

Mega-Game Dire Straits

Assessment by Analyst Team US

October 1, 2017

Connections UK 2017 Wargaming Conference King's College, London University, London UK

Stephen Downes-Martin PhD

Research Fellow, Naval War College, US
stephen.downesmartin@gmail.com

Phillip Pournelle (USN Ret)

Senior Director for Gaming and Analysis, Long Term Strategy Group, US
phillip.pournelle@verizon.net

Paul Vebber (USN Ret)

Asst. Director, Concept Development and Wargaming, HQ Naval Undersea Warfare Center, US
paul.w.vebber@navy.mil

Connections UK 2017 kicked off with the Dire Straits Mega-Game played on September 5th. In accordance with the desire of the Conference Planning Committee to raise the professional bar of the conference, a research component was added to this game to address convergence and divergence in wargame analysis, specifically to answer the question “would three different groups of analysts, each observing the same game and with access to similar materials and documentation, reach similar conclusions about the validity of the wargame methodology adopted and the substantive findings of the game?”

This document is the report of the US Analyst Team. Note that this document addresses how well the game achieved the game designer’s objectives. The research question concerning convergence and divergence in wargame analysis can only be answered by a meta-analysis of the three analyst team reports (Team Europe, Team UK, Team US).

This report does not provide an introduction or overview of the game. That can be obtained from the game designer at <https://paxsims.wordpress.com/2017/09/09/dissecting-dire-straits/>.

Introduction

Dire Straits is a large scale game set in the western Pacific depicting a “perfect storm series of international crisis events region over a two month period in 2020”.¹ The game combines techniques taken from seminar style wargames, boardgames, and role playing games to create a narrative game experience for over 100 players. While the game is “not a military simulation” the designer considers it a simulation² (at least aspirationally) of the “process of intercommunications, political and diplomatic signaling in a period of escalating tensions and international crisis that is the primary area of interest and focus”. The central question the game was intended to address was:

“How would the unpredictability of the US policy under the Trump Administration, and the growing strategic power of China affect crisis stability in East and Southeast Asia?”³

In keeping with the goal of Connections conferences to improve the art and science of wargaming, three teams of analysts were given the task to observe the game and offer conclusions about the validity of the wargame methodology adopted and the substantive findings of the game⁴. However, since the lead game designer has already published his own analysis of game findings,² this report focusses on whether those findings are supported by the game methodology, design and execution.

The USA team allocated a member to each of the three subgames of North Korea, US White House and the UN Security Council. However, most of the players were in teams playing countries of interest other than North Korea, the US and the UN, and did so in a single large room along with the North Korea subgame. Therefore, the single analyst shadowing the North Korea subgame was able to obtain only a cursory overview of these players’ progress.

¹ Dire Straits player’s guide. Text in quotes indicates verbatim statements from it.

² “Dissecting DIRE STRAITS”, <https://paxsims.wordpress.com/2017/09/09/dissecting-dire-straits/>, an article by one of the game’s designers, Rex Brynan.

³ ibid

⁴ ibid

North Korean Subgame

The North Korean subgame was a worker placement game⁵, within a matrix game facilitation/adjudication framework. The North Korean generals jockeyed for the favor of the Dear Leader, with success allowing the promotion of “cadre” game pieces up to more influential tiers within the games representation of a power structure. Players also had cards available to them that allowed them to take or veto specific types of actions.

This internal politicking focused the North Korean players on a limited, inwardly focused incentive/risk structure that had little influence on, and was little influenced by, what was going on in the rest of the game. The exception was in regard to taking actions of various kinds that were intended to get visibility in the media. This was indicated through the relationship between those actions and a track that represented the approval level of the population. How exactly this relationship worked was not discernable and appeared mostly to be a white cell interpretation of the events transpiring during the turn. In some cases there was a disconnect involving the “how would they know” regarding things like a failed missile test, or movement of friendly and enemy military forces. These generally had a negative effect on the approval rating of the NK government, despite issues with just how they would know, given the insular, information-controlled nature of the NK media. This degradation in approval rating had little effect on NK decision-making, particularly due to the lack of understanding of the risk/reward structure of a successful or failed action on the approval track. There also was not a direct effect of a negative approval until it got so bad as to be “revolutionary” and the government falls.

This insularity and limited feedback directly connecting the game play in the NK subgame to the central question of the game, stated above. In large measure, this prevented substantive analysis of the effect of NK gameplay to the broader game, as the effect of US policy on it where marginal. Discussion focused on what actions could be taken to influence the news cycle, without a lot of attention paid to the reaction by the White House. If the actions of NK

⁵ <https://boardgamegeek.com/boardgamemechanic/2082/worker-placement>

dominated the news cycle, it was considered to be benefiting the overall NK narrative that the Dear Leader was keeping its international foes to heel. This was an emergent player behavior resulting from the “buzz in the room” from other players when the newsfeed carried reports of NK activity. Whether there is a real world insight in that dynamic or if it was a game artifact of the manner in which the game was implemented (with “newsfeed” twitter displays around the room that garnered the attention of other players when pithy quotes regarding NK were posted. Provoking this sort of reaction became an objective of sorts, with some fairly outlandish actions being executed as much to get a response from the crowd, as to advance the game toward a NK objective. An example of this “influencing the room” was the frequent rhythmic clapping by the players at the North Korean table, which was taken as a sign to other players of something significant happening. Once this was observed by the NK players, they would frequently just start clapping to observe the impact on the other players in the room. Is there an insight there into an analogy between gameplay and the real world (strategic signaling of one sort or another?), or was this just an artifact of everyone being in the same room having a bit fun with each other. Its unclear.

US White House Subgame

This was a combined role playing and card based game with various key positions within the White House represented, but with the key difference that the US President was represented by a “Snakes and Ladders” like board onto which players placed policy ideas and attempted to move them up the board until they became policy again using cards. Players could interfere with the progress of other team member’s ideas, and gained or lost points by how well their ideas became policy. Players who lost too many points were fired and replaced. A Chief of Staff acted as gate keeper to the game board that represented the President and he or she was also at risk of being fired (and indeed this happened during the game). The US President was represented by a set of cards based on past real tweets which would get played and adjudicated by the game designer. The stated reason for representing the US President in this manner was to avoid “excessively crazy (or reasonable) behavior” which “might adversely affect the entire game.” The conflicting objectives they were each given appears on the surface to be realistic and generated the expected behavior: the players mostly focused on winning the

internal game (else they could not get policy implemented) with occasional attempts to generate consensus.

There are two major problems that the analyst allocated to the White House subgame believes invalidates any claimed correlation between how the White would behave in a real crisis and how the game proceeded. First is a conflict of interest problem. The game designer of the White House subgame is well known to have strong hostile opinions of the US President but nevertheless selected how the President was portrayed in the cards and design of the game board, was also the prime adjudicator of how the “tweet cards” affected policy and policy ideas, and interpreted “tweets” as policy. Second, the game designer on several occasions modified the scoring mechanism that decided who got fired and replaced within the White House subgame. There was little evidence of the players paying attention to the North Korea subgame or other country players.

The macro level game behaviors were “baked in” to the game by the design and the game designer acting as adjudicator. These behaviors might be what one could expect during a real crisis but if so they were an input to the game, not an emergent behavior one can claim as an insight.

Did the Game Structure Support Answering the Central Question?

This gets at a fundamental issue attempting to apply analytic techniques that are used in wargames that are structured and facilitated with data collection to answer specific analytic questions in mind, to experiential, narrative games. The events in such a game can be reconstructed to a certain extent (the game artifacts ordered and correlated) but that is a long way from deciding “the validity of the wargame methodology adopted and the substantive findings of the game”. Given the central question is somewhat “baked into the game” (through the use of a somewhat capricious method of policy adoption by President Trump as represented by a “chutes and ladders” style game board), together with a representation of Chinese strategic power that was a matter of interpretation by whichever player team they were trying to influence.

The opaque nature of the connection between the placement of markers on the military situation map made assessing the strategic power of China a matter of the Chinese player's personality and engagement style with other players, than a function of a strategic correlation of respective forces or economic strength. This made diplomacy (and to some extent in-game information exploitation) the primary expression of power, thus a function of individual player personality, than the game constructs itself. For an experiential narrative game this is definitely a feature, but does pose a problem for validity and substantiation of findings that are derived from the player's experience playing the game, and corresponding ability to draw effective analogy between their gameplay and the real world.

For this reason, asking about the validity of the wargame methodology is not appropriate to a game that is not designed around answering specific analytic questions. Without those questions, there is no reference frame for assessing what is "valid" and what is not. Was the game methodology effective at providing insight into the central question? Each player will take away an individual assessment of that based on their particular game-play experience. Rather than analysis teams attempting to collect information in the manner of an analytic seminar-style game, a post-game questionnaire about what the players felt enhanced their ability to engage within the "magic circle" and remain in the flow of the game, and what things broke their willing suspension of disbelief would be valuable. Experiential games require players to recount the highlights of their experience to provide a basis for assessing a games methodology. These answers will not be found in a detailed reconstruction of the series of events in the game.

Game Designer's Claims versus Analysts' Counter-Arguments

As to assessing the substantive findings of the game, again, the answer lies more in the perception of the players, rather than data collected by observation. The broader lessons taken away by one of the game designers² follow along with counter-arguments by the analysts:

Game Designer’s Claims	Analysts’ Counter-Arguments
<p>“US policy under the Trump Administration is much less predictable than under any other president in modern times, a function of both the President’s mercurial and populist political instincts, and the clash between differing priorities and world-views within the White House. True, we had designed the game system to encourage this, but none of it was predetermined, and players could have taken a more cooperative route (as they did when deciding to increase the American investment in ABM systems). As White House Control, I was pleased to see how realistically and enthusiastically participants role-played their roles. Debate centered around different political views and goals, and not the manipulation of game mechanics. Domestic political concerns often trumped geopolitics. In short, if one builds a game system that models the existence of factions, rivalries, and differences within the current White House, one gets game outputs that look very much like current US foreign policy.”</p>	<p>Despite the assertion that this conclusion was not predetermined, there was no perception by the NK players of “unpredictability” on the part of the US president, quite the contrary the tweets that resulted from NK instigation were entirely predictable. The text of at least some of the tweets was provided by cards that were part of the game design, with intent to drive gameplay in certain directions. That is a valid technique used in many seminar-style wargames, but the injects can’t then be turned around and declared to be insights. This insight was not emergent from the gameplay. That players assumed their role and enthusiastically pursued appropriate goals, is a mark of the level of interest of the players in the gameplay, not an indication that the game system is particularly good or bad at depicting specific inter-relationships between any other set of players (let alone real world personalities) or being susceptible or not to being “gamed”. The final statement about the game system producing outputs that look like current foreign policy based on the impact of “factions, rivalries, and differences” is true of any political staff. Since we have not seen the current administration deal with the type of crises depicted in the game, it is pure speculation whether they reflect an accurate extrapolation from current policy.</p>

Game Designer’s Claims	Analysts’ Counter-Arguments
<p>“The mixed and sometimes wildly oscillating signals coming out of Washington do less damage than might be the case because they are quietly spun, nuanced, and moderated by cabinet officials and ambassadors in the field. In DIRE STRAITS the Secretary of Defense, Chairman of the Joint Chiefs of Staff, Secretary of State, and various ambassadors played a key role in this. Indeed, it was precisely because he spent so much time trying to patch over problems arising in Washington that our simulated Secretary of State found himself with little influence in the Oval Office and was ultimately sacked.”</p>	<p>Once again, the assumption that “mixed and wildly oscillating signals...do less damage” is again a speculation that is not supported through observed gameplay, or was a part of the game design. We have little actual knowledge of damage being done by the manner in which the current White House influences the domestic news cycle on the effectiveness of foreign policy objectives. The signaling that goes on through military, diplomatic and intelligence channels is historically “spun, nuanced, and moderated” to some degree or another by the career diplomats to protect the generally inexperienced political appointees that serve as Ambassadors and delegation heads in other countries. That the players, when put into a chaotic situation, tried to mitigate the adverse effects of the chaos, is a specific insight particular to the current White House is not supported by game observations from the NK subgame perspective.</p>
<p>“Despite this, uncertainties in US policy generate anxiety among American friends and allies. Neither South Korea nor Japan seemed to feel they could fully rely on Washington, as evidenced by the former secret decision to prepare a potential nuclear weapons programme. Taiwan was never quite sure how much latitude and US support it had, and Beijing was also left guessing about American commitment to the “One China Policy.” ASEAN countries increased regional security cooperation in part because US backing seemed uncertain. Several countries diversified their relations to counterbalance China and hedge their bets regarding American support.”</p>	<p>This is a generality that is simply a fact of life in a multi-polar competitive international regime. Ambiguity, mixed messaging, mis-interpretation of what were thought to be clear signals, and countries acting in their own self-interest to acquire the weapons they deem required are not unique to this game, or to the current administration’s policy dynamics. The game may have made these points in a manner that caused the players to understand the manner in which these uncertainties arise, from experiencing the frustrations first hand, but that the frustrations are experienced, is not a basis for inferring that the game specific narrative resulting from those pathologies has any significant analogy to real world policy decisions.</p>

Game Designer’s Claims	Analysts’ Counter-Arguments
<p>“The game clearly showed that there are no good policy options regarding North Korea’s nuclear capacity, only less-bad ones. Everyone was wary of pushing Pyongyang too far. Toppling the Kim Jong-un regime was seen by most (but not all) as dangerous, since it risked retaliation or chaos in a nuclear-armed state. In this sense, Pyongyang’s nukes demonstrated their value as a deterrent. Rather than punitive strikes or intervention, a messy mix of threats, deterrence, sanctions, and diplomatic dialogue appeared to offer the best path to crisis management. US-Chinese cooperation was important, but undermined by mutual suspicion, as well as tensions between Washington and Beijing on other issues (such as trade or the South China Sea). Overall, the game seemed to suggest no meaningful path to denuclearization, a real risk that South Korea (or even Japan) might consider a future nuclear weapons option, and the reality of having to live with a nuclear-armed DPRK while mitigating the threat and deterring North Korean adventurism.”</p>	<p>The experience of being faced with a no-win situation, within an afternoon of highly time-constrained information gathering and decision-making, as mentioned in the paragraph above, can be a valuable educational experience, but to claim that a busy chaotic afternoon of deliberations by non-expert policy analysts “suggest no meaningful path to denuclearization” is considerable over-reach.</p>
<p>“DIRE STRAITS ended with a few incidents at sea over illegal fishing and a some major cyber-attacks, but no open warfare.”</p>	<p>The PACOM player covertly attempted to sink an NK submarine but failed. The player portraying dear leader of NK was aware of the attempt. The game directors did not appear to know about this act of kinetic warfare. This also demonstrates the challenge of mega games to keep track of information flowing in the game.</p>

Game Designer’s Claims	Analysts’ Counter-Arguments
<p>“It was noteworthy than many/most non-American players were unaware that “Donald Trump” was a game system rather than a human player until after it was all over.”</p>	<p>The game designer of the White House subgame selected how the President was portrayed in the cards and design of the game board, was the prime adjudicator of how the “tweet cards” affected policy and policy ideas, interpreted “tweets” into policy, and on several occasions modified the scoring mechanism that decided who got fired and replaced within the White House subgame. “Donald Trump” was to all practical purposes played by the adjudicator, guided but not directed by the “board game and cards”.</p>
<p>“As someone who is arguably an SME in crisis stability and foreign policy, and who spent years doing cabinet-level strategic threat assessment, I’m happy to assert that one doesn’t need to be an SME in the subject to make some useful comments about whether a game like this offers insights in this area. In fact, our internal work on the quality of strategic assessment product highlights Tetlock’s findings on the limits of SMEs and the value of diverse teams of fox-like thinkers.”⁶</p>	<p>No evidence was provided that the random selection of the world’s wargamers who participated in the Connections UK 2017 Conference consisted of fox-like thinkers. The hypothesis that they were in fact hedgehog-like thinkers focused on “how to design and play board-games” cannot a priori be dismissed. Since decisions made in a game rely on the player’s ability to forecast the likely impact of their decisions then Tetlock’s more recent work into forecasting must also be considered along with his earlier work on expert judgement. This later work implies that active training in how to forecast is necessary, and no evidence was provided that the players had that training.</p>

Conclusion

The game design overall provides an evocative venue for giving players a small taste of the elephant that is the real world of global politics and policy formation. It lacks, however, a set of well-defined analytic questions, a data collection package to inform formulating answers to those questions, and an overarching control mechanism to ensure player interactions (and tier motivations) are captured in each of the country cells. This level of preparation and data

⁶ The game designer made this claim in an email to the US Analysis Team dated August 11, 2017.

collection would require a substantial organization to plan and execute. To provide access to the player cells, and photos of game board configurations doesn't provide the information required to judge the applicability of actions and decisions made in the game to any real-life policy deliberations. Without a considerable effort to apply rigorous game design process, formulation of analytic questions, and a reasonably controlled environment for data collection by rapporteurs (the players wandering hither and yon to find each other and talk face to face poses an extreme challenge in this regard by creating unmanageably interactive complexity). A set of very carefully crafted player questionnaires can provide some of this data in an informal manner, but as currently implemented, the mega-game format is highly entertaining, provides a valuable experiential appreciation of confusion, miscommunication, stovepiped and insular intrateam dynamics, and ultimately the difficulty "doing this sort of thing for real", but even with players expert in policy formulation, diplomacy, and intelligence matters, the time-constraints and information limitations prevent this type of game from being used as a serious analytic gaming method.

Summary

Game Design Findings: The design achieved the objective of providing participants with an enjoyable gaming and social experience and an introduction (for those who needed it) with a collection of interesting game dynamics.

1. The game design lacked a set of well-defined analytic questions, a data collection package to inform formulating answers to those questions, and an overarching control mechanism to ensure player interactions (and tier motivations) were captured in each of the country cells.
2. There was no evidence that the selection of the world's wargamers who played were experts in the subject matter of the game, neither was there any evidence they were fox-like thinkers or trained forecasters which would be a required mitigation.
3. The designer of the White House subgame is well known to be hostile to President Trump and chose to be the adjudicator of the White House thus creating a conflict of interest which drove the behaviors of the White House and the NK subgames (both how they behaved and how they ignored each other) to be predictable from the design.

4. Competition between players within the White House subgame, while perhaps realistic in isolation, overwhelmed any signaling between the White House cell and the rest of the game due to the lack of any other component of the US Government.
5. The lack of adequate real-time data collection generated a lack of understanding by the game designers/directors of what was actually happening during the game.

Subject Matter Findings: The design was inappropriate for an analytical game, and thus generated no subject matter findings that could be credibly attributed to the game.