Bringing Analytical Rigor to Joint Experimentation

CDR Steve “Swede” Swittel
3 Oct 06

Nature of Warfighting

- Warfare (experimentation) is a human interaction ≡ nonlinear dynamics
  - Many variables with unknown states and relationships
  - Results highly dependent on both the specific initial conditions and path taken
  - Whole ≠ sum of the parts
- Nonlinear ≠ completely unpredictable
  - Macro-level stability often exhibited

2

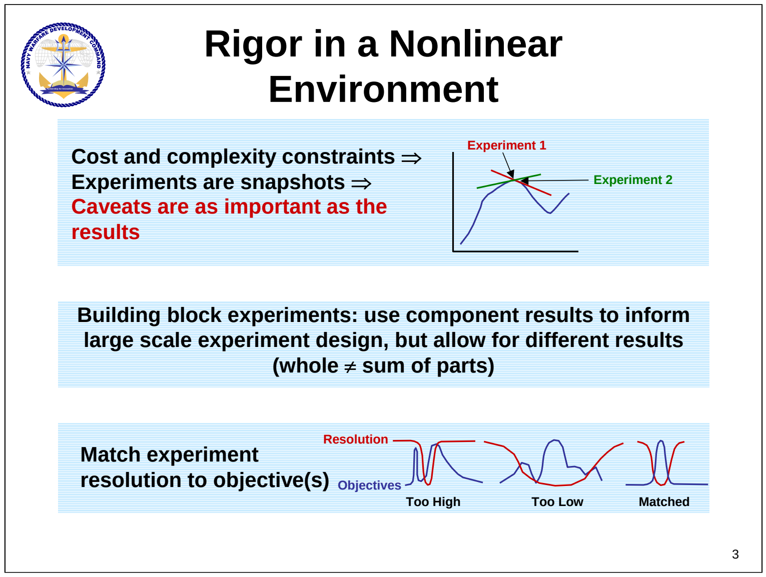
Warfighting and by extension, the associated experimentation, are fundamentally complex sets of human interactions. For the purposes of this discussion, the term "warfighting" is taken to encompass all phases of military operations including those that do not involve combat, such as shaping, deterrence or stability operations.

Although nonlinear, many of the individual component elements that make up "warfighting" can be described in linear, deterministic cause and effect terms, for example - hitting an aircraft parked in an open revetment with a 500 pound bomb results in destruction of the aircraft, but the overall dynamic is in fact nonlinear. One should consider this to be the equivalent of making a linear, straight line approximation of curve.

Two of the defining characteristics of nonlinear systems such as warfighting are:

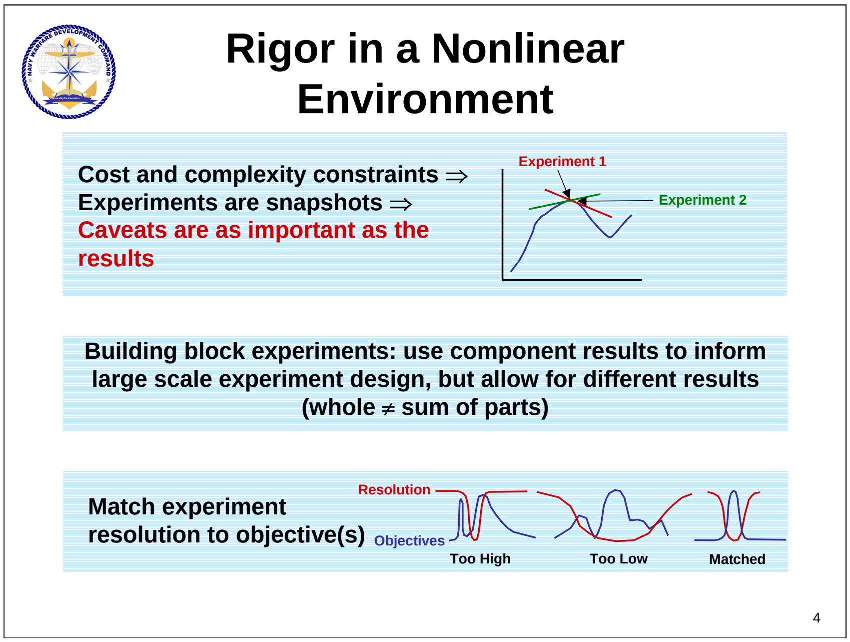
- **The results of an action are highly dependent upon both the specific initial conditions and specific path of the action (scenario). As a result, even seemingly insignificant changes in either the initial conditions or the path taken can result in vastly different results.**
- **The superposition principle doesn't necessarily apply. i.e., the overall results are not uniquely defined by the sum of the results of the component parts. How often have we heard that the sum of a capability is greater than the component parts? The whole could also be less than the sum of the parts!**

This doesn't mean that nonlinear systems like warfighting are completely unconstrained or produce purely random results. In nature, many nonlinear, chaotic systems exhibit surprising stability and predictability at macro levels while remaining indeterminate and essentially random at higher resolution. For example the Great Red Spot in Jupiter's atmosphere has remained essentially unchanged over the entire period of time that people have had telescopes strong enough to see it, yet if you look at the spot at higher magnification, the details of its structure are constantly changing.



The inherent nonlinearity of warfighting has significant impact on how one might bring analytical rigor to warfighting experimentation. Some of the issues include:

- In most cases experiment the design and associated analysis really make a localized linear approximation at a given point or points of interest on the underlying nonlinear dynamic. Hence, the results of the analysis/experiment will be valid only for the local region consistent with the linearization and are most likely not "general results" that are universally applicable. The caveats that provide the context of the experiment/analysis are as important as the results themselves and really are inseparable from the results. For example, the Navy examined processes and tools to shorten the detect to engage (DTE) sequence for fleeting and time sensitive targets in several Fleet Battle Experiments and most recently during the Joint Expeditionary Force Experiment 06. These experiments showed that a networked environment with the bandwidth and tools to proved common situational awareness among the participating nodes could significantly reduce the DTE timeline. Taken at face value and decoupled from the numerous caveats associated with experiment design and the results, one can easily draw the conclusion that the use of common tools in a high bandwidth, networked environment will produce 50+% reductions in the DTE timeline. In reality, however, most of these experiments looked at the TST portion of the fires process in isolation, without many of the competing factors that would exist in a real situation nor did they include the potentially lengthy decision delays associated with ROE, political or other issues. Thus, without the associated caveats, these result could lead to spending a lot of money to buy a system that won't consistently produce the expected results because other factors that were specifically excluded in the experiments may be the more significant drivers of many real world DTE timelines.
- The inapplicability of the superposition principle means that experiment designers and analysts need to take care in how the results from smaller, focused experiments based on parsing the "big picture" into "manageable bites" are put together to get the "big picture" results. For example, we often design experiments that focus on a single warfare area like ASW in order to control the cost and complexity of the experiments. This further highlights the importance of the caveats associated with the results and also points to the value of incorporating large scale, multidimensional experiment environments in a building block approach. The caution here is that there is no guarantee that the conclusions reached in the smaller experiments will remain valid in the large experiment environment. One of the specific analysis goals of the smaller building block experiments should be to assess which of the issues examined or identified in the experiment are likely to be unstable in a more complex, multidimensional experiment. Additionally, experiment designers and analysts need to design the larger experiment objectives to stress these areas and analyze the results without undue influence from the results seen in the smaller building block experiments with the understanding that completely different results are possible. (Assuming that the reason for the deviation was not due to a fundamental experiment design flaw.)

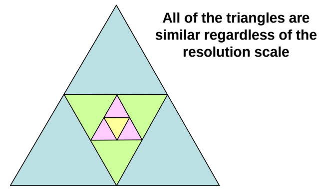


- Since many of the dynamics associated with warfighting are in fact chaotic or fractal in nature, examining the problem at the wrong resolution may give poor or inconclusive results. Too much whitewashing/white celling can make a warfighting process seem far more stable than it actually may be leading to conclusions or results that are never actually seen in the real application because of the friction created by the variability of the inputs to the process. Conversely, an experiment examining an operational level process using unfiltered high fidelity tactical level feeds to drive the problem may drive the experiment off on several different tangents, negating an ability to assess the validity of the operational level process. The concept here is similar to that of bandwidth matching an analysis window to a signal of interest. If one selects too narrow of a filter, significant portions of the signal are excluded if the filter is narrower than either the base frequency fluctuations or the width of the signal of interest. Alternatively if one selects too wide of a filter excessive background noise is brought in with the signal of interest.

Thus the analytical rigor associated warfighting experimentation is really more about identifying and accurately stating the constraints and caveats concerning the experiment design and the associated results than making sure that the results meet statistical goodness tests. Don't abandon the "standard" analytical tools used to tease relevant conclusions from the data and provide some measure of the "goodness" of those conclusions – just make the constraints and limitations of the results clear AND an integral part of the conclusions.

Rigor ⇒ Eliminate Low Value Experiment Objectives

- **Do the prep work needed to scope and frame the experiment objectives**
  - **Determine what was already learned**
  - **Include analysts in developing objectives**
    - **Do the objectives lead to "new" information?**
    - **Avoid – "because we can, we must"**
  - **Know when a more detailed look adds no greater fidelity (fractal processes that feature self-similarity)**



5

One other aspect indirectly related to bring analytic rigor to experimentation is associated with the point of diminishing returns – part of being rigorous is knowing and enforcing when not to do an experiment or when not to include certain objectives because there is really nothing new to be learned. The rigor here is in the form of doing adequate preparatory study and assessment of past experimentation and the specific experiment objectives to determine what still needs to be learned. As an example, the Services and the joint experimentation communities have experimented with chat as a collaboration tool for at least a decade. Yet, one frequently finds experimental objectives to collect data concerning the utility of chat as a medium for sharing information. Sometimes these objectives are focused on the utility of a specific tool (results highly dependent on the look and feel or user familiarity). While the specific anecdotes change, little if any new is learned about the underlying process of communicating in a collaboration environment.

In some cases, the underlying fractal nature of the issue being studied means that more data or higher resolution looks are unlikely to elicit better or more refined results. Using collaboration communications as the example again, one could argue that the information is shared and processed buy the operators is essentially the same regardless of the level or extent of the collaboration net. Thus in a case of nested collaborative processes, examining the way that information is shared in the subnets is not all that likely to provide any insight that differs from that which would be derived from examining the sharing and human processing dynamics in the parent net.

Thus a key part of rigor is in weeding out superfluous objectives or forcing examination of these objectives to an appropriate venue with properly focused metrics.