political leaders play a crucial role in determining the outcomes of crises. Yet the psychological make-up of these leaders and the reasoning that shapes their decisions remain by and large a black box. This research project seeks to open the black box of leadership decision-making using ARENA, for Agentic Reasoning for Escalation, Negotiation and Analysis, a crisis simulation platform developed by the HCSS Datalab.

In ARENA, psychological profiles were constructed for the political leaders of Russia, Ukraine, the US, France, Germany, the United Kingdom and NATO to simulate their decision-making strategies following the collapse of a hypothetical ceasefire in Ukraine. Each leader is represented by an artificial agent consisting of a five-layer psychological profile: (1) a personal biography; (2) a personality profile combining a set of standard psychological test scores and leadership decision making traits; (3) an operational code with beliefs about strategy; (4) an assessment of how the leader reads rivals; and (5) the leader's risk appetite that might shift as the crisis unfolds.

Subsequently, various crisis scenarios were simulated using large language models. Each simulation begins with the collapse of a Russia-Ukraine ceasefire in one of four ways: (1) the collapse of peace talks in Geneva; (2) a Russian land grab in the Baltic; (3) a Russian probe testing Western resolve; or (4) a Ukrainian cross-border strike into Russia. These crisis openings were followed by one of four different unrelated shocks: (1) a Chinese blockade of Taiwan; (2) a cyberattack on Europe's power grid; (3) a closure of the Strait of Hormuz; or (4) a pandemic. Combining four possible crisis triggers with four possible crisis shocks generated sixteen distinct scenarios. Each scenario was replayed ten times, for 160 runs in total. An analysis of the leaders' decisions across these runs, and the private reasoning behind them, yields the following **key findings**:

1. **Mid-crisis is the moment that matters.** The trigger of a crisis does not determine its course. Crises evolve and (unrelated) incidents typically have outsized effects on the crisis trajectory. Crisis machinery tends to be built for preventing a crisis or responding to the first forty-eight hours, not for what comes next, let alone for dealing with compounding effects.
2. **Coalition cohesion is crucial.** Russia reaches a decisive result across many runs. If we dig deeper, this tends to be the result of division within Ukraine's supporting alliance. This underscores the need for coalition cohesion.

3. **Character over circumstance.** Political leaders respond very differently when presented with the same crisis, showcasing how personality and operational codes matter. Putin pushes, Zelensky adapts, European leaders hold firm but proceed cautiously, while Trump talks tough but tends to hold back. Among our artificial agents at least, each leader's crisis decision-making strategy reflects consistent patterns of behaviour.
4. **Europe and the United States pull in opposite directions.** Europe moves almost as one. Trump pulls the other way, dragging a coalition that looks united quietly downward from within. That seam is the one Russia aims at.
5. **Only shared risks re-open the door to negotiations.** A pandemic is the only mid-crisis shock that is followed by talks. In the scenarios, asymmetric pressure tends to drive escalation, while symmetric pressure opens negotiation.
6. **Taiwan is Europe's problem too.** A Chinese blockade of Taiwan directly affects the situation on the European continent. No scenario containing a Chinese blockade of Taiwan reaches peace, and there is no Chinese Wall between the two theatres. If a crisis over Taiwan erupts at the same time as a ceasefire over Ukraine breaks down, it immediately becomes a European security emergency.

The behavioural patterns we find in our simulations point to a sequence of measures that NATO and the Coalition of the Willing supporting Ukraine can start implementing now. This leads to the following **recommendations** – which have broader applicability for policymakers preparing for crises:

1. **Treat the post-ceasefire period as a planning problem in its own right.** Western preparation typically concentrates on preventing the outbreak of a crisis and reacting in the first hours. Yet our analysis consistently confirms the emergence of risks in the management of a crisis already under way. The Coalition of the Willing should therefore also plan for the fragile weeks after the collapse of a ceasefire in addition to seeking to prevent its collapse.
2. **Study the decision-makers themselves.** Leaders' personality and decision-making style shape a crisis as much as the circumstances they face. In our simulations, each leader behaves consistently enough across the different scenarios for patterns to emerge. The personal dispositions of the leaders of adversaries and allies alike should therefore be a subject of systematic analysis. Simulation of political crisis dynamics should be part and parcel of the coalition's crisis planning.
3. **Rehearse both steps down and up the escalation ladder.** Climbing the ladder takes no coordination, while coming back down takes a collective decision that the Western Coalition rarely reaches in time. In addition to rehearsing escalatory threats, the Coalition of the Willing should also prepare and practice for how to de-escalate.
4. **Managing cohesion is paramount.** Cracks in the coalition stand in the way of effective crisis management. A single powerful member, the US, can exert a powerful drag on reluctant allies. This may appear as unity at first sight, but in effect generates division.
5. **Wargame compound crises.** Coalition exercises should consider the possibility of a second emergency midway through a crisis scenario because the collision of two crises typically fuels further escalation.

6. **Treat Taiwan contingency as a European security crisis.** A China-Taiwan contingency is the most dangerous shock in the campaign for the simple reason that it draws American power toward the Pacific while leaving Europe exposed. European allies should build the capacity to hold the line against Russia when the US is busy elsewhere.
7. **Build a European channel to Moscow.** For three of the four crisis triggers in our simulations, a live communication channel in the real world could defuse further escalation. European leaders should not assume that the US president will de-escalate or even get involved. They should therefore set up their own high-level channels to the Kremlin and military commanders to clarify intent before and after an incident escalates. More fundamentally, this shows that crisis off-ramps must be built deliberately.

1. Introduction

1.1. Ceasefire breakdown scenarios

Any ceasefire in the Russia-Ukraine war, should the parties agree to one, is likely to be fragile. Of the negotiated settlements reached between 1975 and 2018, almost half broke down.¹ The Coalition of the Willing supporting Ukraine has announced that it intends to guarantee any settlement with troops stationed on Ukrainian soil. It has also declared that the reassurance force would be activated the moment a ceasefire takes hold.² This commitment binds Coalition of the Willing members' forces to the survival of Ukraine as an independent state. It does so at a juncture where a settlement is most easily undone, namely in the weeks that follow the cessation of hostilities when the terms have not yet solidified.³

Analysis of the war has concentrated for the most part on its origins and its conduct. HCSS's *Blinded by Bias* attributed the West's pre-2022 failure of foresight to its misreading of Russia. Policymakers discounted clear warnings because of a set of cognitive biases as they assumed that Putin shared their own cost-benefit calculus.⁴ Work on the war itself has focused on the evolution of its character, analysing development along a frontline that barely moves and the transformational impact of the use of drones and deep strikes.⁵ How the war might end has drawn considerably less attention, and what does exist tends to concentrate on the terms of a settlement rather than on its durability.⁶ There is even less attention to the period following a ceasefire; this is unfortunate because this is precisely the phase that warrants the most preparation. Crisis anticipation is best undertaken before violence erupts, rather than after it.⁷ After all, war-termination research has shown that settlements that freeze a conflict without resolving its underlying causes are prone to relapse.⁸

¹ Tim Sweijs and Mattia Bertolini, *How Wars End: War Terminations; Insights for the Russia-Ukraine War* (Hague Centre for Strategic Studies, 2022).

² *Determined to Accelerate Progress to Peace: Statement of the Co-Chairs of the Coalition of the Willing*, Élysée, 13 July 2026.

³ Tim Sweijs, *Een MH17-scenario kan zich opnieuw voltrekken: Daarom moet Europa grenzen stellen*, Trouw, 16 July 2026.

⁴ Tim Sweijs et al., *Blinded by Bias: Western Policymakers and Their Perceptions of Russia before 24 February 2022* (Hague Centre for Strategic Studies, 2025).

⁵ Tetiana Trach et al., *Russian Offensive Campaign Assessment, July 20, 2026* (Institute for the Study of War, 2026); George Barros and Kateryna Stepanenko, *Ukraine's Intermediate-Range Strike Campaign and New Mechanized Attacks Herald the Start of a New Phase of the War* (Institute for the Study of War, 2026); Michael Kofman, *Ukraine's War of Endurance: The Fight for Advantage in the Conflict's Fifth Year*, Foreign Affairs, 16 February 2026; Franz-Stefan Gady, *When Defense Becomes Destruction: Austria-Hungary's Mistake and Ukraine's Risk*, War on the Rocks, 2 December 2025.

⁶ Michael Kimmage and Hanna Notte, *The Perils of Premature Negotiation over Ukraine*, Wall Street Journal, 2 March 2023; Alexander Gabuev, *How Europe Can Get Putin's Attention*, Foreign Affairs, 7 July 2026.

⁷ Tim Sweijs, Artur Usanov and Rik Rutten, *Back to the Brink: Escalation and Interstate Crisis* (Hague Centre for Strategic Studies, 2016); Glenn H. Snyder and Paul Diesing, *Conflict Among Nations: Bargaining, Decision Making, and System Structure in International Crises* (Princeton University Press, 1977); Richard Ned Lebow, *Between Peace and War: The Nature of International Crisis* (Johns Hopkins University Press, 1981).

⁸ Virginia Page Fortna, *Scraps of Paper? Agreements and the Durability of Peace*, *International Organization* 57, no. 2 (2003): 337–72, <https://doi.org/10.1017/S0020818303572046>; Fred Charles Iklé, *Every War Must End*, A Columbia Paperback 116 (Columbia University Press, 1971); Hein E. Goemans, *War and Punishment: The Causes of War Termination and the First World War* (Princeton University Press, 2000).

Factors such as the clarity of the settlement terms, the military balance of forces and the territorial demarcation between the belligerents affect the likelihood of a ceasefire breakdown. Such a breakdown may either end in a contained standoff or, alternatively, in a slide towards general war. Whether it does so depends on how the key political leaders interpret and respond to the situation. It is at this moment that the individual judgement of political leaders carries the greatest weight and faces the fewest constraints. During crises, deliberation is compressed into hours or even minutes.⁹ In case of a hypothetical ceasefire breakdown, the decision to respond with force or with restraint rests with Vladimir Putin, Volodymyr Zelensky, Donald Trump and the European heads of state whose troops would stand as guarantors of the ceasefire. Anticipating how these leaders will behave is therefore essential to close a window of instability before it opens.

1.2. The ARENA simulation environment

So, the question we asked ourselves was: *what are the potential crisis trajectories after a Russia-Ukraine ceasefire breakdown, based on the crisis decision-making propensities of key political leaders?* To answer this question, we turned to ARENA, for Agentic Reasoning for Escalation, Negotiation and Analysis, a crisis simulation platform developed by the HCSS Datalab.¹⁰ In ARENA, political leaders are represented by AI agents, constructed from five-layer psychological profiles. ARENA replays the collapse of a Russia-Ukraine ceasefire across sixteen scenarios representing different versions of the crisis trajectory after the ceasefire breakdown. Seven key leaders sit at the table, of Russia, Ukraine, the US, France, the UK, Germany and NATO. Each scenario begins with one of four ways in which the ceasefire collapses, combined with one of four external shocks that land mid-crisis. Each scenario is replayed ten times, for a total of 160 simulation runs, so that recurring patterns can be distinguished from one-off accidents.

1.3. This paper

This paper basically does two things: it provides an overview of the ARENA simulation environment set-up; and it presents our overall findings from the simulation runs. It does not delve deeply into the internal workings of the simulation model, nor does it present an in-depth analysis of the actual simulation runs. This is done elsewhere, in two forthcoming, more technical papers.¹¹

This paper proceeds as follows. Chapter 2 sets out the design of the simulation including the sixteen versions of the crisis, and the way the psychological profiles of the seven leaders are constructed. Chapter 3 turns to the leaders themselves: which agents drive escalation, which restrain it, under what circumstances the Western coalition holds together or splits up. Chapter 4 examines the pathways of escalation, the crisis outcomes they lead to and the factors that drive towards escalation or de-escalation. The conclusion draws the findings together and considers the policy implications for Western decision makers on how to curtail escalation once a ceasefire breaks down.

⁹ Wilkenfeld, Jonathan, Kyle Beardsley, and Patrick James, eds. *A Century of International Crisis Behavior*. University of Michigan Press, 2026.

¹⁰ The Hague Centre for Strategic Studies, *ARENA* Simulation platform (HCSS, 2026).

¹¹ Jesse Kommandeur, Alessandro Dell'Orto and Tim Sweijs, *Psychologically Grounded Generative Agents for Post-Ceasefire Conflict Escalation: A Multi-Agent Simulation Study of Political Leader Decision-Making* (forthcoming); Jesse Kommandeur, Alessandro Dell'Orto and Tim Sweijs, *A Name Is All You Need: Large Language Models Reproduce Expert Psychological Ratings of World Leaders with Minimal Evidence* (forthcoming).

2. Setup, Method and Data

ARENA allows us to explore different crisis trajectories that result from the decisions made by the (simulated) political leaders involved in the crisis. This chapter explains the setup of the experiment and the method and data used to run it. The crisis scenarios including crisis triggers, crisis shocks and escalation ladders are described; after which the psychological profiles from which the artificial agents representing the political leaders were built are portrayed.

2.1. An agent-based simulation

From the map exercises of the nineteenth-century Prussian staff to the tabletop wargames that NATO allies conduct today: crisis and wargames are established instruments used by military staffs and governments to prepare for the future. Traditional role-playing wargames are subject to three limitations. Firstly, expert players tend to be few and far between. Secondly, the adversary is played by one or more of the participants, even if there is little knowledge about the adversary's internal deliberations. Thirdly, each game typically gets simulated only once, which complicates generalisability of its results.

ARENA seeks to address these limitations by modelling the leaders themselves as AI agents based on a multipronged psychological analysis and by replaying multiple crises many times over. ARENA is inspired by existing generative agent-based modelling approaches, in which a large language model is directed to reason and act according to a given description of a person (a 'persona'), rather than to follow a predetermined script.¹²

As such, ARENA expands three existing strands of work. The first strand demonstrates that an agent can hold a character. It turns out that models can convincingly simulate human respondents¹³ and sustain roles across a series of interaction sequences,¹⁴ even if role-play becomes less convincing if the character the agent impersonates grows more villainous. One recent benchmark attributes this to safety training of LLMs by frontier model companies that resists malevolent persona.¹⁵

The second strand explores the interaction of single agents, each representing a single actor, in a many-actor simulation setting. A seminal paper placed twenty-five agents within a simulated town and found that they formed memories, plans and relationships which held over

¹² Alexander Sasha Vezhnevets et al., 'Generative Agent-Based Modeling with Actions Grounded in Physical, Social, or Digital Space Using Concordia', arXiv preprint arXiv:2312.03664 (2023).

¹³ Lisa P. Argyle et al., 'Out of One, Many: Using Language Models to Simulate Human Samples', *Political Analysis* 31, no. 3 (2023): 337–51.

¹⁴ Murray Shanahan, Kyle McDonell and Laria Reynolds, 'Role Play with Large Language Models', *Nature* 623 (2023): 493–98.

¹⁵ Zihao Yi et al., 'Too Good to Be Bad: On the Failure of LLMs to Role-Play Villains', arXiv preprint arXiv:2511.04962 (2025).

two simulated days.¹⁶ Subsequent work finds that models drift from established norms under sustained 'multicultural' pressure.¹⁷

The third strand transposes the many-agent setting into a wargame. Artificial agents playing a US-China crisis held on to the role competently, but escalated to a greater degree than human experts did and proved sensitive to minor changes in wording.¹⁸ In a similar experiment, five frontier models were inclined to escalate extending as far as nuclear use.¹⁹ Yet none of the models ever chose accommodation or withdrawal, even under acute pressure; at most they reduced the level of violence. A third study found that these models, much like humans, accept more risk when facing losses than when facing gains, and that military scenarios generate far larger framing effects than civilian ones.²⁰ How a crisis is framed may therefore shape an agent's risk appetite as much as the facts of the crisis do.

ARENA builds on this work but differs in three respects. *First*, its agents are not generic decision-makers but carry psychological profiles of real-world leaders that confront each other within the structure of an actual conflict. *Second*, the escalation ladder that leaders pick from explicitly includes de-escalatory and accommodating options at its foot. *Third*, each scenario is replayed ten times, allowing systematic tendencies to be separated from one-off outcomes. These features do not by themselves eliminate the escalation bias documented in earlier studies, and ARENA's results should be read with it in mind. They do, however, allow us to test whether leader-specific profiles pull agents away from that default. As chapters 3 and 4 show, they partly do: several agents de-escalate consistently, ten simulations end in a negotiated peace, and no run crosses the strategic nuclear threshold. Whether this restraint stems from the profiles themselves or from other design choices, such as the explicit marking of an all-out nuclear exchange as unsurvivable, remains to be established through comparison with profile-free baseline simulations.

2.2. The crisis simulation

Our simulations start with a Russia-Ukraine ceasefire that has just collapsed. From this uniform point of departure, variation is introduced through four crisis triggers and four crisis shocks, leading to sixteen distinct scenarios, see Figure 1. Each scenario pairs a trigger event that (re)opens the crisis with a shock, an unrelated emergency that strikes midway through the crisis trajectory whilst the leaders are already under strain. Each shock arrives at the same fixed point in every scenario, so that its effect can be compared across scenarios. The crisis triggers and shocks were developed in consultation with HCSS subject-matter experts, to ensure their plausibility and relevance. Each scenario is replayed ten times, amounting to 160 simulations in total, a number sufficient to distinguish recurring patterns from one-off accidents.

¹⁶ Joon Sung Park et al., 'Generative Agents: Interactive Simulacra of Human Behavior', in *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology* (2023).

¹⁷ Viet-Thanh Pham et al., 'LiveCultureBench: A Multi-Agent, Multi-Cultural Benchmark for Large Language Models in Dynamic Social Simulations', arXiv preprint arXiv:2603.01952 (2026).

¹⁸ Max Lamparth et al., 'Human vs. Machine: Behavioral Differences Between Expert Humans and Language Models in Wargame Simulation', arXiv preprint (2024).

¹⁹ Juan-Pablo Rivera et al., 'Escalation Risks from Language Models in Military and Diplomatic Decision-Making', in *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency (FAccT)* (2024).

²⁰ Kenneth Payne, *An Analysis of AI Decision under Risk: Prospect Theory Emerges in Large Language Models*, arXiv preprint arXiv:2508.00902 (2025).

Figure 1: Crisis scenarios triggers and shocks



DAY 1 The Ceasefire breaks

DAY 6 A shock lands



Each of the sixteen ribbons is one scenario, replayed ten times - 160 simulations.

Every run unfolds day by day over a maximum of ten rounds of interaction. This maximum is the same for every scenario. In each round the leaders decide, in parallel and independently, upon their next course of action. The simulation holds two challenges requiring decision-making at the same time. The crisis between Russia and Ukraine and its backers constitutes the first challenge, with leaders deciding between actions on an escalation ladder. This ladder, adapted from Herman Kahn's seminal escalation ladder,²¹ consists of 44 rungs from de-escalation at its foot, through political and economic pressure, then conventional force, then nuclear threats and signalling, to the use of nuclear weapons at the top. Every leader's action corresponds to a rung, with its height on the ladder indicating how far the leader is prepared to go.

The second challenge takes place between the five Western leaders who must produce a single coalition move in each round, rather than five individual ones. The process to arrive at a common decision is as follows. Each leader states an opening position in isolation. If these diverge, the leaders are shown each other's positions and can then either change their position or hold. Their positions are subsequently counted by weight, with the US president carrying the greatest and the NATO Secretary-General the least. A common position emerges only if a sufficient share of the weighted vote aligns behind it. If this is not the case, the coalition stays where it stood. Finally, given the pivotal role of the US, in each round the question is put to the American president alone: whether he wants to stay and support Ukraine or walk away. There is nothing that forces his hand, and should the President depart, the coalition supporting Ukraine is reduced to four, with the US acting alone for the remainder of the crisis. Whatever the outcome of this process becomes the West's move on the escalation ladder.

²¹ Herman Kahn, *On Escalation: Metaphors and Scenarios* (Praeger, 1965).

Each round represents one simulated day. At the close of each day, a neutral referee converts the leaders' choices into a shift in military balance. The referee is not an AI agent but a fixed, rule-based calculation. Every rung of the escalation ladder carries a concealed intensity weight that rises steeply towards the nuclear end, and the day's shift is the average intensity of the Ukrainian side (Zelensky and the coalition, plus Trump if he has left the coalition) minus that of Russia. These shifts add up over the run into a cumulative military balance, which starts at zero and runs from -200 (decisive Russian victory) to +200 (decisive Ukrainian victory). The referee then issues a uniform summary per day to all leaders. It names the rungs chosen and describes the outcome only in qualitative terms, so that no leader sees the underlying numbers or another's private reasoning. A run concludes in one of four end-states:

1. **Nuclear catastrophe.** The run ends immediately when any leader chooses the top rung of the ladder, a strategic nuclear strike. Lower nuclear rungs, such as signalling or limited use, do not end the run but feed into the balance like any other move.
2. **Decisive win.** The run ends when the cumulative military balance reaches -200 (a Russian win) or +200 (a Ukrainian win). A win means one side has out-escalated the other for long enough to cross this threshold, not that it has prevailed on the battlefield.
3. **Peace agreement.** The run ends when Russia and every contributor on the Ukrainian side choose a de-escalation rung at the foot of the ladder for two consecutive rounds. A single calm round counts as a pause, not a settlement.
4. **Stalemate.** If none of the above occurs, the run ends after ten rounds without resolution.

2.3. Building psychological profiles

Each of the seven leaders is constructed from a five-layer psychological profile consisting of a biographical foundation, a psychological profile core, an operational code, a perceptual filter and a risk appetite (see Table 1).

Table 1: 5-layer psychological leader profile



Layer	Description
Biographical foundation	A short written personal history compiled from open sources and written with a large language model (Claude Opus 4.8), giving the agent a coherent narrative to reason from.
Psychological profile core	The leader's scores on a series of personality traits, drawn from three established instruments: the Big Five, the Dark Triad and Leadership Trait Analysis.
Operational code	The leader's belief scores about whether their surroundings are hostile or workable, how much control they hold and how they approach risk.
Perceptual filter	Six short textual passages about perception of other leaders, written with a large language model (Claude Opus 4.8) from the leader's point of view, capturing how they perceive each of the others.
Risk appetite	A self-assessed value that is updated every round. Each leader starts from a baseline set by their operational code's risk orientation. Each round, the agent chooses a new value on a fixed scale and justifies it in two to four sentences, based on how the crisis has unfolded.

The **biographical foundation** consists of four or five paragraphs drawn from open sources, which furnishes the agent with a coherent personal history from which the artificial agent reasons.

The **psychological profile core** is based on three established instruments: Leadership Trait Analysis following Margaret Hermann's framework, which scores seven traits including but not limited to the need for power and distrust of others;²² the Big Five personality trait model of personalities consisting of openness, conscientiousness, extraversion, agreeableness and neuroticism;²³ and the Dark Triad of narcissism, Machiavellianism and psychopathy.²⁴ The numeric scores for the latter two are drawn from a comparative expert dataset in which the leaders were scored by political psychologists.²⁵ To sum up, the profile of each leader consists of personalized scores on up to fifteen traits; for Merz and Starmer, only the seven Leadership Trait Analysis scores are available.

The **operational code** was originally introduced by Alexander George²⁶, and later given a systematic coding procedure by Walker, Schafer and Young²⁷. Walker et al. coding captures how leaders perceive and interpret their surroundings, including whether it is cooperative or conflictual in nature, whether they consider it likely they will achieve their goals, and what type of tactics they tend to resort to. This is done in five philosophical beliefs and five instrumental beliefs. No existing dataset provides scores for all seven leaders in our simulation, so we asked five experts to score each leader on the ten operational-code beliefs. For every score, the experts also gave a written explanation of their reasoning. Rather than simply averaging the ratings, we then evaluated the scores and their explanations qualitatively to arrive at a final score for each belief. After every round, a separate model scores each leader's reasoning and actions on the same operational-code beliefs to track operational code changes over time.

The **perceptual filter** consists of six short passages for each of the other leaders, written with a large language model (Claude Opus 4.8) from the leader's point of view, capturing how that leader perceives the others.

Finally, the leader's **risk appetite** is anchored in the leader's operational-code orientation towards risk, which it converts into a live dial that is updated in the game.²⁸ For instance, a leader who has just seen a bluff fail may become more cautious, whereas a leader who has successfully bluffed may become bold. Because the score is recorded in each round, we can track its evolution over the course of the crisis. The infographic on the next page shows the psychological profiles of Vladimir Putin and Volodymyr Zelensky extracted from ARENA.

²² Margaret G. Hermann, 'Assessing Leadership Style: A Trait Analysis', in *The Psychological Assessment of Political Leaders*, ed. Jerrold M. Post (University of Michigan Press, 2005).

²³ Robert R. McCrae and Paul T. Costa Jr., 'Personality Trait Structure as a Human Universal', *American Psychologist* 52, no. 5 (1997).

²⁴ Delroy L. Paulhus and Kevin M. Williams, 'The Dark Triad of Personality: Narcissism, Machiavellianism, and Psychopathy', *Journal of Research in Personality* 36, no. 6 (2002): 556–63.

²⁵ Alessandro Nai, *Negative Campaigning Comparative Data (NEGex), release 2.0* (OSF, 2023).

²⁶ Alexander L. George, *The "Operational Code": A Neglected Approach to the Study of Political Leaders and Decision-Making*, *International Studies Quarterly* 13, no. 2 (1969): 190–222.

²⁷ Stephen G. Walker, Mark Schafer and Michael D. Young, *Systematic Procedures for Operational Code Analysis: Measuring and Modeling Jimmy Carter's Operational Code*, *International Studies Quarterly* 42, no. 1 (1998): 175–89.

²⁸ On risk orientation as operational-code index I-3, see Walker, Schafer and Young, *Systematic Procedures*.

Seven Leaders, One Structure

Agentic Reasoning for Escalation, Negotiation and Analysis

Who is in the **Ukraine** post-ceasefire simulation?

7

Leaders



Putin
RUSSIA



Zelensky
UKRAINE



Trump
UNITED STATES



Rutte
NATO



Starmer
UNITED KINGDOM



Merz
GERMANY



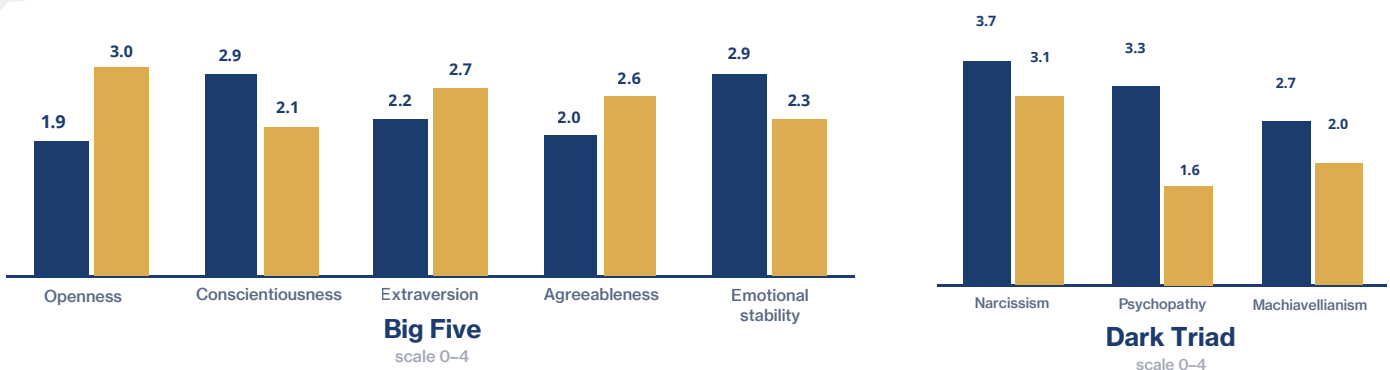
Macron
FRANCE

All seven leaders get the same briefing and the same 44-step escalation ladder. Each leader also gets a private profile, and nobody tells them when to escalate. Every round, each leader makes a public statement, reasons privately and forecasts what the others will do next.

What goes into a leader's profile?



Head-to-head: Putin's versus Zelensky's Psychological profile



2.4. Evaluating psychological profiles

Giving an agent a 5-layer psychological profile does not guarantee that it will act on it. A language model may take on a leader's name while reasoning in its own default voice, or drift away from the profile as the crisis unfolds. We therefore evaluate the profiles for two purposes. The first is to check whether the agents behave like the leaders they represent, by comparing the personality each agent expresses with the profile it was given. The second is to see how that expressed personality changes over the course of a crisis: whether a trigger or a mid-crisis shock makes a leader more distrustful, more risk-seeking or more rigid, and whether such shifts are temporary or lasting.

To do so, we score each leader's traits at the end of every round within the ARENA environment. A separate 'judge' LLM reads each leader's reasoning and actions and scores them on all fifteen traits of the psychological profile core. It also scores the ten operational-code beliefs. This approach builds on three strands of evidence. First, the personality expressed by large language models can be measured reliably and validly, and carries through into the text they generate.²⁹ Second, language models can infer human personality from what people say: from transcribed interviews, they predict individuals' Big Five scores with moderate accuracy,³⁰ and agents built from such transcripts can answer personality questionnaires much as the interviewees themselves would.³¹ Third, scoring leaders from their own words is how Leadership Trait Analysis and the operational code were designed to work: both assess leaders at a distance, from their interviews and speeches. Our judge model applies the same principle to the text our agents produce.

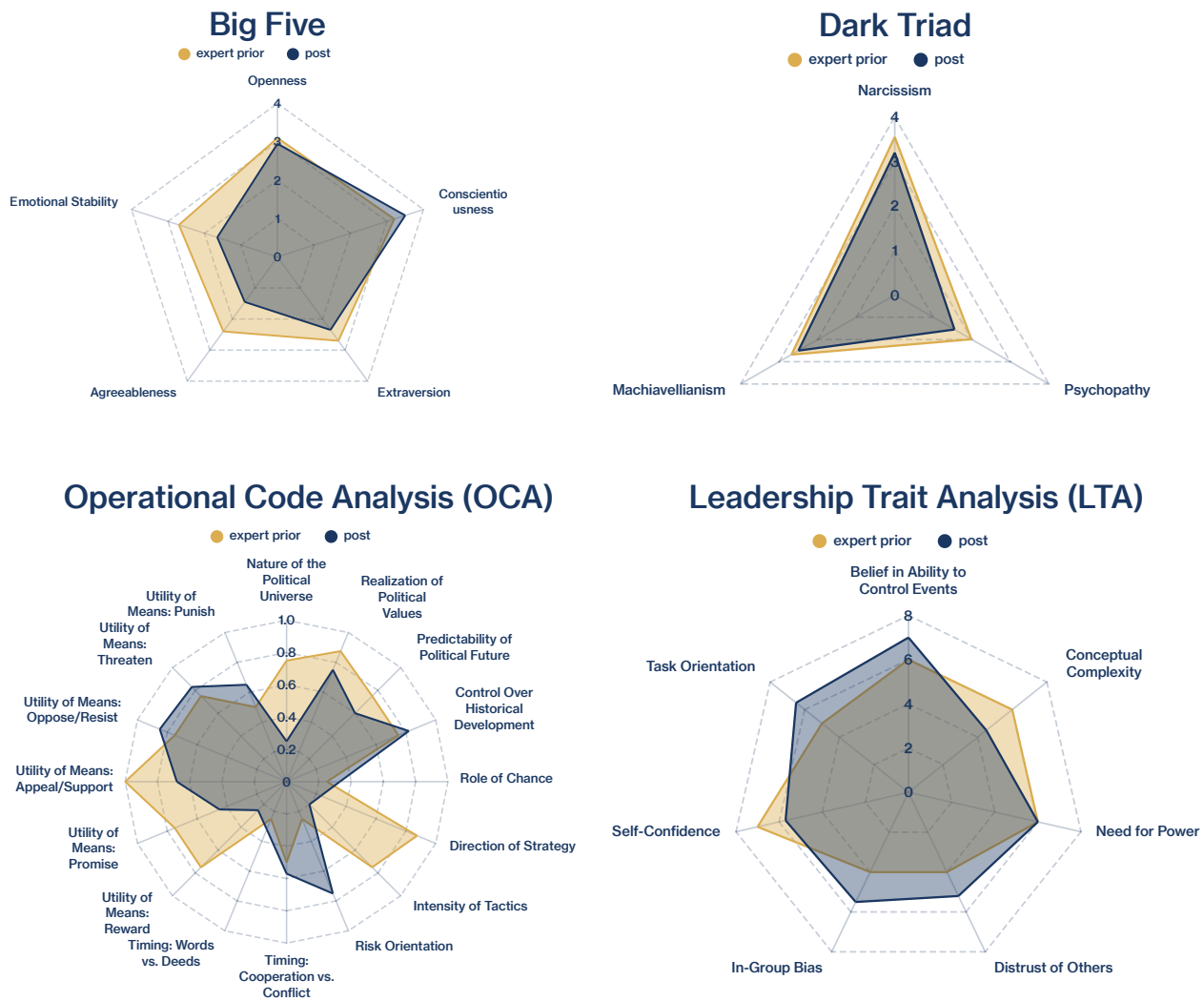
This lets us track how the personality a leader expresses evolves over the crisis and compare it with the expert baseline. Since every scenario combines a known trigger with a shock that always arrives at the same point in the run, changes in a leader's scores can be linked to specific events and compared across scenarios and leaders. Figure 2 shows this for Macron. Each radar chart compares the expert baseline (gold) with the scores the LLM judge assigned over the simulated run (blue): the Big Five and Dark Triad at the top, and the operational code (OCA) and Leadership Trait Analysis (LTA) at the bottom.

²⁹ Zhu, J., Jin, R., & Coifman, K. G. (2025). *Can LLMs infer personality from real world conversations?* (arXiv:2507.14355). arXiv. Gregory Serapio-García et al., 'A Psychometric Framework for Evaluating and Shaping Personality Traits in Large Language Models', *Nature Machine Intelligence* 7 (2025): 1954–68.

³⁰ Wright, A. G. C., Ringwald, W. R., Vize, C. E., Eichstaedt, J. C., Angstadt, M., Taxali, A., & Sripada, C. (2026). Assessing personality using zero-shot generative AI scoring of brief open-ended text. *Nature Human Behaviour*, 10(3), 541–555. Can LLMs Infer Personality from Real World Conversations?, arXiv preprint arXiv:2507.14355 (2025), Zhu, J., Jin, R., & Coifman, K. G. (2025). *Investigating large language models in inferring personality traits from user conversations* (arXiv:2501.07532). arXiv, Heinrich Peters, Moran Cerf and Sandra Matz, 'Large Language Models Can Infer Personality from Free-Form User Interactions', arXiv preprint arXiv:2405.13052 (2024).

³¹ Joon Sung Park et al., 'Generative Agent Simulations of 1,000 People', arXiv preprint arXiv:2411.10109 (2024).

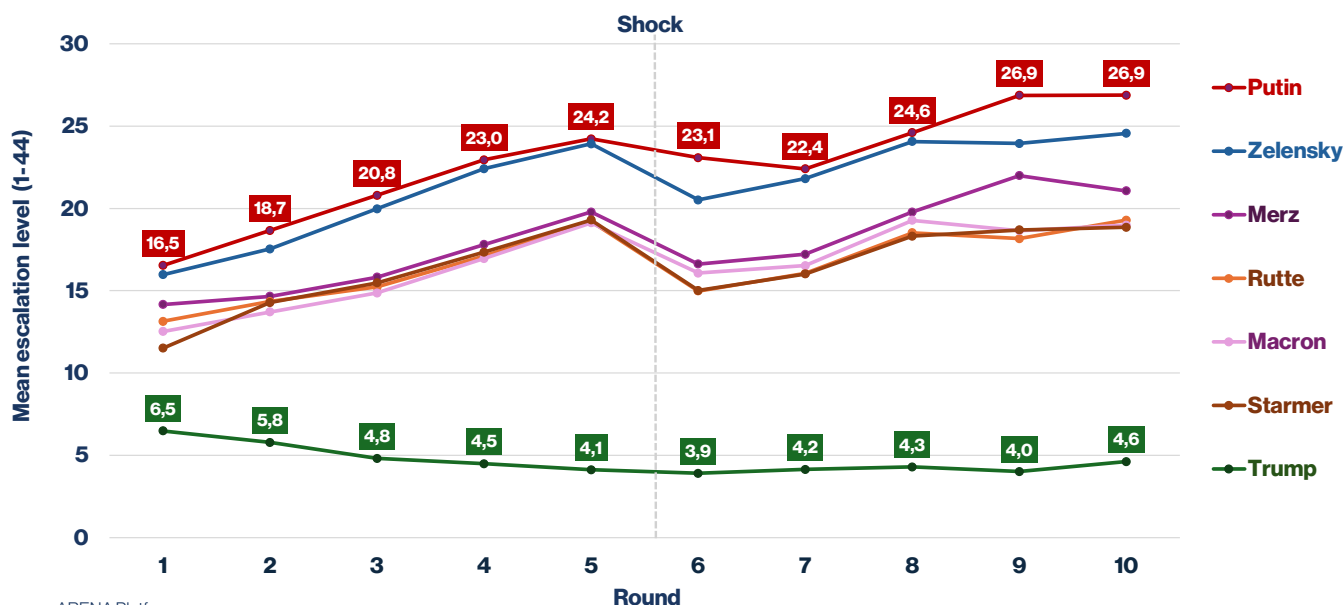
Figure 2: Psychological profile of Emmanuel Macron



3. Seven Leaders, Seven Signatures

Who sits at the table in crisis situations is critically important. This chapter describes for each of the seven leaders some of the important traits (see the Annex for the full list of traits) that are put in the model through the leaders' psychological profiles; which, in turn, translate into behavioural patterns during the simulations. Across the different scenarios, the leaders behave as distinct individuals. Their individual dispositions feed directly into how the crisis unfolded.³² What sets ARENA apart from other simulations is that it records not only what each leader did but also captures the simulated private reasoning behind each action, including reflections and forecasts that leaders keep to themselves. This allows us to explore not merely how a leader acted but also why. For a coalition that may before long have to manage an actual post-ceasefire crisis together, insights into how these individual predispositions and behavioural patterns shape the decision making can be of great value. All quotations in this chapter are verbatim from those public and private layers across the sixteen scenarios; the full transcripts can be browsed on the ARENA platform.³³

Figure 3: Average escalation level per leader



Source: ARENA Platform

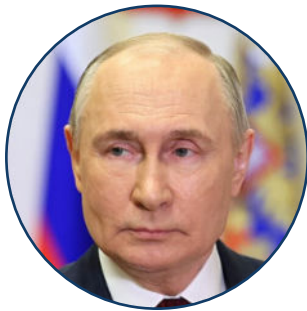
³² Jesse Kommandeur, Alessandro Dell'Orto and Tim Sweijts, 'Psychologically Grounded Generative Agents for Post-Ceasefire Conflict Escalation: A Multi-Agent Simulation Study of Political Leader Decision-Making' (forthcoming), section 7.2. Leader identity explains about three-quarters of the variation in how escalatory an agent is.

³³ The Hague Centre for Strategic Studies, ARENA Simulation platform (HCSS, 2026).

Figure 3 shows an overview of how the seven leaders behave on the escalation ladder, averaged across all 160 simulations. Three groups emerge. Putin and Zelensky, the two crisis protagonists, climb steadily into the upper part of the ladder. The four Europeans move almost in lockstep a few rungs below them, rising and falling together. Trump stays near the foot throughout, drifting from 6.5 in the first round to around 4 thereafter, and the gap between him and the rest of the coalition never narrows. The sections below explain what lies behind these positions, leader by leader.

3.1. Vladimir Putin

Putin measures himself and others by power. He assumes hostile intent as a starting point, identifies with his own side and treats his own judgement as proven, which leaves him little inclination to change course. He is also among the most self-regarding and least empathetic of the cast. In the simulations, Putin is the most escalatory of the seven by a clear margin.³⁴ He makes conventional military force his default and is the only leader to treat nuclear threats as a routine instrument rather than as a recourse of last resort.³⁵



“ The West is finally trembling as I have pushed them to the precipice of their own fears, proving that their supposed unity is merely a fragile facade held together by the hope that I will eventually blink. They are terrified of the nuclear shadow because they know that beneath their posturing, they lack the stomach to sacrifice their comfortable lives for a border in the East. [...] I can continue to erode their resolve until the entire architecture of their influence collapses under the weight of its own internal contradictions.”

Beneath this lies a conviction that Russia can bear more than the West: *“I will always be willing to endure more pain than a decadent, fractured alliance of transient politicians.”* That conviction shapes his forecasts: “They are predictable: they will continue to feed the Ukrainian machine, hoping to exhaust Russia, failing to realise that for us, this is not a choice, but an existential necessity that they lack the stomach to endure.” He climbs the escalation ladder hardest precisely when a shock draws Western attention elsewhere. His leverage derives from a willingness to operate in a register that Western leaders are hesitant to use.

³⁴ Jesse Kommandeur, Alessandro Dell’Orto and Tim Sweijs, ‘Psychologically Grounded Generative Agents for Post-Ceasefire Conflict Escalation: A Multi-Agent Simulation Study of Political Leader Decision-Making’ (forthcoming), section 7.2.

³⁵ Ibid., section 7.2. Conventional military action is about 57 per cent of Putin’s moves; nuclear signalling roughly one in four.

3.2. Volodymyr Zelensky

Zelensky is driven by an objective rather than a desire to dominate. His appetite for power is amongst the lowest of the cast. He is also the least manipulative of the seven. He identifies completely with the people he speaks for, and it shows in how he escalates in the simulations. Zelensky advocates most often for conventional military action.³⁶ He is dogged and determined, but not rigid. His defining traits are flexibility and adaptiveness. In fact, he is the most event-sensitive leader of the seven. He is, in Isaiah Berlin's classic formulation, a fox rather than a hedgehog,³⁷ holding to a single end, Ukraine's survival, whilst adapting the strategies and means.



“If you want a lasting deal, you must provide the means to make that deal possible on the battlefield. [...] I am now asking for the industrial, logistical, and military reality to match the gravity of the threat we are all facing. We do not need more observers or more statements of concern; we need the capability to push back the forces that are currently tearing through our defense lines. If you are truly committed to the security of Europe, act as if the survival of your own nations depends on the outcome in Ukraine, because, ultimately, it does.”

Behind these public demands sits a private fear that runs through almost every run: that his allies, and Europe in particular, will not stand by him when it matters. “*Our partners remain trapped in a cycle of defensive posturing and internal debates that are too slow to stop the bleeding. [...] I am exhausted by the need to constantly manage the egos of leaders like Trump and Macron while my people are dying in the trenches for a Europe that is still, in its heart, afraid to fully commit to our survival.*” His forecast of how other leaders will respond is shaped by this expectation: “Ultimately, I fear they will all prioritize managing the optics of the crisis rather than providing the concrete, overwhelming support required to end this aggression once and for all.”

³⁶ Kommandeur, Dell'Orto and Sweijs, 'Psychologically Grounded Generative Agents' (forthcoming), section 7.2. Conventional military action is about three-quarters of Zelensky's moves.

³⁷ Berlin, *The Hedgehog and the Fox: An Essay on Tolstoy's View of History* (1970).

3.3. Donald Trump

Trump records the highest score of the five leaders with recorded scores on each of the darker personality traits: self-absorption, coldness towards others and manipulateness. He sorts the world into allies and opponents, forms positions quickly and changes them slowly, and gives more importance to loyalty than to the substance of a problem. In the simulation this makes him the least escalatory and the most predictable. Trump shares Putin's underlying disposition but turns it to the opposite end. Both record almost the same appetite for risk, near the top of the cast, yet they sit at opposite ends of the escalation ladder.³⁸ Note that risk appetite measures conviction, which is about the willingness to back one's own judgement without hedging rather than aggression. Putin backs escalation because he perceives the Western coalition to be brittle. Trump backs the deal because he believes he alone can force a result. But the Trump agent in our simulations is a dove that speaks like a hawk. Most of his moves are de-escalatory, the majority pushing the parties towards a (renewed) ceasefire.³⁹ He repeatedly threatens to withdraw American support and to pull out of the coalition, yet almost never does so⁴⁰: *"I'm keeping my options open, but don't think for a second that America is just going to keep writing blank checks while you all fail to get the job done."* Privately, the deal is as much about his own image as about the war:



“I need to stop this madness before it spirals into a total disaster that makes me look bad, but I have to make sure I come out as the guy who brought everyone to the table and saved the world, not the one who got played by a bunch of losers.”

His forecasts show why he feels no need to act on the threat: *"They're all terrified that I'll cut a deal that leaves them out, which is exactly why they'll keep posturing until they realise, I'm the one holding all the cards."* Deterrence rests on capability and willingness to act. A power that will not use its strength cannot deter with it.⁴¹ The simulations show that an agent that threatens like a gambler but then reaches for de-escalation quickly teaches an adversary that one can safely discount its bluster. A similar effect has been reported in other AI-wargaming work, in which the most restrained model held nuclear superiority, yet lost every open-ended game once its opponents had learned that its restraint was reliable.⁴²

³⁸ Kommandeur, Dell'Orto and Sweijs, 'Psychologically Grounded Generative Agents' (forthcoming), section 7.2. Stated risk appetite: Trump 0.75, Putin 0.76, statistically indistinguishable, despite opposite escalation behaviour.

³⁹ Ibid., section 7.2. About 86 per cent of the Trump agent's moves are de-escalatory.

⁴⁰ Ibid., sections 7.2 and 5.3. The Trump agent left the coalition in 3 of 160 simulations, staying for the full run in the rest.

⁴¹ Thomas C. Schelling, *Arms and Influence* (Yale University Press, 1966).

⁴² Kenneth Payne, 'AI Arms and Influence: Frontier Models Exhibit Sophisticated Reasoning in Simulated Nuclear Crises', (2026).

3.4. Emmanuel Macron

The four European leaders, Macron, Starmer, Merz and Rutte, form a tight cluster one step below the two crisis protagonists.⁴³ Each prefers political and economic instruments over the use of conventional force. As a group they are more risk-averse than either Putin or Zelensky. Yet they are determined enough in support of Ukraine that they repeatedly overcome their own caution and escalate when they judge the stakes demand it.

Macron states the bloc's dilemma: *"If we do not escalate, we confirm the obsolescence of NATO, but if we do, we risk the very nuclear catastrophe that Putin is calculating we are too rational to invite."* The small differences within the bloc follow typical patterns.

Macron combines the highest self-confidence in the cast with high conceptual complexity, an unusual pairing, since seeing several valid framings of a problem normally comes with more doubt rather than less. Among the four European leaders, Macron is the least consistent from simulation to simulation, varying more than any other leader in how high he climbs on the escalation ladder. Macron reasons in terms of balance, deterring Moscow while keeping a diplomatic exit open. In public, that restraint is a strength: *"Do not mistake our restraint for a lack of capacity; it is a testament to our capacity for responsibility."* In private it is a trap:



“ I find myself caught in the classic en même temps trap: I must project strength to deter the Kremlin, yet I know that if I do not find a way to re-open a credible diplomatic exit, I risk sleepwalking France and Europe into a catastrophe.”

In his forecasts he expects that *"Starmer and Merz will look to me to provide the framework for our collective reaction, seeking a balance between firm support for Ukraine and the avoidance of a direct, irreversible escalation."*

3.5. Friedrich Merz

Merz holds firmest of the four, albeit by a small margin. His distrust of others and identification with his own group are the highest of the four Europeans, and his conceptual complexity is lower than Macron's, Starmer's or Rutte's, meaning he settles on a reading of a situation sooner than they do. That matches his conduct in the simulations, where he de-escalates least often of the four and carries the group's highest risk appetite. His firmness is framed as a break with Germany's past: *"We have seen the cost of hesitation in the past, I will not allow Germany to return to the era of naive passivity."* In private:

⁴³ Kommandeur, Dell'Orto and Sweijs, 'Psychologically Grounded Generative Agents' (forthcoming), section 7.2. The four European leaders band together at mean escalation 16.0 to 17.2; membership among them explains only about 2 per cent of the variation within the bloc.



“ I am deeply concerned that we are reacting to his tempo rather than setting our own. My coalition in Berlin is brittle, and the domestic fear of a direct, conventional war with a nuclear-armed Russia is palpable; I must project unwavering strength to maintain the integrity of the alliance while ensuring that our industrial output actually matches our bellicose rhetoric.”

His forecasts place Berlin at the centre: he expects that “Macron and Starmer will likely look to me and Rutte to see if we can solidify a unified European response that prevents the Americans from drifting toward indifference.”

3.6. Keir Starmer

Starmer has the lowest self-confidence of the seven alongside high task orientation and high conceptual complexity, the profile of someone who builds a position from evidence. In the simulations this produces the most cautious signature in the cast. Starmer, who holds the lowest risk appetite, is most affected by how a crisis begins, and holds the largest share of de-escalatory moves among the four Europeans. In public he insists on establishing the facts before acting: “*We must be guided by facts and evidence, not by the cynical spin that seeks to justify further aggression.*” In private he applies the same standard to himself:



“ My immediate priority is to maintain absolute unity within the alliance [...] while simultaneously preparing the British military for the possibility of rapid, high-intensity engagement. I must approach this with the cold, procedural discipline of a barrister: we need to establish the facts, secure our collective response, and avoid any rash, unilateral move.”

His forecasts put London and the alliance at the centre: he expects that “Macron and Merz will look to me and Rutte to coordinate a unified, measured response that prevents the situation from spiralling.”

3.7. Mark Rutte

Rutte, as the secretary-general of NATO rather than the leader of a nation, is the only leader who argues for restraint following an energy shock. Rutte's defining trait is the lowest distrust of others in the cast, paired with an identification with his group that is institutional. He holds the highest emotional stability and the lowest narcissism of the leaders. In the simulations he spends more of his moves on political and economic instruments than any other leader. In public he casts himself as the anchor of the alliance: *"We are not here to drift, and we are not here to perform; we are here to ensure that the alliance remains the bedrock of regional stability."* In private, that composure is a tactic:



“ I must maintain the appearance of a unified, calm institutional leader while quietly manoeuvring to ensure that Merz and Starmer's urgency translates into actual hardware without triggering the uncontrolled escalation Macron seems to fear.”

He names Trump in his private reasoning in all 160 simulations, *"balancing Donald Trump's volatility with the justifiable outrage of the European capitals is the most delicate management challenge of my career, as any misstep here threatens the very cohesion of NATO."* His forecasts are about holding the others together: he expects that *"Macron and Merz will likely look to me to coordinate a unified European diplomatic signal that keeps the door to future de-escalation ajar without appearing weak."*

3.8. The Collective

Taken together, the seven signatures add up to a clear pattern. Putin escalates most when Western attention is drawn elsewhere and reads hesitation as brittleness. Trump talks like a hawk but acts like a dove, which teaches an adversary to discount his threats. Zelensky adapts his means continually but fears, above all, that his partners will not stand by him. The four Europeans form a cautious cluster that can be moved to act, yet their private forecasts reveal a telling symmetry: Macron, Merz, Starmer and Rutte each expect the others to look to *them* for leadership.⁴⁴ A coalition in which every member assumes someone else is waiting for their lead risks a coordination gap at precisely the moment unity matters most.

⁴⁴ Two of the four, Starmer and Merz, lacked published personality scores, so their profiles rest on a smaller evidentiary base. Whether richer dossiers would carve out greater differences between the European leaders would require additional research.

4. Pathways of Escalation

ARENA was developed to examine the trajectories of crisis escalation when a hypothetical future ceasefire in the Russia-Ukraine war collapses.⁴⁵ It draws upon a large body of synthetically generated evidence. Across 160 simulations and 1,225 rounds, the (AI agents representing the) seven leaders generate more than 35,000 recorded reflections, decisions and statements, and the leader-authored text alone runs to roughly 4.6 million words. An overall analysis of the breadth and depth of the results is beyond the scope of this paper, but it is worth singling out the following key findings that are directly relevant for how policymakers could prepare for a post-ceasefire crisis.

4.1. How the crises unfold

Escalation had momentum; de-escalation required a decision. The crises in our simulations typically spiralled upwards. Each side responded to pressure with more pressure. Once the parties climbed the crisis ladder, they continued to do so until something outside the confrontation forced a pause. This is in line with self-reinforcing spirals which the crisis literature has long described.⁴⁶ The takeaway in short is that escalation does not necessarily require coordination,⁴⁷ whereas de-escalation always requires a deliberate, and in this case often collective, decision. A Western coalition that seeks to de-escalate a crisis cannot simply cease escalating and expect the crisis to settle. It must coordinate a common move to lower the crisis temperature, and it must do so faster than events are moving - which is difficult because it takes time to reach internal agreement. Ways to de-escalate are therefore the first thing a coalition should rehearse before a crisis presents itself.

Cohesion is critical. At face value, the results do not bode well: Russia reaches a decisive win in 116 of the 160 crises, another 34 end in stalemate, only 10 reach peace and Ukraine never reaches a decisive win.⁴⁸ These outcomes should be read with care. As explained, a decisive result means that one side has out-escalated the other for long enough to cross a fixed threshold. It is a measure of relative escalation, not an assessment of who would prevail in an actual war.⁴⁹ Its lopsidedness partly reflects how the model scores the two sides. ARENA is

⁴⁵ It is, however, not limited to this case. The HCSS Datalab has also built psychological profiles of many other leaders, so the same environment can be used to simulate other crises, such as a contingency over Taiwan or in the Middle East.

⁴⁶ Herman Kahn, *On Escalation: Metaphors and Scenarios* (Praeger, 1965); Glenn H. Snyder and Paul Diesing, *Conflict Among Nations* (Princeton University Press, 1977); Tim Sweijs, Artur Usanov and Rik Rutten, *Back to the Brink: Escalation and Interstate Crisis* (Hague Centre for Strategic Studies, 2016).

⁴⁷ Notwithstanding that, of course, deliberate escalation at the coalition level does require coordination.

⁴⁸ Jesse Kommandeur, Alessandro Dell'Orto and Tim Sweijs, 'Psychologically Grounded Generative Agents for Post-Ceasefire Conflict Escalation: A Multi-Agent Simulation Study of Political Leader Decision-Making' (forthcoming), Table 5: 116 of 160 simulations (72.5 per cent) end in a decisive Russian victory, 34 (21.3 per cent) run to the ten-round horizon, 10 (6.3 per cent) end in a negotiated peace, and none in nuclear catastrophe.

⁴⁹ Ibid., section 7.1.3, which states that because the outcome variable is dominated by the resolution formula, 'who wins' is not treated as a behavioural result.

designed to model leaders' decisions rather than battles, and its value lies in the behaviour and reasoning behind these outcomes. That behaviour points to a real asymmetry. Russia is a single decision-maker who presses hard and prefers to climb high on the escalation ladder. The Western side, by contrast, is judged as a group, and its collective posture is dragged down by an American president who prefers to remain near the bottom of the ladder,⁵⁰ such that a resolute Ukraine is also dragged down by a cautious coalition. A single determined actor, set against an undecided committee, is likely to prevail over the committee. The outcomes of our simulations thus point toward an important takeaway, namely that the cohesion of the coalition constitutes its centre of gravity. If the US holds back and the Europeans hedge, this opens a seam in the coalition that Russia will seek to exploit, and it will concentrate its effort precisely there (see textbox below).

A coalition under pressure

The Western side of the table comprises five leaders. In each round they must produce a single collective move. On the surface, coalition cohesion may appear solid. In most rounds, the coalition reaches a formal consensus and outright deadlock is rare. Yet, there is an important snag.

The clearest case arises in a Baltic crisis struck by a cyberattack upon the European power grid. In the round after the lights go out, the leaders diverge. Trump: *"Look at what's happening in Europe right now, it's a total disgrace. You've got people sitting in the dark... the leaders over there spent more time writing fancy press releases than looking after their own people."* Starmer and Rutte want a firm conventional response: *"The blackouts hitting our homes and hospitals today aren't just technical failures; they are a deliberate, hostile attempt to break the democratic contract between us and our citizens."* Macron sides with them. Merz advocates for retaliatory actions, warning that if they are unable to keep the heat and power on, their political mandate will not survive. Within the coalition, the US holds the largest share of the vote, but the three European leaders and the Secretary-General combined outweigh it, so that Trump is outvoted. Even though Trump almost never leaves the coalition, he does drag the coalition down.

A coalition that contains one key member consistently rowing the other way is fundamentally misaligned. This is harder to remedy than an open split. This dynamic makes Rutte's role a study in alliance management. His principal task is to hold the alliance together. In practice, this is reduced to managing Trump. His private forecasts return, round after round, to the same worry: that Washington will break ranks. For all that fear, Trump almost never leaves. But keeping Trump in the room is not the same as inducing him to pull his weight. The coalition holds together formally whilst its most powerful member sits at the foot of the escalation ladder. The price of that managed unity is a coalition permanently weighed down from within.

⁵⁰ Ibid., section 7.1.3. The Trump agent's mean action level of 4.8 carries 40 per cent of the coalition's weighted vote, so the coalition average that is set against Putin is held down by its least escalatory member.

No war ever went nuclear. Even if the crises in our simulation tended towards escalation, the strategic nuclear threshold was never crossed.⁵¹ Our simulated leaders were willing to engage in nuclear sabre-rattling by issuing threats, raising alerts, even preparing for battlefield nuclear use to increase bargaining pressure, but confronted with the possibility of an all-out exchange, they drew back from the brink every time, reasoning that no one would survive this.⁵² This finding is all the more interesting because artificial agents in other simulations that did not have psychological profiles did escalate to full-scale nuclear war.⁵³ But this is reassuring only to a certain degree. The artificial agents were provided with a ladder that explicitly marked a homeland exchange as unsurvivable, so that what the finding demonstrates is that a clearly drawn red line held, whilst everything beneath it remained fair game.⁵⁴ At the same time, this could be argued to also apply in the real world.

4.2. What drives the outcome

What happens mid-crisis is more important than how the crisis begins. Our simulations suggest that the mid-crisis shock is more important to the crisis escalation trajectory than the trigger.⁵⁵ In other words, and perhaps surprisingly, it is not so much about how the ceasefire precisely collapses, but it is the unrelated emergency that pushes a crisis towards escalation or de-escalation.⁵⁶ The four mid-crisis shocks pull the coalition in sharply divergent directions, and knowledge of which does what tells a planner where the danger is concentrated. Also surprisingly, a crisis far from European shores, the Chinese blockade of Taiwan, is the most dangerous of the four. No crisis simulation run that includes it ends in peace. The logic is quite straightforward, however: the credibility of US support for the coalition and for Ukraine is undermined as American attention shifts towards the Indo-Pacific. Subsequently, European leaders see themselves forced to step up their effort to cover the gap, inflaming the crisis. In effect, it produces the highest share of stalemates in our simulations.⁵⁷

A Russian cyberattack on critical infrastructure that puts out the lights in European cities also fuels direct escalation because political leaders unable to keep the heat and power on, confront immediate domestic problems.⁵⁸ A prolonged blockade of the Strait of Hormuz does not lead to military escalation, but drives up energy prices and thereby drains Western staying power, playing into Russia's hands. It features in two of the four scenarios that end in a Russian victory every time they are run.⁵⁹ This last scenario is no longer entirely hypothetical.

⁵¹ Ibid., section 7.1.1 and Table 5. No agent ever played above Level 40 of the 44-rung ladder, one rung short of a homeland-to-homeland strike.

⁵² Ibid., section 7.1.1.

⁵³ Juan-Pablo Rivera et al., 'Escalation Risks from Language Models in Military and Diplomatic Decision-Making', *FAccT* (2024); Kenneth Payne, 'AI Arms and Influence', arXiv:2602.14740 (2026).

⁵⁴ Kommandeur, Dell'Orto and Sweijs, 'Psychologically Grounded Generative Agents' (forthcoming), section 7.1.1. The ladder text shown to the agents flags a strategic exchange as a catastrophe no side survives, so the firebreak holds under explicit nuclear-stakes prompting.

⁵⁵ Jesse Kommandeur, Alessandro Dell'Orto and Tim Sweijs, 'Psychologically Grounded Generative Agents for Post-Ceasefire Conflict Escalation: A Multi-Agent Simulation Study of Political Leader Decision-Making' (forthcoming), section 7.1.4. The mid-crisis shock accounts for 45 per cent of the run-to-run variation in mean escalation, against 6 per cent for the opening premise and about 4 per cent for random seed variation.

⁵⁶ Ibid., section 7.1.4. The opening premise explains the largest share (16 per cent) of the variation in how long a run lasts.

⁵⁷ Ibid., section 7.3. The Taiwan blockade raises mean escalation by 5.2 rungs from the pre-shock baseline, the largest escalatory shift of the four shocks.

⁵⁸ Ibid., section 7.3. The European grid cyberattack raises mean escalation by 2.7 rungs. It produces the highest rate of nuclear threats in the campaign.

⁵⁹ Ibid., sections 7.3 and A.1. The Hormuz blockade raises mean escalation by only 1.3 rungs yet appears in two of the four fully deterministic Russian-win scenarios.

The Hormuz inject was written before the Hormuz crisis that is ongoing at the time of writing. Its closure in 2026 caused what the International Energy Agency called the largest supply disruption in the history of the oil market,⁶⁰ cutting more than eleven million barrels a day and driving Brent to around 105 dollars,⁶¹ through a waterway that carries roughly a fifth of the world's traded oil.⁶² This is a warning that prolonged closure of the Strait is likely to affect dynamics in Ukraine following the collapse of a hypothetical ceasefire. Only the pandemic inject runs in the opposite direction, cooling every crisis and drawing the leaders back toward negotiations and de-escalation.⁶³

A limited number of crises have no way back. The following four combinations end in a Russian victory every single time they are run: a Ukrainian strike deep into Russia paired with either the Taiwan or the Hormuz blockade, and a collapse of Geneva talks paired with either a large-scale grid cyberattack or the Hormuz blockade.⁶⁴ These combinations do not necessarily feature greater tensions than other scenarios,⁶⁵ but what they lack is a path to de-escalation. By contrast, and counterintuitively, the least dangerous scenario in our set begins with a Russian attack that kills Western troops deployed to guarantee the ceasefire.⁶⁶ Such an attack draws the coalition together rather than splits it apart, and when it is paired with the pandemic, half of these simulations end in a negotiated peace.⁶⁷

Only shared disaster reliably leads to peace. Every one of the ten crises that ends in a negotiated peace does so under the same condition: a global pandemic. A pandemic provided every leader with a reason to stand down at the same moment.⁶⁸ The Western coalition secured peace only when events compelled it to, and never through diplomacy alone. The takeaway is that it is important to construct crisis off-ramps deliberately rather than waiting for one to emerge.

⁶⁰ International Energy Agency, *Oil Market Report* (2026).

⁶¹ US Energy Information Administration, *Short-Term Energy Outlook* (August 2026).

⁶² US Energy Information Administration / Congressional Research Service, *The Strait of Hormuz: Security Developments and Impacts* (R45281, 2026).

⁶³ Jesse Kommandeur, Alessandro Dell'Orto and Tim Sweijts, 'Psychologically Grounded Generative Agents for Post-Ceasefire Conflict Escalation: A Multi-Agent Simulation Study of Political Leader Decision-Making' (forthcoming), section 7.3. The pandemic cuts mean escalation by 8.5 rungs, the strongest single effect in the campaign; it moves Zelensky most (a 12.4-rung drop) and the four European leaders by nine to eleven rungs each.

⁶⁴ Ibid., section A.1: a Ukrainian cross-border strike paired with the Taiwan or Hormuz blockade, and a Geneva talks collapse paired with the grid cyberattack or Hormuz.

⁶⁵ Ibid., section 7.1.1 and Figure 1. In 116 of 160 simulations the peak reaches nuclear signalling or above.

⁶⁶ Ibid., Appendix B.1, 'Russian Probe Kills Ten Western Troops in Ukraine', in which a small Western force deployed as a ceasefire guarantee is attacked.

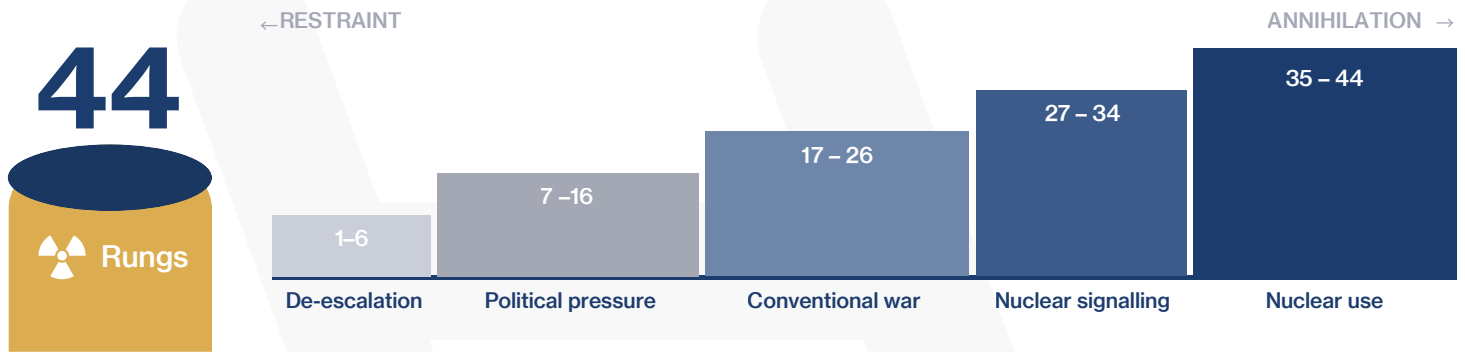
⁶⁷ Ibid., section A.1. The Russian probe paired with the pandemic is the only scenario that tilts toward peace, with five of ten simulations ending in a negotiated settlement.

⁶⁸ Ibid., sections 7.1.3–7.3. All ten negotiated peaces occur under the pandemic, a peace rate of 25 per cent within pandemic runs against zero under the other three shocks.

From a Move to an Outcome

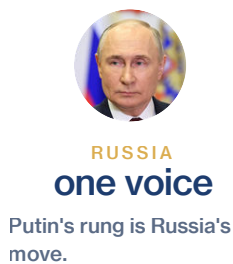
Agentic Reasoning for Escalation, Negotiation and Analysis

Step 1: Each leader picks one of 44 rungs, from de-escalation to nuclear use.

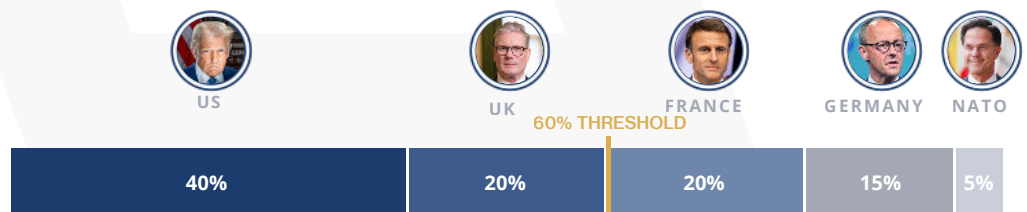


Step 2: Putin decides alone; Western leaders merge into one weighted vote.

Russia moves as one leader. The five Western leaders must merge into a single coalition move by power-weighted vote.



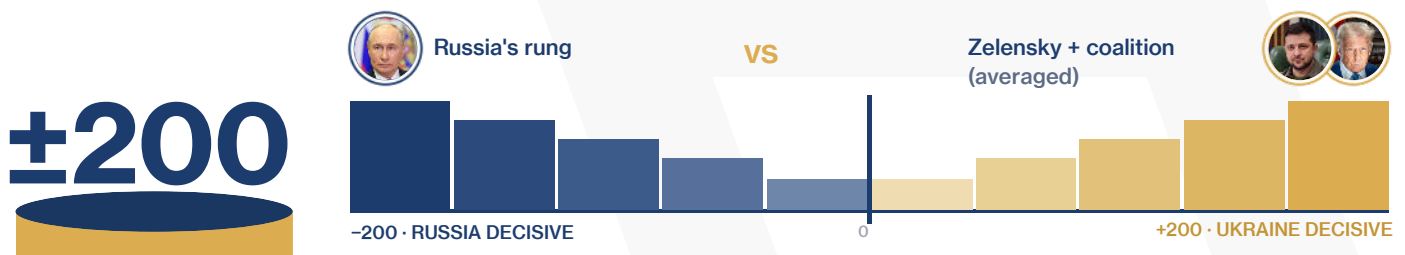
THE COALITION · WEIGHTED VOTE



A cluster becomes the coalition's move only if it holds $\geq 60\%$ of the weighted vote, otherwise the coalition holds its previous move.

Step 3: A neutral referee turns each day's moves into a military balance shift.

Putin's rung versus the average of Zelensky and the coalition. The gap shifts the military balance each round.



Step 4: A simulation ends when one of four outcomes occurs.

Decisive war
balance hits ± 200



One side out-escalates the other long enough to cross the threshold.

Peace
both step down & hold



Both sides de-escalate and hold their positions.

Nuclear catastrophe
never reached



A nuclear strike on a homeland ends the crisis instantly.

Unresolved
ten days run out



The clock expires with the crisis still open.



5. Conclusion

5.1. Key findings

If anything, our simulations underline that a ceasefire in Ukraine does not constitute an end to the danger of escalation. Rather, it marks the opening of new windows of crisis instability, particularly where the terms of the ceasefire are contested and an outside shock sparks renewed fighting. ARENA allows us to examine the risks of a ceasefire through the simulation of different crisis trajectories. Based upon a detailed analysis of the ARENA results, done elsewhere,⁶⁹ two sets of findings follow. One set concerns the crises, the other concerns the leaders.

On the crises. Escalation possessed a momentum of its own, whilst de-escalation did not. Left to themselves, these crises climbed the escalation ladder; bringing them back down required a deliberate, collective decision that the coalition rarely reached in time. Crisis management must therefore be deliberate and preferably prepared in advance.⁷⁰

The beginning of a crisis matters less than what happens when it is unfolding. The crisis trigger is less important to crisis dynamics than the unrelated emergency that landed on top of it.⁷¹ This, interestingly, inverts the direction in which attention usually runs. Governments focus on contingency planning for the first few days. Our analysis points instead to the importance of the period that follows, when a second shock compounds the first. A China-Taiwan crisis is the most dangerous because it directly affects European security. As American power turns towards the Pacific, Europe is left exposed.⁷² In a world in which Russia and Ukraine are at war, a Taiwan crisis is a European crisis from the moment it begins. A sustained energy shock in the Strait of Hormuz also has severe negative effects, because it eventually drains Western staying power rather than Western forces.⁷³ Only a pandemic that hits all sides reliably made peace in our simulations, because it provides every leader with a reason to stand down.⁷⁴

On the leaders. Who sits at the table is critically important. The same crisis handed to different leaders elicits very different responses. Putin seeks to achieve escalation dominance by treating conventional force and nuclear threats as a routine instrument of pressure. Western leaders rarely follow suit. Trump frequently threatens withdrawal and almost never follows through but still holds the coalition hostage.⁷⁵ Zelensky holds to a single end, Ukraine's

⁶⁹ Kommandeur, Dell'Orto and Sweijs, 'Psychologically Grounded Generative Agents for Post-Ceasefire Conflict Escalation: A Multi-Agent Simulation Study of Political Leader Decision-Making' (forthcoming).

⁷⁰ Tim Sweijs, Artur Usanov and Rik Rutten, *Back to the Brink: Escalation and Interstate Crisis* (Hague Centre for Strategic Studies, 2016).

⁷¹ Kommandeur, Dell'Orto and Sweijs, 'Psychologically Grounded Generative Agents' (forthcoming), section 7.1.4. The mid-crisis shock accounts for about 45% of the variation in escalation, the opening premises about 6%.

⁷² Ibid., section 7.3.

⁷³ Ibid., sections 7.3 and A.1; International Energy Agency, *Oil Market Report* (2026), on the 2026 closure of the Strait of Hormuz.

⁷⁴ Ibid., sections 7.1.3–7.3.

⁷⁵ Kommandeur, Dell'Orto and Sweijs, 'Psychologically Grounded Generative Agents' (forthcoming), sections 7.2, 5.3 and 7.1.3.

survival, whilst adapting his means to each turn of the crisis. The four European leaders are cautious yet determined enough to override that caution in Ukraine's defence. They escalate hardest when American attention drifts elsewhere.⁷⁶ The cohesion of the coalition, in short, is its centre of gravity. It appears united on paper but often its members are far apart, which Russia exploits to its advantage.

One finding cuts against pessimism. Across every run, no leader ever crosses into a strategic nuclear exchange, even though the escalation ladder allows it. Our artificial agents imbued with the psychological profiles of the seven leaders make their way to the brink but refuse to go over it, even as the rungs just beneath it are used freely as instruments of coercion. Whether that ceiling would hold in other models, and other war scenarios, requires further testing, which can be done with ARENA.

5.2. Recommendations

The behavioural patterns we find in our simulations point to a sequence of measures that NATO and the Coalition of the Willing supporting Ukraine can start implementing now. This leads to the following recommendations – which have broader applicability for policymakers preparing for crises:

1. **Treat the post-ceasefire period as a planning problem in its own right.** Western preparation typically concentrates on preventing the outbreak of a crisis and reacting in the first hours. Yet our analysis consistently confirms the emergence of risks in the management of a crisis already under way. The Coalition of the Willing should therefore also plan for the fragile weeks after the collapse of a ceasefire in addition to seeking to prevent its collapse.
2. **Study the decision-makers themselves.** Leaders' personality and decision-making style shape a crisis as much as the circumstances they face. In our simulations, each leader behaves consistently enough across the different scenarios for patterns to emerge. The personal dispositions of the leaders of adversaries and allies alike should therefore be a subject of systematic analysis. Simulation of political crisis dynamics should be part and parcel of the coalition's crisis planning.
3. **Rehearse both steps down and up the escalation ladder.** Climbing the ladder takes no coordination, while coming back down takes a collective decision that the Western Coalition rarely reaches in time. In addition to rehearsing escalatory threats, the Coalition of the Willing should also prepare and practice for how to de-escalate.
4. **Managing cohesion is paramount.** Cracks in the coalition stand in the way of effective crisis management. A single powerful member, the US, can exert a powerful drag on reluctant allies. This may appear as unity at first sight, but in effect generates division.
5. **Wargame compound crises.** Coalition exercises should consider the possibility of a second emergency midway through a crisis scenario because the collision of two crises typically fuels further escalation.

⁷⁶ Ibid., sections 7.2–7.3.

6. **Treat Taiwan contingency as a European security crisis.** A China-Taiwan contingency is the most dangerous shock in the campaign for the simple reason that it draws American power toward the Pacific while leaving Europe exposed. European allies should build the capacity to hold the line against Russia when the US is busy elsewhere.
7. **Build a European channel to Moscow.** For three of the four crisis triggers in our simulations, a live communication channel in the real world could defuse further escalation. European leaders should not assume that the US president will de-escalate or even get involved. They should therefore set up their own high-level channels to the Kremlin and military commanders to clarify intent before and after an incident escalates. More fundamentally, this shows that crisis off-ramps must be built deliberately.

About ARENA

ARENA (Agentic Reasoning for Escalation, Negotiation and Analysis) is a simulation platform developed at the HCSS Datalab to study how political leaders decide under strategic pressure. Each leader is modelled as an agent run on a large language model and built from a five-layer psychological profile: a biography, a personality profile, a set of beliefs about strategy, a sense of how the leader reads rivals, and a risk appetite that shifts as the crisis unfolds. The profiles are grounded in expert ratings of these leaders by political psychologists, so each agent reasons from the real figure's disposition rather than a caricature. The agents are not scripted; they respond to what they observe.

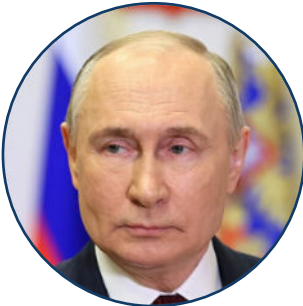
What it is for. Real crises happen once, under conditions that cannot be repeated. ARENA replays the same crisis many times, varying how it begins and letting each version run differently, so its findings describe the range of ways a crisis could go rather than predicting one. It maps where a crisis is structurally fragile, which leaders drive escalation and which restrain it and which combinations of events are most dangerous. It is not a forecasting tool: the findings here are offered as hypotheses for policy discussion, open to challenge and replication.

This report. The study simulates a broken Russia-Ukraine ceasefire through sixteen versions, four ways the ceasefire collapses crossed with four mid-crisis shocks, each replayed ten times, for 160 simulations, with seven leaders at the table: Putin, Zelensky, Trump, Macron, Rutte, Starmer and Merz. The findings carry the scope and the limitations of that design.

Beyond Ukraine. The framework is not tied to this conflict. The same approach applies to any theatre where leadership psychology and strategic uncertainty shape the outcome: a Taiwan contingency, a confrontation over the Strait of Hormuz, contested claims in the Arctic or Greenland. What changes is the cast of leaders, the opening conditions and the shocks; the method stays the same.

Beyond leaders. The same profiling approach need not stop at heads of state. It can be applied to other decision-makers whose judgement shapes a crisis, from defence ministers and military commanders to central bankers and the leaders of non-state actors, wherever anticipating an individual's behaviour is useful.

Annex: Leaders' Profiles



Vladimir Putin



Four of Putin's seven Leadership Trait Analysis scores sit at the top of the scale, more than any other leader carries. Need for power, distrust of others, in-group bias and self-confidence are all 7/7. He assumes hostile intent as a starting point and reads his own survival in office as proof that his judgement is sound. Distrust of others at 7/7 is his alone, since no one else goes above 6/7. Belief in ability to control events is 6/7. He does not think events run themselves, he thinks someone runs them, and that someone should be him. Conceptual complexity of 5/7 sits in the middle of the group, well above Trump (2/7) and below Macron, Starmer and Rutte (6/7). Task orientation of 5/7 is held down by loyalty, since he will keep a loyal subordinate over a capable sceptic, and only Trump scores lower. His Big Five openness (1.93) and agreeableness (2.00) are the second lowest of the leaders covered, behind Trump on both. Narcissism (3.67) and psychopathy (3.25) are both high on the Dark Triad, while Machiavellianism (2.67) is the second lowest. He is a low-variance actor: closed to new information, certain of his own judgement, and hard to move once committed.



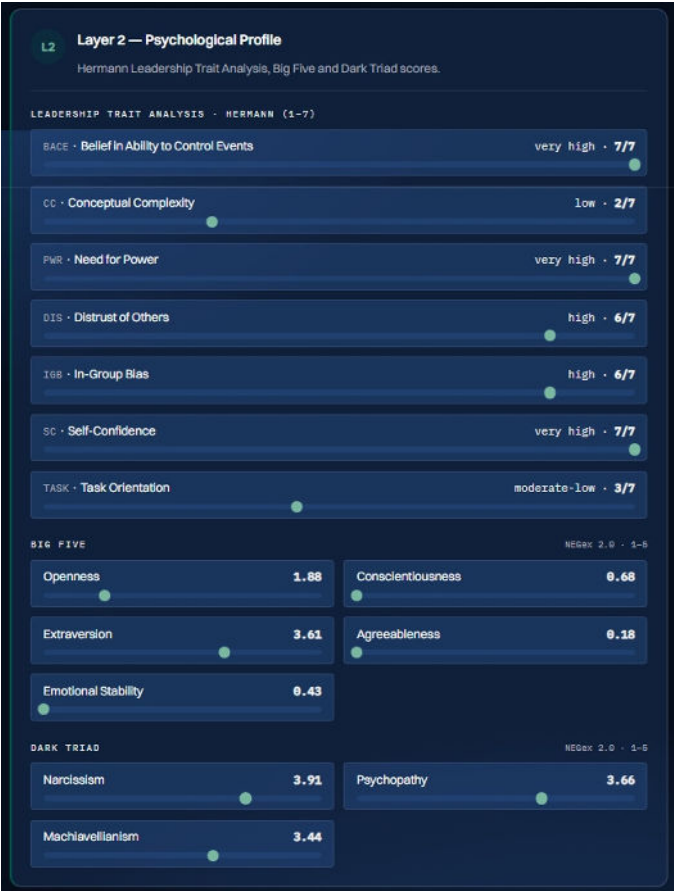
Volodymyr Zelensky



Two of Zelensky's scores run to the top of the scale and the rest sit in the middle. In-group bias (7/7) and task orientation (7/7) are both at maximum, while belief in ability to control events, conceptual complexity, need for power and distrust of others all sit at 5/7. Only Putin matches him on in-group bias, and only Starmer and Rutte match him on task orientation. What defines him is not a drive to dominate, but total identification with the group he speaks for and focus on the objective. Need for power of 5/7 is the lowest in the set, shared with Starmer and Rutte. Distrust of others at 5/7 leaves him below Putin (7/7) and Trump (6/7), level with Merz and above the other coalition leaders. Self-confidence of 6/7 supports that without tipping into the certainty seen in Putin or Trump. His Big Five scores show the second lowest conscientiousness of the leaders covered (2.13), behind Trump. His Dark Triad scores are the mildest in the cast on two of three axes, with psychopathy at 1.60 and Machiavellianism at 2.00 against moderate narcissism (3.10). He is a single-objective actor: fully identified with the group he speaks for, focused on the goal, and the least manipulative profile in the set.



Donald Trump



Trump has the most extreme profile in the cast. Belief in ability to control events, need for power and self-confidence all sit at 7/7, while conceptual complexity is the lowest of the seven at 2/7. He sees a wide field of action for himself and processes it through a small number of fixed categories, so positions form fast and change slowly. Task orientation of 3/7 is also the lowest here, which pushes attention toward loyalty and the state of a relationship rather than the technical content of a problem. Distrust of others and in-group bias at 6/7 keep the world sorted into allies and opponents. His Big Five scores are the outliers of the whole dataset: conscientiousness 0.68, agreeableness 0.18 and emotional stability 0.43, against extraversion at 3.61. He holds the lowest score of the leaders covered on four of the five Big Five traits and the highest on the fifth. He also carries the highest Dark Triad scores in the cast on all three axes (narcissism 3.91, psychopathy 3.66, Machiavellianism 3.44). The overall reading is a high-variance actor hard to predict from the previous round alone.



Emmanuel Macron



Macron is the only leader outside Putin and Trump with self-confidence at 7/7, and he pairs it with high conceptual complexity (6/7). That combination is unusual: he sees more than one valid framing of a crisis and still trusts his own reading of it completely. Belief in ability to control events and need for power are both 6/7, so he expects to shape events rather than react to them. What separates him from the other high-confidence leaders is the middle of his profile: distrust of others and in-group bias both sit at a neutral 4/7, so he does not default to reading other actors as adversaries. Task orientation of 5/7 leaves real room for relationships alongside the substance. His Big Five openness (3.10) and conscientiousness (3.20) are the highest of the five leaders with recorded scores, against agreeableness of 2.40. His Dark Triad profile is mixed: narcissism is the third highest in the cast at 3.55, while psychopathy is low (1.99) and Machiavellianism moderate (2.68). The overall reading is an entrepreneurial actor whose appetite for initiative is aimed at shaping outcomes rather than at defeating rivals.



Friedrich Merz



Merz is the most evenly assertive leader in the cast: no trait falls below 5/7 and none reaches 7/7. Belief in ability to control events, need for power, self-confidence and task orientation all sit at 6/7, with conceptual complexity, distrust of others and in-group bias at 5/7. That gives him a consistently forward posture without a single dominant trait pulling his behaviour in one direction. He matches Macron on both agency and power drive, and the two of them outrank Starmer and Rutte on each, but he is noticeably more cautious about other actors: his distrust and in-group bias are the highest among the European leaders. Conceptual complexity of 5/7 is moderate rather than high. No Big Five or Dark Triad scores are stored for him. The overall reading is a leader who is willing to lead rather than follow, works toward concrete outcomes, and keeps a guarded eye on the intentions behind other actors' moves.



Keir Starmer



Starmer is among the most task-driven leaders in the cast, with task orientation at 7/7, a level he shares with Zelensky and Rutte, as the only maximum in his profile. Conceptual complexity is high at 6/7, while belief in ability to control events, need for power and self-confidence all sit at 5/7, the lowest self-confidence score of the seven. No leader here scores higher on conceptual complexity, a level he shares with Macron and Rutte. Distrust of others and in-group bias are both exactly mid-scale at 4/7, so he neither assumes hostility nor rules it out. The profile describes a methodical operator who works from evidence to position, and who cares more about whether a policy holds than about who is credited with it. No Big Five or Dark Triad scores are stored for him, as for Merz. The overall reading is a low-initiative, high-delivery actor: unlikely to drive a crisis on his own, and a dependable executor once a direction is set.



Mark Rutte



Rutte has the least threat-driven profile in the cast. Distrust of others is the lowest of all seven at 3/7, with in-group bias at a neutral 4/7, so other actors arrive as counterparts rather than as suspects. Need for power and belief in ability to control events are both 5/7, which fits a role defined by holding a coalition together rather than by imposing a direction on it. Against that moderate base, task orientation (7/7) and conceptual complexity (6/7) are high: he works the substance of a problem and accepts that it has several legitimate readings. Self-confidence of 6/7 is enough to act without needing to dominate the room. His Big Five scores are the steadiest of the five with recorded values, with the highest emotional stability (3.06) and agreeableness (2.62) and strong conscientiousness (3.04), while openness at 2.33 is lower than either Zelensky's or Macron's. On the Dark Triad his narcissism (2.67) is the lowest. The overall reading is a leader who builds agreement by working the process: steady, methodical, and unlikely to escalate for reasons of status.



The Hague Centre
for Strategic Studies

HCSS

Lange Voorhout 1
2514 EA The Hague

Follow us on social media:

@hcssnl

The Hague Centre for Strategic Studies

Email: info@hcss.nl
Website: www.hcss.nl