Designing for Validity

A Methodology for Professional Wargame Design

Disclaimer

The opinions of the speakers do not represent the opinions of the U.S. Army, U.S. Army War College, or any other body of the U.S. Government.

We are using the word “Validity” in a colloquial sense, not the narrow sense it is used in in science.

Introductions

Josh Kovan is an instructional wargaming fellow at the U.S. Army War College and a non-resident fellow at the Irregular Warfare Initiative.

Evan D’Alessandro is a PhD student at King’s College London studying Immersion in Professional Wargaming and also works as a freelance professional wargamer. A portfolio of free games can be found at https://evandalessandro.com/

Commercial Disease

We use commercial games as if they are professional games

There are inherent differences in purpose of the game and the nature by which they are designed which means commercial wargames are not fit for professional use

We need to differentiate between design methodologies for commercial and professional wargames

When Worlds Collide

The commercial wargaming industry has extensive overlap with professional wargaming

-Common Members

-Common Lineage

-Common Practices

Commercial Design Practices

<u>Top Down</u>

1. Theme/Topic
2. Core Rules
3. Prototype Components
4. Iterate
5. Transition to development

<u>Bottom Up</u>

1. Theme/Topic (Maybe)
2. Prototype Components
3. Iterate
4. Write core rules
5. Transition to development

Commercial design is whatever you want it to be

Professional Design Practices

Professional Process:

1. Requirements
2. Research
3. Initial Design
4. Revision/Playtesting
5. Execution/Delivery
6. Analysis/Educational Lesson

Professional design is conducted linearly to maintain validity

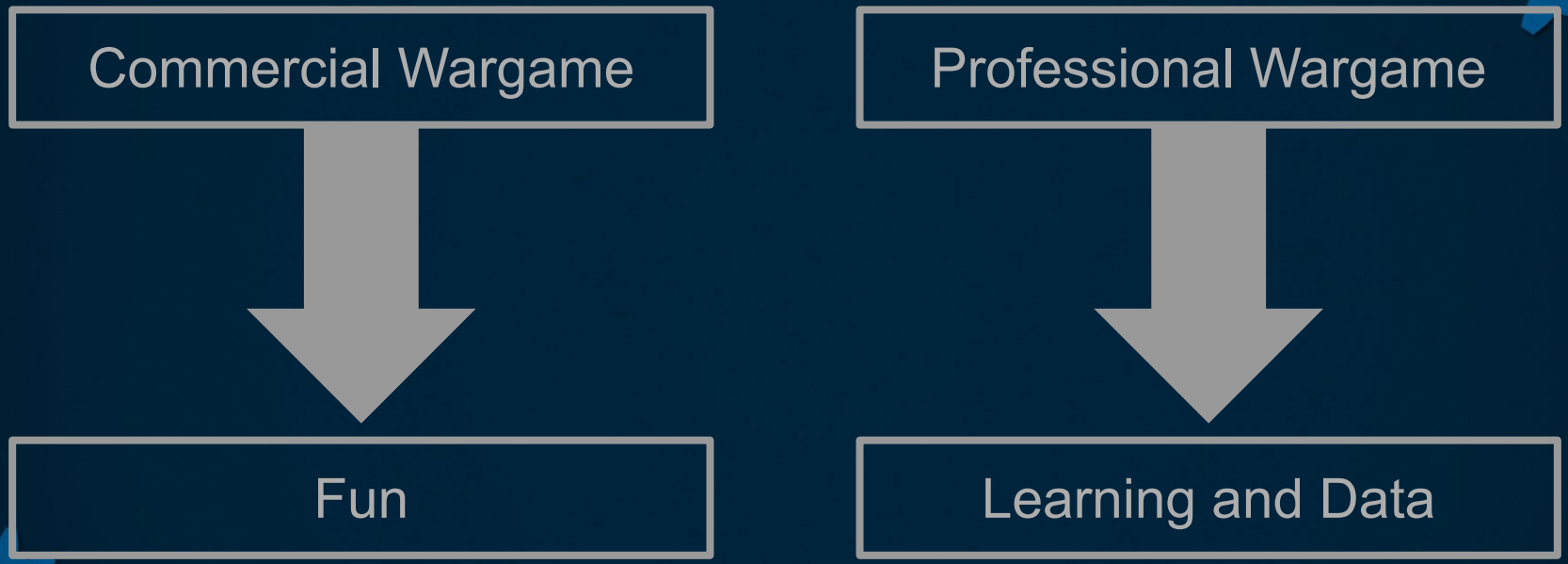
Differences in Practice and Purpose Lead to Issues with Using Commercial Games

Commercial wargames are not fit for professional use:

1. Complex
   a. Not designed with the beginner in mind, even if it is, it often has a lot of assumed knowledge baked in
2. Typically low reflexivity on the game
   a. Poor designers notes, sourcing, etc.. – see our other presentation!
3. Internal model is opaque
4. Often designed by non-SME's
   a. Often pop history creeps in, or systems are not well understood
5. Designed with fun as the primary goal, no representativeness or education.
6. Will never meet your purpose out of the box [1]

[1] Blockbuster: Dstl Urban Manual Wargame COTS Assessment, Paul Beaves, and discussion with UK Wargamers on other COTS Assessments

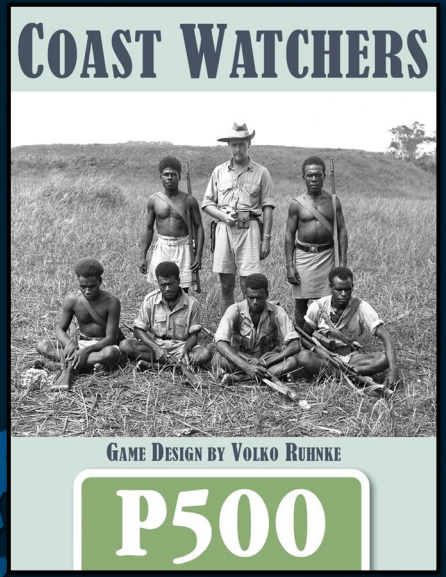
Apples and Oranges



Commercial games maximize fun at the cost of education/analysis

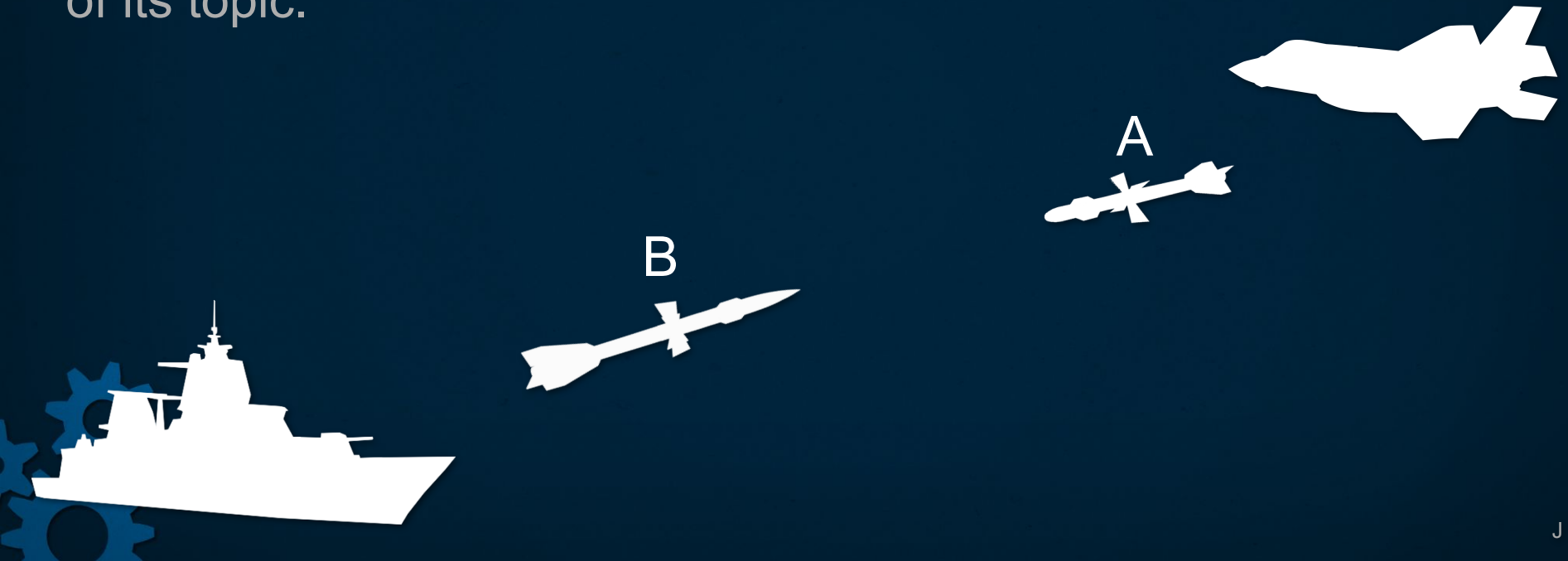
I Think, Therefore I Wargame

- All wargame designs are guided by the designer's biases and expertise
- Those will determine the degree of representation



What is Representation?

The degree to which a wargame accurately models the realities of its topic.



Representation in Commercial Wargames

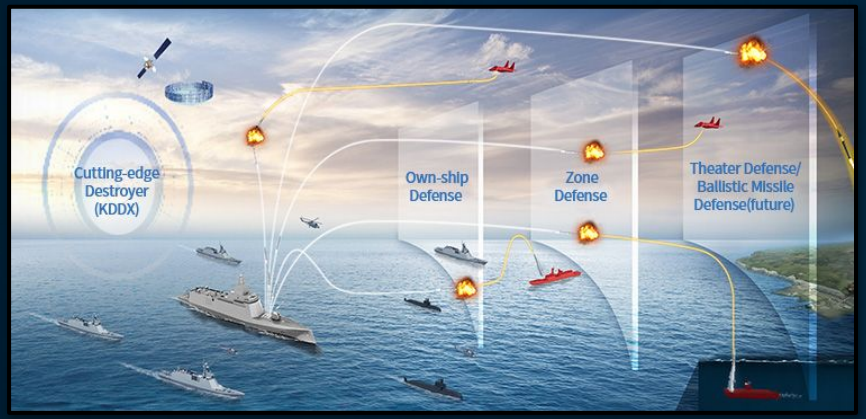
WARHAMMER 40,000

Expectation

Reality

The design/experience does not match the theme

Representation in Professional Wargames

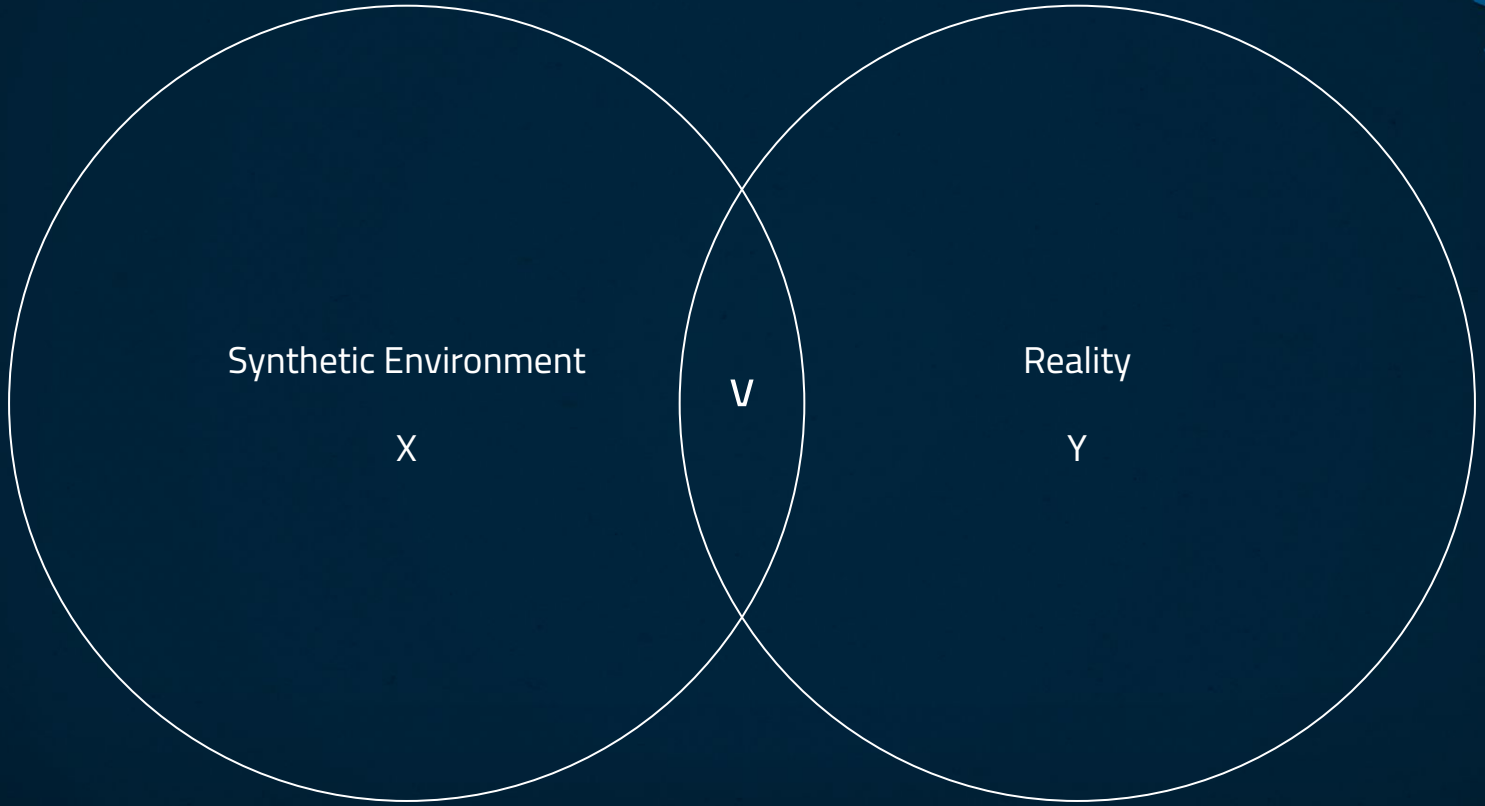


Expectation

Reality

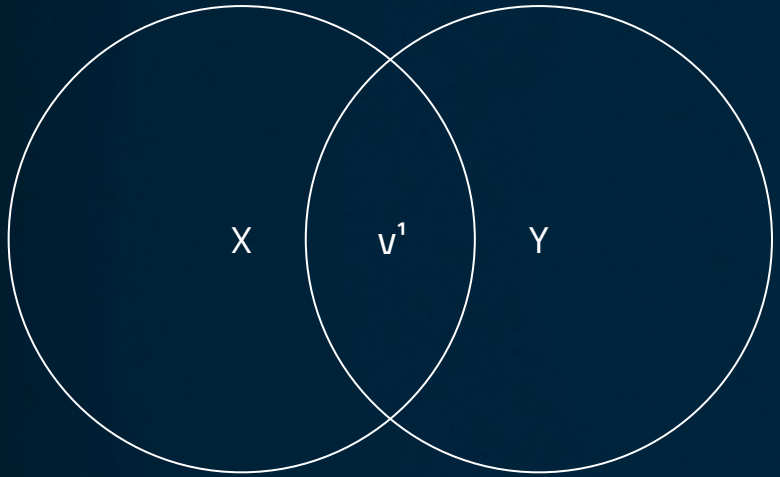
Bad Design = Bad Decisions = Lives Lost

How it Works (or Doesn’t)

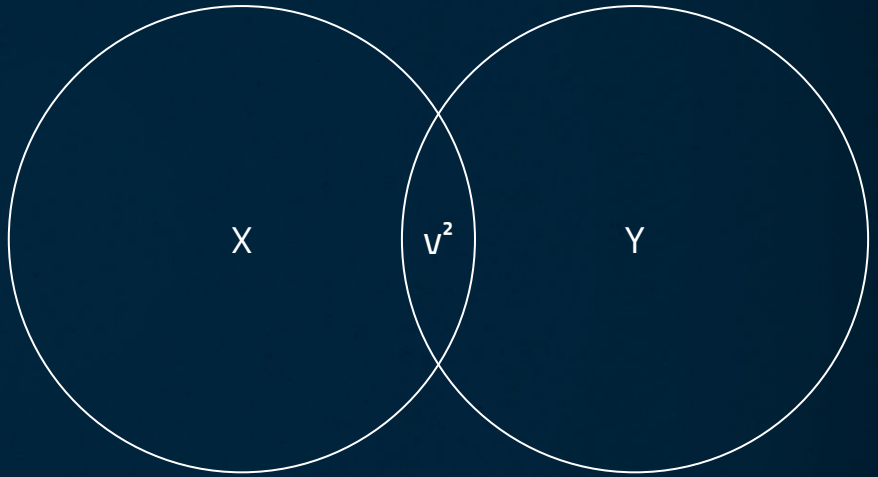


False Utility in Wargames

Perceived Utility



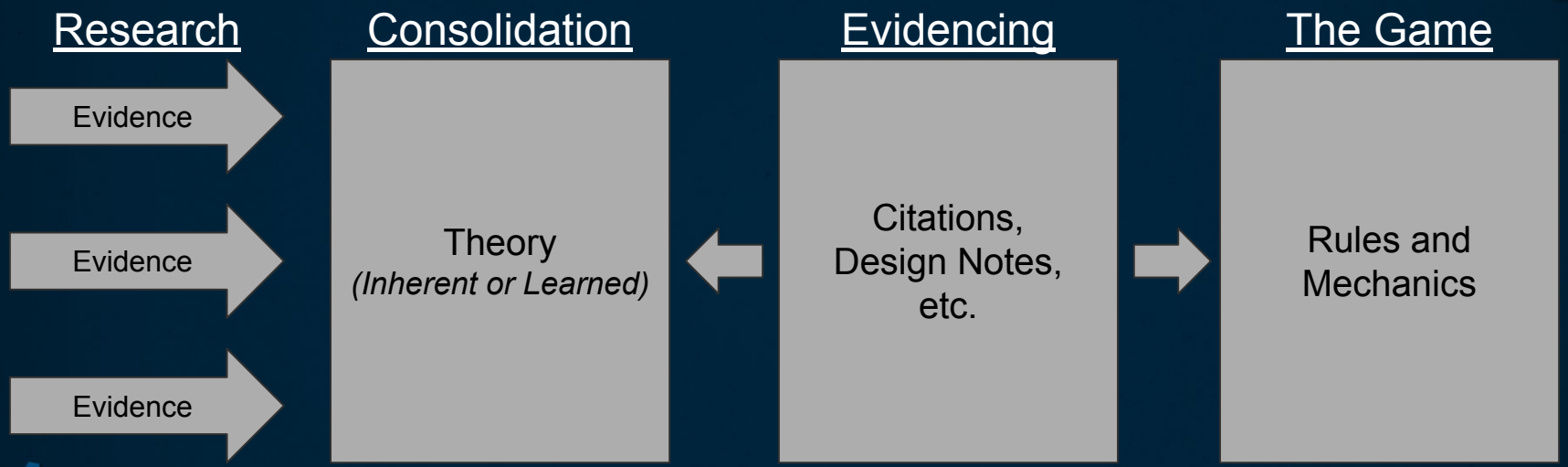
Actual Utility



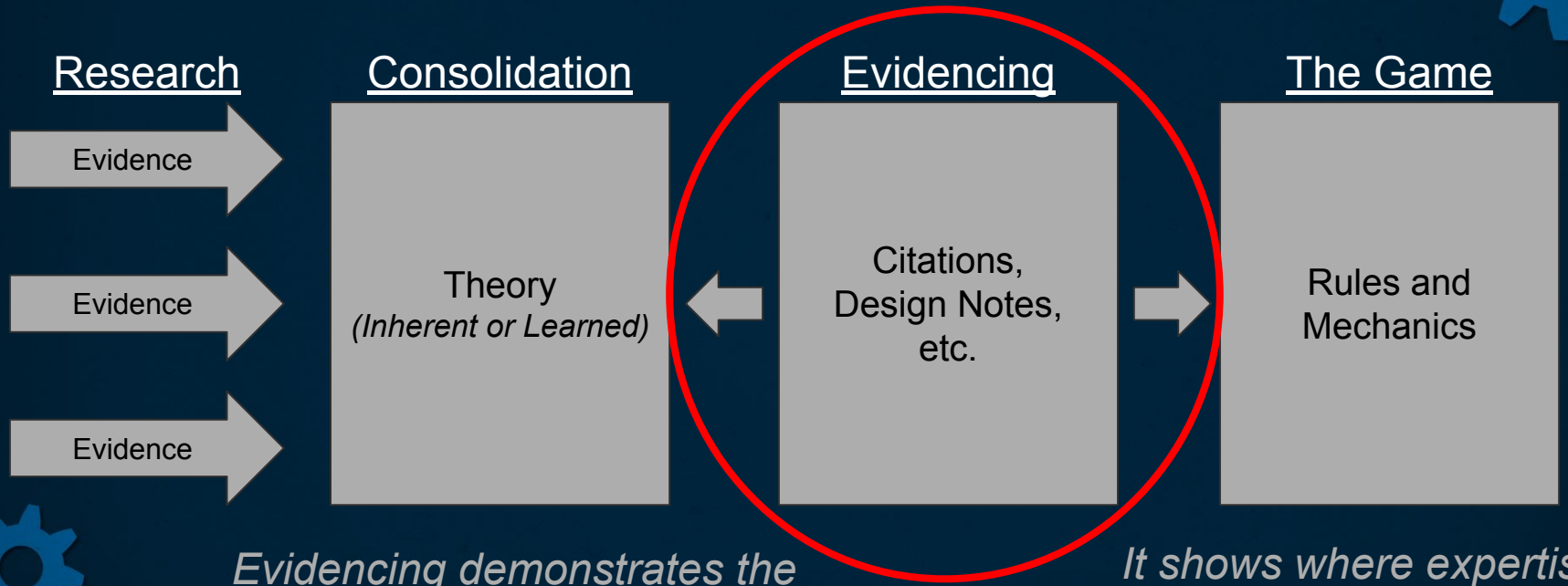
$v^1 \neq v^2$

ΔV= False Utility

The Methodology



The Methodology



Evidencing demonstrates the degree to which expertise is applied to design decisions

It shows where expertise is lacking and where design decisions are biased

How do We Do This?

Designing for validity is not about designing for validity, it's about writing so the design can be validated.

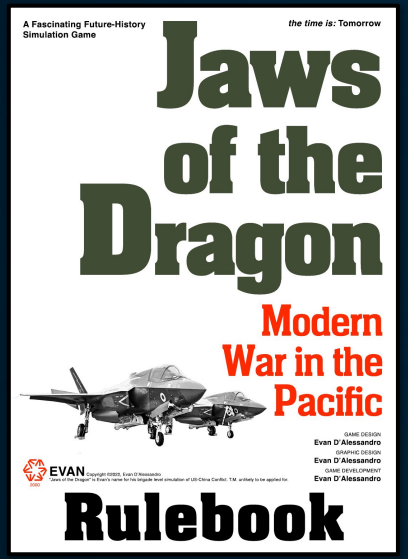
How to you Acquire Evidence?

1. Be a Subject Matter Expert (takes a long time, and a lot of work)
2. Become knowledgeable (read as much as you can possibly get your hands on)
3. Talk to SME's (typically at the start or towards the end, to get a general idea of the system, or to check details/the game matches to the system)

Generally Professional Wargamers use a mix of 2 and 3 when designing games.

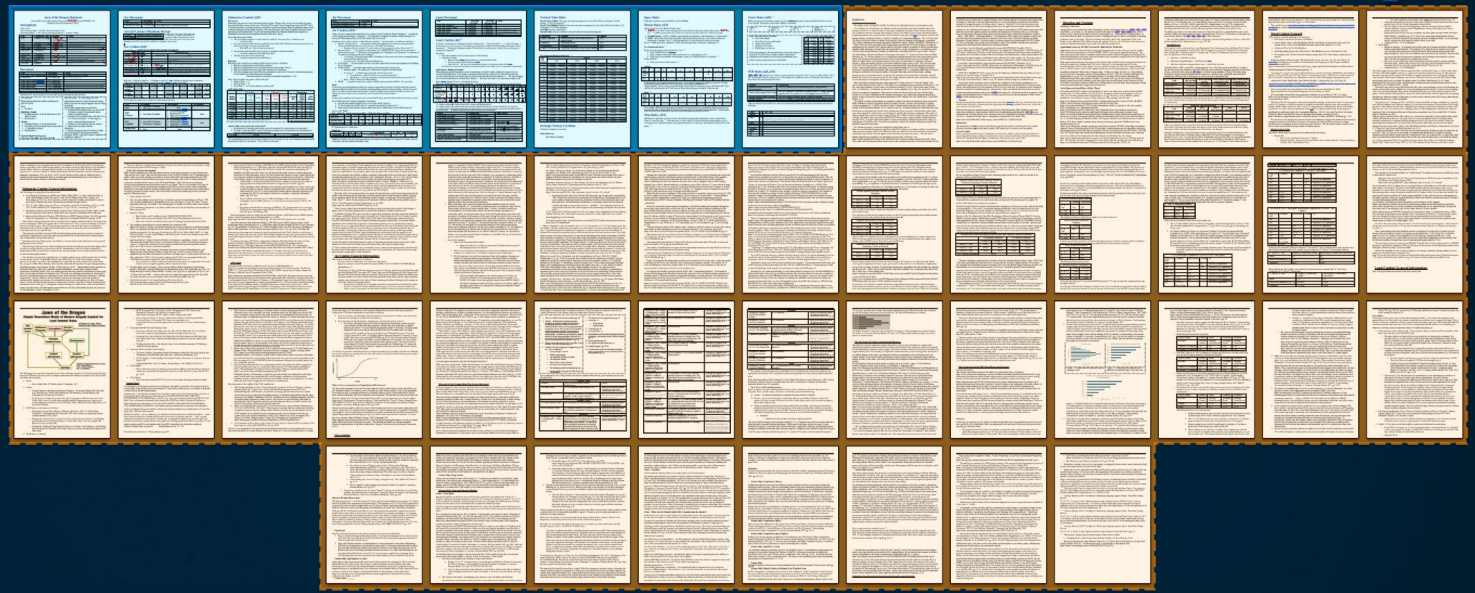
Sometimes you can bring a SME into the design team (getting the benefit of 1 and 3.

How Much Research is Good Enough? - An Example



Rules

8 Pages, ~4,500 words



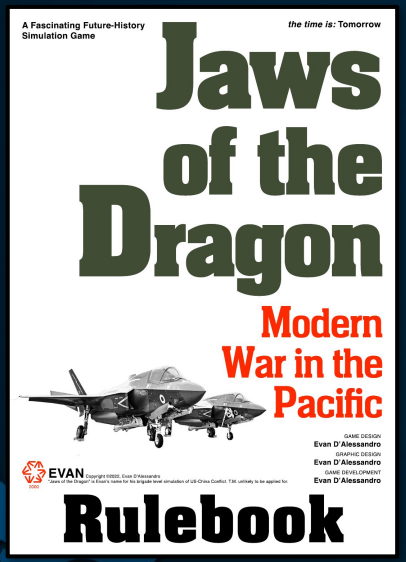
Endnotes

42 Pages, ~25,000 words, 151 unique sources

How Much Research is Good Enough?

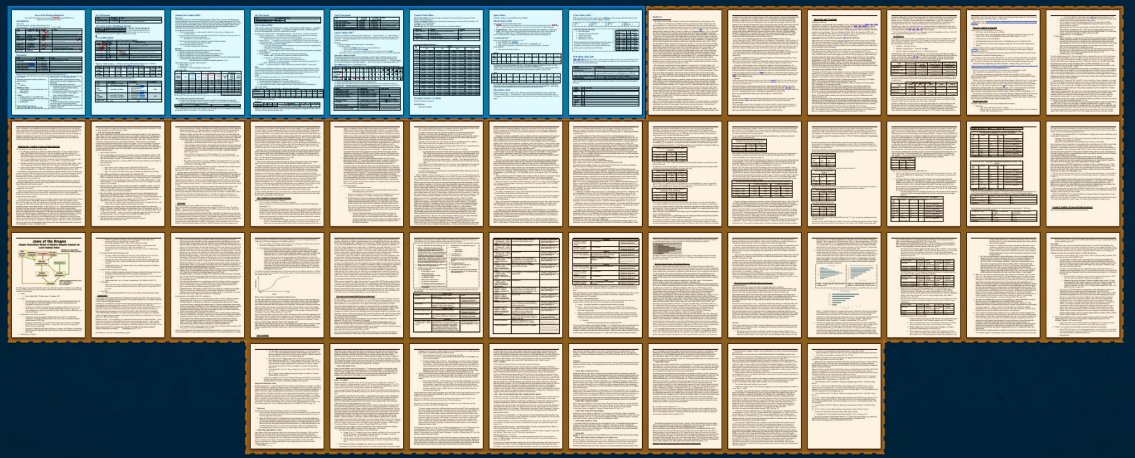
This is the result of hundreds of hours of reading and discussion over 3 years.

Is it good enough? - Not for me (but I don't owe it to anyone right now)



Rules

8 Pages, ~4,500 words



Endnotes

42 Pages, ~25,000 words, 151 unique sources

Remaining Sources to Read	
Air Warfare	30
Cyber	11
"Cognitive"	8
Detection	8
Land Warfare	62
Logistics	12
Naval Warfare	18
Space	2
Other	16

Is there Enough Research to be Good Enough?

Games are not finished, they are run.

You need to have done enough research that the purpose of the game can been achieved.

Enough Research - Rules of Thumb

- *A Good Guideline* – if this game/part of a game was a research paper, how much research would I need to do?
  - Undergrad Paper – Part of a game that has some importance
  - Masters Thesis – Important game
  - phd: fate of the world
- *When to Do More Work* – When you don't understand the debate or historiography around the topic.
  - If you can't explain the complexity of the topic, you don't understand it.
  - You should understand all the arguments of the sides, know your side, who you agree/disagree with, and to what degree you do.
  - Avoid Dunning-kruger

Enough Research - What to do with Imperfect Knowledge?

- Dealing with imperfect knowledge (because it will occur):
  - Flag it to sponsors and players
  - Talk about it in debrief
  - Make the system transparent so others can decided if it's good enough for them!
- None of this is an excuse to do poor/little research!

Showing your Work - How do you Make a Game Transparent?

You show your research and your implementation of the research into the game by writing it down in a place where others can see it!

Evidencing - Structure of the Evidence

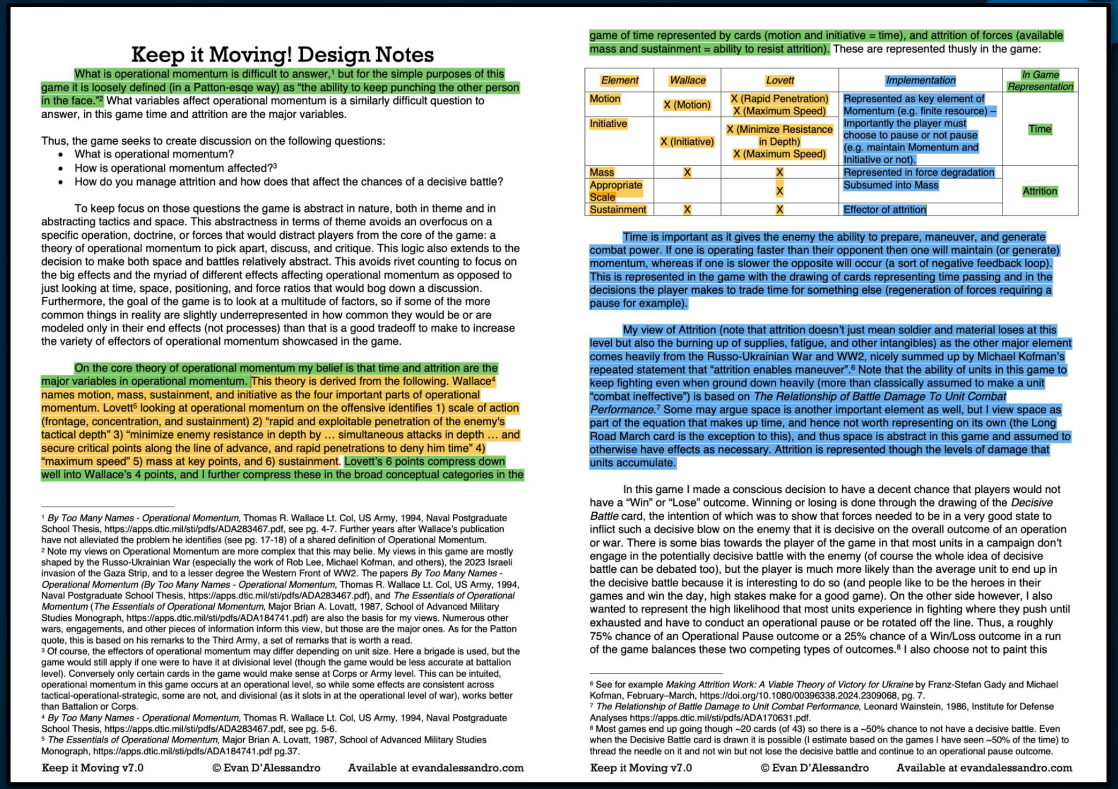
1) Explain the overall theory - what are the elements that make up the system

2) Explain how individual elements of the system function

3) Explain how elements and the way in which the elements function are implemented in the game

- Biases disclosed throughout the text

- Sources and explanations footnoted



Keep it Moving! Design Notes

What is operational momentum is difficult to answer,[1] but for the simple purposes of this game it is loosely defined (in a Patton-esqe way) as "the ability to keep punching the other person in the face."[2] What variables affect operational momentum is a similarly difficult question to answer, in this game time and attrition are the major variables.

Thus, the game seeks to create discussion on the following questions:

- What is operational momentum?
- How is operational momentum affected?[3]
- How do you manage attrition and how does that affect the chances of a decisive battle?

To keep focus on those questions the game is abstract in nature, both in theme and in abstracting tactics and space. This abstractness in terms of theme avoids an overfocus on a specific operation, doctrine, or forces that would distract players from the core of the game: a theory of operational momentum to pick apart, discuss, and critique. This logic also extends to the decision to make both space and battles relatively abstract. This avoids rivet counting to focus on the big effects and the myriad of different effects affecting operational momentum as opposed to just looking at time, space, positioning, and force ratios that would bog down a discussion. Furthermore, the goal of the game is to look at a multitude of factors, so if some of the more common things in reality are slightly underrepresented in how common they would be or are modeled only in their end effects (not processes) than that is a good tradeoff to make to increase the variety of effectors of operational momentum showcased in the game.

On the core theory of operational momentum my belief is that time and attrition are the major variables in operational momentum. This theory is derived from the following. Wallace[4] names motion, mass, sustainment, and initiative as the four important parts of operational momentum. Lovett[5] looking at operational momentum on the offensive identifies 1) scale of action (frontage, concentration, and sustainment) 2) "rapid and exploitable penetration of the enemy's tactical depth" 3) "minimize enemy resistance in depth by … simultaneous attacks in depth … and secure critical points along the line of advance, and rapid penetrations to deny him time" 4) "maximum speed" 5) mass at key points, and 6) sustainment. Lovett's 6 points compress down well into Wallace's 4 points, and I further compress these in the broad conceptual categories in the game of time represented by cards (motion and initiative = time), and attrition of forces (available mass and sustainment = ability to resist attrition). These are represented thusly in the game:

Element	*Wallace*	*Lovett*	*Implementation*	*In Game Representation*
Motion	X (Motion)	X (Rapid Penetration) X (Maximum Speed)	Represented as key element of Momentum (e.g. finite resource) – importantly the player must choose to pause or not pause (e.g. maintain Momentum and Initiative or not).	Time
Initiative	X (Initiative)	X (Minimize Resistance in Depth) X (Maximum Speed)		
Mass	X	X	Represented in force degradation	Attrition
Appropriate Scale		X	Subsumed into Mass	
Sustainment	X	X	Effector of attrition	

Time is important as it gives the enemy the ability to prepare, maneuver, and generate combat power. If one is operating faster than their opponent then one will maintain (or generate) momentum, whereas if one is slower the opposite will occur (a sort of negative feedback loop). This is represented in the game with the drawing of cards representing time passing and in the decisions the player makes to trade time for something else (regeneration of forces requiring a pause for example).

My view of Attrition (note that attrition doesn't just mean soldier and material loses at this level but also the burning up of supplies, fatigue, and other intangibles) as the other major element comes heavily from the Russo-Ukrainian War and WW2, nicely summed up by Michael Kofman's repeated statement that "attrition enables maneuver".[6] Note that the ability of units in this game to keep fighting even when ground down heavily (more than classically assumed to make a unit "combat ineffective") is based on *The Relationship of Battle Damage To Unit Combat Performance*.[7] Some may argue space is another important element as well, but I view space as part of the equation that makes up time, and hence not worth representing on its own (the Long Road March card is the exception to this), and thus space is abstract in this game and assumed to otherwise have effects as necessary. Attrition is represented though the levels of damage that units accumulate.

In this game I made a conscious decision to have a decent chance that players would not have a "Win" or "Lose" outcome. Winning or losing is done through the drawing of the *Decisive Battle* card, the intention of which was to show that forces needed to be in a very good state to inflict such a decisive blow on the enemy that it is decisive on the overall outcome of an operation or war. There is some bias towards the player of the game in that most units in a campaign don't engage in the potentially decisive battle with the enemy (of course the whole idea of decisive battle can be debated too), but the player is much more likely than the average unit to end up in the decisive battle because it is interesting to do so (and people like to be the heroes in their games and win the day, high stakes make for a good game). On the other side however, I also wanted to represent the high likelihood that most units experience in fighting where they push until exhausted and have to conduct an operational pause or be rotated off the line. Thus, a roughly 75% chance of an Operational Pause outcome or a 25% chance of a Win/Loss outcome in a run of the game balances these two competing types of outcomes.[8] I also choose not to paint this

[1] *By Too Many Names - Operational Momentum*, Thomas R. Wallace Lt. Col, US Army, 1994, Naval Postgraduate School Thesis, https://apps.dtic.mil/sti/pdfs/ADA283467.pdf, see pg. 4-7. Further years after Wallace's publication have not alleviated the problem he identifies (see pg. 17-18) of a shared definition of Operational Momentum.

[2] Note my views on Operational Momentum are more complex that this may belie. My views in this game are mostly shaped by the Russo-Ukrainian War (especially the work of Rob Lee, Michael Kofman, and others), the 2023 Israeli invasion of the Gaza Strip, and to a lesser degree the Western Front of WW2. The papers *By Too Many Names - Operational Momentum (By Too Many Names - Operational Momentum*, Thomas R. Wallace Lt. Col, US Army, 1994, Naval Postgraduate School Thesis, https://apps.dtic.mil/sti/pdfs/ADA283467.pdf), and *The Essentials of Operational Momentum* (*The Essentials of Operational Momentum*, Major Brian A. Lovatt, 1987, School of Advanced Military Studies Monograph, https://apps.dtic.mil/sti/pdfs/ADA184741.pdf) are also the basis for my views. Numerous other wars, engagements, and other pieces of information inform this view, but those are the major ones. As for the Patton quote, this is based on his remarks to the Third Army, a set of remarks that is worth a read.

[3] Of course, the effectors of operational momentum may differ depending on unit size. Here a brigade is used, but the game would still apply if one were to have it at divisional level (though the game would be less accurate at battalion level). Conversely only certain cards in the game would make sense at Corps or Army level. This can be intuited, operational momentum in this game occurs at an operational level, so while some effects are consistent across tactical-operational-strategic, some are not, and divisional (as it slots in at the operational level of war), works better than Battalion or Corps.

[4] *By Too Many Names - Operational Momentum*, Thomas R. Wallace Lt. Col, US Army, 1994, Naval Postgraduate School Thesis, https://apps.dtic.mil/sti/pdfs/ADA283467.pdf, see pg. 5-6.

[5] *The Essentials of Operational Momentum*, Major Brian A. Lovatt, 1987, School of Advanced Military Studies Monograph, https://apps.dtic.mil/sti/pdfs/ADA184741.pdf pg.37.

[6] See for example *Making Attrition Work: A Viable Theory of Victory for Ukraine* by Franz-Stefan Gady and Michael Kofman, February–March, https://doi.org/10.1080/00396338.2024.2309068, pg. 7.

[7] *The Relationship of Battle Damage to Unit Combat Performance*, Leonard Wainstein, 1986, Institute for Defense Analyses https://apps.dtic.mil/sti/pdfs/ADA170631.pdf.

[8] Most games end up going though ~20 cards (of 43) so there is a ~50% chance to not have a decisive battle. Even when the Decisive Battle card is drawn it is possible (I estimate based on the games I have seen ~50% of the time) to thread the needle on it and not win but not lose the decisive battle and continue to an operational pause outcome.

Keep it Moving v7.0 © Evan D'Alessandro Available at evandalessandro.com

Showing Your Work - What Arguments you Need

		Research Argues:	
		Element is NOT Relevant	Element is Relevant
Game Element	Element Included in Game	Explanation must be given of why *and* how the element was included despite it not matching research	Explanation on how the element was implemented in game
	Element not included in game	No explanation needed	Explanation must be given of why the element was left out despite research stating it was important

This diagram is drawn from Dr. David Banks (and hopefully should be published soon)

Types of Evidencing - Design Notes

Keep it Moving! Design Notes

What is operational momentum is difficult to answer,[1] but for the simple purposes of this game it is loosely defined (in a Patton-esqe way) as "the ability to keep punching the other person in the face."[2] What variables affect operational momentum is a similarly difficult question to answer, in this game time and attrition are the major variables.

Thus, the game seeks to create discussion on the following questions:

- What is operational momentum?
- How is operational momentum affected?[3]
- How do you manage attrition and how does that affect the chances of a decisive battle?

To keep focus on those questions the game is abstract in nature, both in theme and in abstracting tactics and space. This abstractness in terms of theme avoids an overfocus on a specific operation, doctrine, or forces that would distract players from the core of the game: a theory of operational momentum to pick apart, discuss, and critique. This logic also extends to the decision to make both space and battles relatively abstract. This avoids rivet counting to focus on the big effects and the myriad of different effects affecting operational momentum as opposed to just looking at time, space, positioning, and force ratios that would bog down a discussion. Furthermore, the goal of the game is to look at a multitude of factors, so if some of the more common things in reality are slightly underrepresented in how common they would be or are modeled only in their end effects (not processes) than that is a good tradeoff to make to increase the variety of effectors of operational momentum showcased in the game.

On the core theory of operational momentum my belief is that time and attrition are the major variables in operational momentum. This theory is derived from the following. Wallace[4] names motion, mass, sustainment, and initiative as the four important parts of operational momentum. Lovett[5] looking at operational momentum on the offensive identifies 1) scale of action (frontage, concentration, and sustainment) 2) "rapid and exploitable penetration of the enemy's tactical depth" 3) "minimize enemy resistance in depth by … simultaneous attacks in depth … and secure critical points along the line of advance, and rapid penetrations to deny him time" 4) "maximum speed" 5) mass at key points, and 6) sustainment. Lovett's 6 points compress down well into Wallace's 4 points, and I further compress these in the broad conceptual categories in the game of time represented by cards (motion and initiative = time), and attrition of forces (available mass and sustainment = ability to resist attrition). These are represented thusly in the game:

Element	*Wallace*	*Lovett*	*Implementation*	*In Game Representation*
Motion	X (Motion)	X (Rapid Penetration) X (Maximum Speed)	Represented as key element of Momentum (e.g. finite resource) –	Time
Initiative	X (Initiative)	X (Minimize Resistance in Depth) X (Maximum Speed)	Importantly the player must choose to pause or not pause (e.g. maintain Momentum and Initiative or not).	
Mass	X	X	Represented in force degradation	Attrition
Appropriate Scale		X	Subsumed into Mass	
Sustainment	X	X	Effector of attrition	

Time is important as it gives the enemy the ability to prepare, maneuver, and generate combat power. If one is operating faster than their opponent then one will maintain (or generate) momentum, whereas if one is slower the opposite will occur (a sort of negative feedback loop). This is represented in the game with the drawing of cards representing time passing and in the decisions the player makes to trade time for something else (regeneration of forces requiring a pause for example).

My view of Attrition (note that attrition doesn't just mean soldier and material loses at this level but also the burning up of supplies, fatigue, and other intangibles) as the other major element comes heavily from the Russo-Ukrainian War and WW2, nicely summed up by Michael Kofman's repeated statement that "attrition enables maneuver".[6] Note that the ability of units in this game to keep fighting even when ground down heavily (more than classically assumed to make a unit "combat ineffective") is based on *The Relationship of Battle Damage To Unit Combat Performance*.[7] Some may argue space is another important element as well, but I view space as part of the equation that makes up time, and hence not worth representing on its own (the Long Road March card is the exception to this), and thus space is abstract in this game and assumed to otherwise have effects as necessary. Attrition is represented though the levels of damage that units accumulate.

In this game I made a conscious decision to have a decent chance that players would not have a "Win" or "Lose" outcome. Winning or losing is done through the drawing of the *Decisive Battle* card, the intention of which was to show that forces needed to be in a very good state to inflict such a decisive blow on the enemy that it is decisive on the overall outcome of an operation or war. There is some bias towards the player of the game in that most units in a campaign don't engage in the potentially decisive battle with the enemy (of course the whole idea of decisive battle can be debated too), but the player is much more likely than the average unit to end up in the decisive battle because it is interesting to do so (and people like to be the heroes in their games and win the day, high stakes make for a good game). On the other side however, I also wanted to represent the high likelihood that most units experience in fighting where they push until exhausted and have to conduct an operational pause or be rotated off the line. Thus, a roughly 75% chance of an Operational Pause outcome or a 25% chance of a Win/Loss outcome in a run of the game balances these two competing types of outcomes.[8] I also choose not to paint this

[1] *By Too Many Names - Operational Momentum*, Thomas R. Wallace Lt. Col, US Army, 1994, Naval Postgraduate School Thesis, https://apps.dtic.mil/sti/pdfs/ADA283467.pdf, see pg. 4-7. Further years after Wallace's publication have not alleviated the problem he identifies (see pg. 17-18) of a shared definition of Operational Momentum.

[2] Note my views on Operational Momentum are more complex that this may belie. My views in this game are mostly shaped by the Russo-Ukrainian War (especially the work of Rob Lee, Michael Kofman, and others), the 2023 Israeli invasion of the Gaza Strip, and to a lesser degree the Western Front of WW2. The papers *By Too Many Names - Operational Momentum (By Too Many Names - Operational Momentum*, Thomas R. Wallace Lt. Col, US Army, 1994, Naval Postgraduate School Thesis, https://apps.dtic.mil/sti/pdfs/ADA283467.pdf), and *The Essentials of Operational Momentum* (*The Essentials of Operational Momentum*, Major Brian A. Lovatt, 1987, School of Advanced Military Studies Monograph, https://apps.dtic.mil/sti/pdfs/ADA184741.pdf) are also the basis for my views. Numerous other wars, engagements, and other pieces of information inform this view, but those are the major ones. As for the Patton quote, this is based on his remarks to the Third Army, a set of remarks that is worth a read.

[3] Of course, the effectors of operational momentum may differ depending on unit size. Here a brigade is used, but the game would still apply if one were to have it at divisional level (though the game would be less accurate at battalion level). Conversely only certain cards in the game would make sense at Corps or Army level. This can be intuited, operational momentum in this game occurs at an operational level, so while some effects are consistent across tactical-operational-strategic, some are not, and divisional (as it slots in at the operational level of war), works better than Battalion or Corps.

[4] *By Too Many Names - Operational Momentum*, Thomas R. Wallace Lt. Col, US Army, 1994, Naval Postgraduate School Thesis, https://apps.dtic.mil/sti/pdfs/ADA283467.pdf, see pg. 5-6.

[5] *The Essentials of Operational Momentum*, Major Brian A. Lovatt, 1987, School of Advanced Military Studies Monograph, https://apps.dtic.mil/sti/pdfs/ADA184741.pdf pg.37.

[6] See for example *Making Attrition Work: A Viable Theory of Victory for Ukraine* by Franz-Stefan Gady and Michael Kofman, February–March, https://doi.org/10.1080/00396338.2024.2309068, pg. 7.

[7] *The Relationship of Battle Damage to Unit Combat Performance*, Leonard Wainstein, 1986, Institute for Defense Analyses https://apps.dtic.mil/sti/pdfs/ADA170631.pdf.

[8] Most games end up going though ~20 cards (of 43) so there is a ~50% chance to not have a decisive battle. Even when the Decisive Battle card is drawn it is possible (I estimate based on the games I have seen ~50% of the time) to thread the needle on it and not win but not lose the decisive battle and continue to an operational pause outcome.

Keep it Moving v7.0 © Evan D'Alessandro Available at evandalessandro.com

Types of Evidencing - Use Footnotes/Endnotes

Rulebook

A unit with no combat power remaining is destroyed. [77]

Endnotes/Design Notes

[77] **Why Can Units be Ground all the Way Down in this Game?**

Exact percentages for units to become ineffective are not accurate (see *Casualties as a Measure of the Loss of Combat Effectiveness of an Infantry Battalion*, Dorothy K. Clark, Operations Research Office, Technical Memorandum ORO-T-289. Chevy Chase, 1954., https://apps.dtic.mil/sti/tr/pdf/AD0059384.pdf, pg. 3, 7-8). That said, abstractions and quantifications for the purpose of building a game model must be made and are discussed here (see later tables).

This view of units remaining effective in combat even when ground down to a much lower number than classically assumed to make a unit “combat ineffective” is based on of *The Relationship of Battle Damage To Unit Combat Performance, Leonard Wainstein*, 1986, Institute for Defense Analyses as well as Casualties as a Measure of the Loss of Combat Effectiveness of an Infantry Battalion. See also *Effects of Air Interdiction Attacks on Advancing Armored and Mechanized Ground Forces*, Maj. Daniel Clevenger, March 1997, https://apps.dtic.mil/sti/citations/ADA331762, pg. 24, on the effects of a “combat ineffective” unit still being able to influence the battle.

These arguments seem to be at least somewhat borne out by descriptions of battalions in Ukraine still fighting at ~20% (40 of 200 TOE troops) or 35% strength (https://www.washingtonpost.com/world/2024/02/08/ukraine-soldiers-shortage-infantry-russia/), or a brigade at 40% infantry (https://www.pravda.com.ua/eng/articles/2024/09/17/7475408/). See also the excellent discussion of breakpoints and attrition in *Effects of Air Interdiction Attacks on Advancing Armored and Mechanized Ground Forces*, Maj. Daniel Clevenger, March 1997, https://apps.dtic.mil/sti/citations/ADA331762, pg. 16-25.

Note variance in real world outcomes here is likely to a bevy of factors, as well as the non-inclusion of breakpoints in this discussion. Functionally, there is an importance difference between a unit’s ability to attack which is exhausted before it’s ability to defend is. Clark (*Casualties as a Measure of the Loss of Combat Effectiveness of an Infantry Battalion*) defines these factors as follows…

Types of Evidencing - Short and Sweet

HEX ENCOUNTERS

This is Hex Encounters a very simple introductory[1] hex and counter game on modern brigade combat. As such these rules are short. If a rule doesn't make sense, use your military judgment to ignore, change, or come up with a new rule.

~Turns

The game is played in turns representing 1 day, with hexes of 30km. In one turn:[2]

1) Red Attacks with all their pieces.
2) Red Moves with their pieces.
3) Blue Moves with all their pieces.
4) Blue Attacks with their pieces.
5) Detected units are flipped back up.
6) Supply is traced.
   a. Remove cohesion from unsupplied units.
   b. Spend Reinforcement points.

~Units

Each unit is represented by a block. The icon shows the type, the name is the type and size (e.g. Armor Battalion). The numbers below show the unit's movement (how far it can move), combat power (how good it is at fighting), and cohesion (how much damage it can take before being removed).[3] Orange cubes represent loss of cohesion.[4]

~Detection

Units are on blocks. When a unit ends its movement next to an enemy unit, both sides reveal what their units are.[5] Units that remain adjacent at the end of the turn are not flipped back up to their undetected state.

~Moving

Units have movement points (MP) equal to their movement rating. Different terrain requires a different number of MP to move into (a unit with insufficient MP may not move into a hex requiring a larger number than is available). The number of MP is as follows:

Terrain	*Movement Points*[6]
Road	½
Open	1
Hills/Forest/Urban	2
Hills + Forest	3
River Hexside	Impassible. With collocated Engineers, 4 MP to ferry.

~Attacking

Pick a hex, any adjacent units (or artillery within 2 hexes[7]) may attack all enemy units in that hex (e.g. combat occurs in the attacked hex). Each unit may only attack once per turn, and artillery can never attack on its own.[8] If all units in a hex retreat or are destroyed, the attacking unit(s) may advance one hex into the space they formerly occupied. Players decide how to allocate cohesion losses to their units.[9] To determine combat:

1. Find the ratio of combat power between the attacker and defender and find the matching column on the *Combat Results Table* below (round to nearest ratio).
2. Apply any other column shifts from the Column Shift Table.
3. Each side rolls 1d4[10], and shifts that number of columns in their side's favor.

-Combat Results Table

Ratio[11]	**1:4**	**1:3**	**1:2**	**1:1**	**2:1**	**3:1**	**4:1**	**5:1**	**6:1**	**7:1**	**8+:1**
Result	▲⚐⚐	▲ ⚐	▲	▲	☠	☠	⛊	⛊ ⚐	⛊ ⚐	⛊ ⚐	⛊⚐⚐

▲=Atkr. loses 1 cohesion per dfendr., ⛊=Dfendr. loses 1 cohesion per atkr., ☠=Atkr. and Defndr. lose 1 cohesion per enemy unit
⚐ = Number of hexes damaged unit must retreat. A unit can choose not to retreat (or to retreat some) and loses 1 cohesion for each ⚐ it chooses to ignore.[12] If unit must retreat and cannot retreat to friendly hex it is destroyed.

-Column Shift Table

Factors	**Column Shift**	**Other Effects**
Woods or Hills	1 shift for Defender	
Light Urban	2 shifts for Defender	
Heavy Urban	4 shifts for Defender[13]	May ignore 1 ⚐
Attacking Across a Bridge	2 shifts for Defender[14]	For each ⚐ the Attacker loses 1 cohesion
Per Loss of Cohesion[15]	1 unfavorable shift per	
Attack from 2 or more non-adjacent ●'s	1 favorable shift	

~Supply

A unit is in supply if it can trace a path on the board back to a road that connects to their side which is not interdicted by an enemy unit's Zone of Control (ZOC). An infantry or armor unit exerts a ZOC 1 hex around them, all other units exert a ZOC within their own hex.[16] Friendly units negate an enemy ZOC in the hex they are in. A line of supply is also interdicted over a river if the last unit to cross the river at that bridge was an enemy unit. A unit in supply at the end of turn may receive reinforcement,[17] a unit out of supply loses 1 cohesion and halves the number of movement points it has.[18]

~Stacking Limit[19]

Certain types of terrain also have a maximum number of units that can be in that hex, called a "stacking limit".[20]

Terrain	Open	Woods or Hills	Woods+Hills / Light Urban	Heavy Urban
Stacking Limit	6 units	8 units	8 units	30 units

---

[1] This is designed to teach the following: *Classic Hex and Counters Information*: hexographic representation of space, units with Movement-Combat stats (or Movement-Offense-Defense), terrain (affecting movement and combat, and stacking limits), Combat Results Tables and column shifts, Zones of Control (in this case only for supply, but it gets the point across). *Other Information*: simple combined arms maneuver, blocks as hidden information, simple attrition of combat power, simple logistics

[2] This represents the differences between a more clunky doctrine and a more reactive doctrine. This idea is taken from TSR's *Red Storm Rising*.

[3] Cohesion includes not just the number of soldiers and vehicles in a unit, but also morale, fatigue levels, wear and tear, and other intangibles.

[4] This means that it becomes very easy to see how units are being attritted, but 1) some degree of this would be known to the enemy due to SIGINT intercepts, and 2) It's much easier than doing it in a hidden manner with another sheet of units and cubes.

[5] The basis for the size of the areas of reconnaissance (20km for a regular battalion) comes from *Warfighter 2 (Remote)* by Tim Price

[6] This is based on the TRADOC Pamphlet 350-14, September 1994, Heavy Opposing Force (OPFOR) Operation Art Handbook (https://upload.wikimedia.org/wikipedia/commons/8/8d/TRADOC_Pamphlet_350-14_-_Heavy_Opposing_Force%2C_OPFOR_Operational_Art_Handbook_%28September_1994%29.pdf), pg. 58 (3-3), Note the source assumes hard surface roads (reasonable as the roads here are from a modern country with well-developed road infrastructure), and the given numbers are for divisions not brigades, brigades would likely be faster due to their smaller size.

[7] Given the range of artillery from 20-40km (depending on the type of shell), this allows artillery to support an attack from 1 hex away (30km = 1 hex, combat is assumed to occur on a hex side). This might be a slight stretch of artillery's range, but the intention is to allow artillery to function intuitively, and for people to understand how artillery can support units.

[8] Functionally artillery cannot bring the total mass of its fires without an attack going on in the timeframe of 1 day, as enemies will be undetected or unconcentrated thus not presenting enough targets to allow for artillery to have sufficient effect (under attack or attacking the enemy concentrates and returns fire (becomes detectable), presenting juicy targets). This argument is somewhat correct but is also a simplification for the game.

[9] The idea of cohesion losses being randomly applied was also considered as the lower level combat is the more random, but for ease for new players as well as for speed (as random determination requires dice rolling and therefore time), and for the fact that a commander chooses which units to place in the lead of an attack and where they defend, the player gets to choose where cohesion losses fall on their units.

[10] The use of a d4 here leads to some potential variability (+3/-3), though the use of two dice means that this is (approximately) a normal/Gaussian distribution. While some would argue about the accuracy of such variability in combat, I would note that combat is inherently complicated, confusing, and dependent upon a menagerie of different factors that make outcomes imprecise (though not random, combat is probabilistic). Adding in such variability using the d4 system (a 62.5% chance of being inside the bounds of +1/-1, and a 87.5% chance of being inside the bounds of +2/-2), is less unrealistically "swingy" than might be initially presumed. If combat is viewed by the players to be more or less random than presented here change the dice, a d6 has also been tried but was viewed as too random for battalion level operations.

[11] This combat results table is based on the DSTL Force Ratio Risk Table. Available at https://www.professionalwargaming.co.uk/Force%20Ratio%20Table%20with%20Numbers%20v0.1.png. Note that in these rules all attacks are considered to be prepared attacks verses prepared defenses, given the size of units (brigades) and the turn length (1 day), this is close enough.

[12] This allows the representation of a unit screening the enemy while withdrawing (fall back in correspondence to the strength of enemy push) or choosing to stand and fight (in which case you trade casualties for holding your position).

[13] Ignoring the debates around the accuracy of force ratios, generally a successful attack is likely at 3:1 attacker to defender. For urban operations, a ratio of 4-10:1 seems to be sufficient (average of ~7), so 7-3 = 4, thus 4 shifts for the defender in urban terrain (of course one could argue about different types of urban terrain, but at the game's level of abstraction this is ignored for simplicities sake). For some underlying numbers and discussion on force ratios see: http://www.dupuyinstitute.org/blog/2018/04/25/u-s-army-force-ratios/ (2018). Do note the complexity (and low number of datapoints), as outlined in this further article specifically on urban terrain http://www.dupuyinstitute.org/blog/2022/08/11/the-defensive-value-of-urban-terrain/ (2022). Note however that the type of terrain covered by the dataset (from World War II) would in part not account for megacities, ubiquitous modern rebar-concrete, and earthquake-proof buildings, negating some of the applicability and leading me to disagree with the assessment that "Urban terrain does not favor the defender more so than other terrain (rolling or rugged). In fact, it appears less." I instead conclude that it is (at least in game terms) equal, though more correctly I would argue that urban terrain does not provide as strong a defensive bonus as it slows down fighting (however representing this in this game would be difficult, hence the simple column shift). My views on urban combat overall are more complicated but the 4 column shifts and ability to ignore 1 retreat is a good enough approximation of a multitude of complex dynamics.

[14] Lacking information on river crossings I will attempt to extrapolate from amphibious assaults. It is extremely difficult to determine the force ratio difference from 3:1 an amphibious assault requires on average to be successful (*Charting the Pathway to OMFTS: A Historical Assessment of Amphibious Operations From 1941 to the Present*, Carter A. Malkasian, CNA, 2002, pg. 53-59). Given data from the same source (on pgs. 55-56), it seems like past 5:1 the chance of an operational pause lessens dramatically, thus giving us a 2-column shift (3:1 -> 5:1).

[15] While units remain effective in combat even when ground down to a much lower number than classically assumed to make a unit "combat ineffective" (based on of *The Relationship of Battle Damage To Unit Combat Performance*, Leonard Wainstein, 1986, Institute for Defense Analyses.), units do become less and less effective as they take casualties and accumulate the stress and wear of battle.

[16] This represents the area footprint of a unit + the area of reconnaissance of a unit. This is a rather large ZOC to exert for a combat arms unit (e.g. 90 km in diameter), but is done so that players learn about ZOC. Theoretically it could be somewhat defensible by saying that reconnaissance is focused towards the road as the most likely area of enemy movement, and thus represents a much more focused effort than a whole area one.

[17] This represents the provision of supplies and replacements to the unit, but also intangibles like provision of hot food, mail, etc..

[18] This is a harsh penalty for not being able to be resupplied for 1 day, but it ensures that players will take supply seriously.

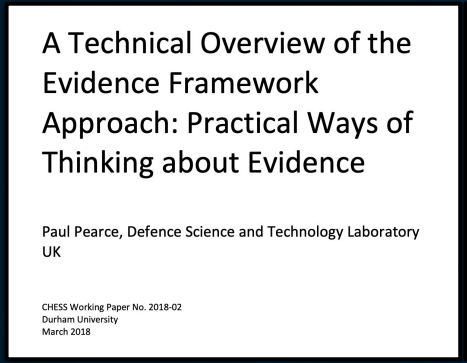
[19] Very low force densities are possible. See Ukrainian generals' comments on Brigades holding 40km lines in https://www.nationaldefensemagazine.org/articles/2022/6/15/ukraine-to-us-defense-industry-we-need-long-range-precision-weapons (2022), note that this is under the conditions of low force densities on both sides on generally open terrain. This also appears to have been the case as per this source - "the 40th Naval Infantry Brigade's two battalions were stretched over tens of km of the front" (see https://twitter.com/RALee85/status/1596128978024079360, archived at https://archive.ph/r7v8W, 2022). It is worth noting that in an urban environment the frontage of a brigade becomes 6-12 blocks, where 1 block is ~100m (*ATP 3-06 Urban Operations*, July 2022, section 4-42).

[20] Given the very low number of forces in this game, the stacking limits are likely to never come into effect, but are included as they are something worth learning.

Validation (or not really…)

There are limited validation tools available (handful of papers and works).

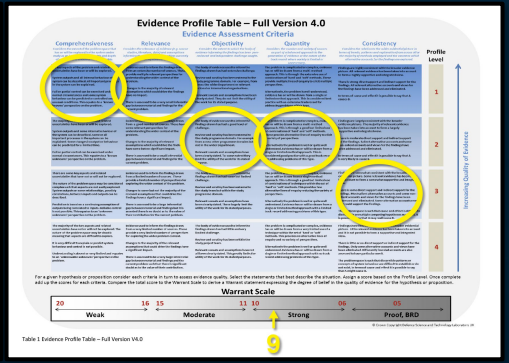
DSTL funded framework, David Burden's speculative work applying validation from PolSci.



A Technical Overview of the Evidence Framework Approach: Practical Ways of Thinking about Evidence

Paul Pearce, Defence Science and Technology Laboratory UK

CHESS Working Paper No. 2018-02
Durham University
March 2018



Evidence Profile Table – Full Version 4.0

Evidence Assessment Criteria

Comprehensiveness | Relevance | Objectivity | Quantity | Consistency

Profile Level

Warrant Scale

20 | 16 15 | 11 10 | 06 | 05

Weak | Moderate | Strong | Proof, BRD

THE NUGGET REPRINT

Wargaming and Validity

Author: David Burden, david.burden21@bathspa.ac.uk

This article originally appeared in The Nugget, Issue 349, Dec 2022, the Journal of the Wargame Developments Group. I wrote it to start to get my head around the idea presented, and comments are welcome as I try and work it up into a more polished article/paper for more formal publication.

A challenge from a client recently got me thinking (perhaps a bit late) about how wargames should be validated. Graham Longley-Brown (in Successful Professional Wargames) notes that wargamers (and system engineers) distinguish between verification (does it follow the design) and validation (is it true to reality), and it was certainly the latter definition that I was interested in. Graham notes that other communities (including the MOD and his book!) exchange the definitions. A recent talk by Kate

https://www.dur.ac.uk/media/durham-university/research-/research-centres/humanities-engaging-sci-and-soc-centre-for/CHESSWP_2018_02_Pearce.pdf

https://taunoyen.com/wiki/lib/exe/fetch.php?media=wargaming:phd:papers:wargaming_and_validity_nugget_reprint.pdf

Validation (It’s not going to happen)

It’s hard to validate wargames (especially when we haven't fought the next peer war, or have poor conceptual understandings of a topic).

No wargaming specific methods for validation - only options cross applied from other areas.

The game model is not validatable unless others can understand it (make it transparent).

What to do about not being able to Validate?

If you cannot validate what you do to mitigate risk?

- Validate parts of the rules that are validatable.
- Research more.
- Peer review (reduces risk of cognitive biases/errors).
- Make internal model transparent so it can be evaluated and changed.

Remember a given run of the game might become invalid due to player actions.

When drawing conclusions know and disclose how the game is accurate/inaccurate and how that might have affected things. A transparent model is one you can draw insight from because…

Objective Evaluation is Possible and Necessary

	Subjective Evaluation *Evaluation based on Personal Views* *"This game feels right/wrong, but I can't tell why..."*	Objective Evaluation *Evaluation based on quality of sources and quality of representation*
Accurate Model	Game has utility, BUT it's utility is unknown *(subjective evaluation kills utility)*	Game has utility, AND it's utility is known *(objective evaluation allows us to know strengths/weakness and what we can/can't take away from the game)*
Inaccurate Model	Game does not have utility, AND it's utility is unknown *(subjective evaluation kills utility)*	Game does not have utility, AND it's utility is known *(objective evaluation lets us know the game is bad and we should not take things away from it)*

Type 1 +Type 2 error will occur with subjective evaluation!

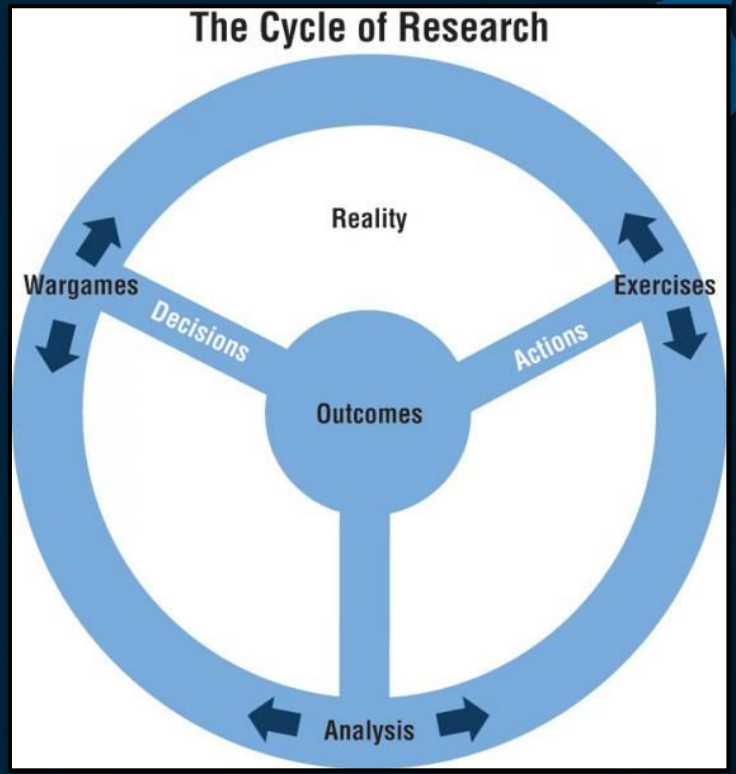
Wargaming Isn’t Enough

Wargaming is a tool in a toolbox

-It’s not magic

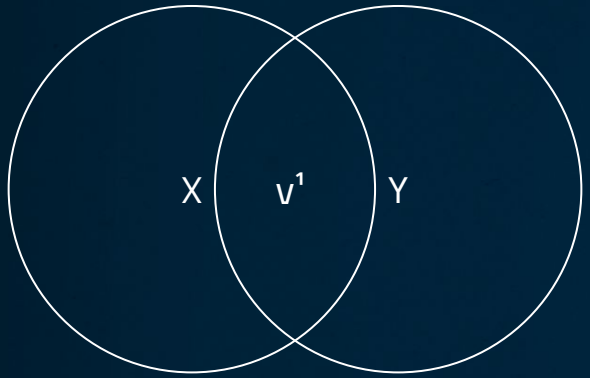
-It’s not perfect

-It cannot validate on its own



Transparent Models Reduce False Utility in Wargames

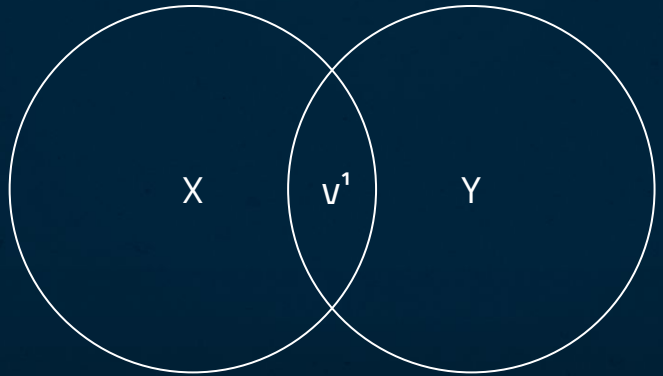
Perceived Utility
(Model Not Transparent)



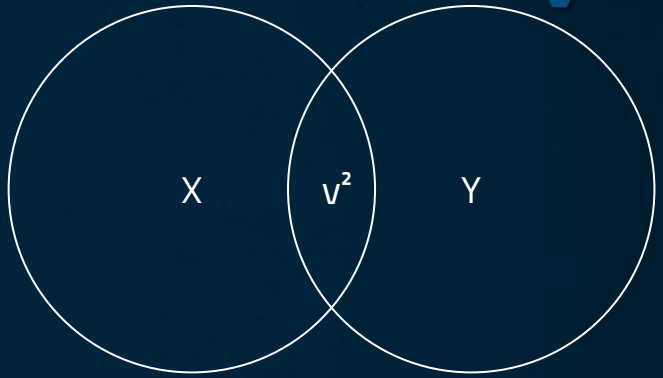
$v^1 \neq v^2$

ΔV= False Utility

Perceived Utility
(Model Transparent)



Actual Utility



$v^1 = v^2$

ΔV= Reduced False Utility

How do we know a game is good?

A: Does it meet its purpose?

Q: How can we tell if it meets it purpose?

A: By assessing the design.

Q: How to we assess the design?

A: By having a transparent model to look at.

Q: How can we tell what a good model is?

A: By being able to see the research that has been done.

Incentives for Implementation

- Improved Outcomes
  - Increase credibility of methods and outcomes
  - In educational settings, players walk away with more than face value takeaways (higher reflexivity)
    - And they have sources they can use to continue their education
- Improved Interactions
  - CYA (Show how and when others screwed your work in or outside of the design)
  - Reduce stakeholder disbelief on methods and outcomes
  - Increased support/leeway from boss/sponsor
  - Demonstrates amount of work done for boss/sponsor
- Ease of Design
  - Easier to bring additional staff/SME's onto the project (process/design is more transparent)
  - Ease of update + adaptation from and to other wargames (design and research is more portable) + ability for others to improve research
  - You you remember why you did that thing that way 4 months ago…

You Can Handle the Truth!

1. Do your research!
2. Do theory top to bottom (system to system elements to implementations)!
3. Create your citations and cite your sources!
4. Make your work public!

Can They Handle the Truth?

As you go through this conference think about what people are saying about their games, and evaluate what they say against this methodology

-Have they done enough research?

-Have they evidenced their model?

-Have they validated their model?

-Have they made all of this public for you to check their work?