Understanding Joint Warfighting Experiments

The Logic of Warfighting Experimentation

Presented to

10th International Command and Control, Research and Technology Symposium (ICCRTS)
14 June 2005

Rick Kass
Chief, Analysis Division
Joint Futures Lab, J9
US Joint Forces Command

“Take-Away’s”

Experimentation is uniquely suited to Capability Development

Develop Capabilities to cause increased effectiveness
… and design experiments to assess causality

Logic of Experimentation is not difficult:

2, 3, 4, 5, 21

Can apply principles of science and achieve robust defensible results in Experiments

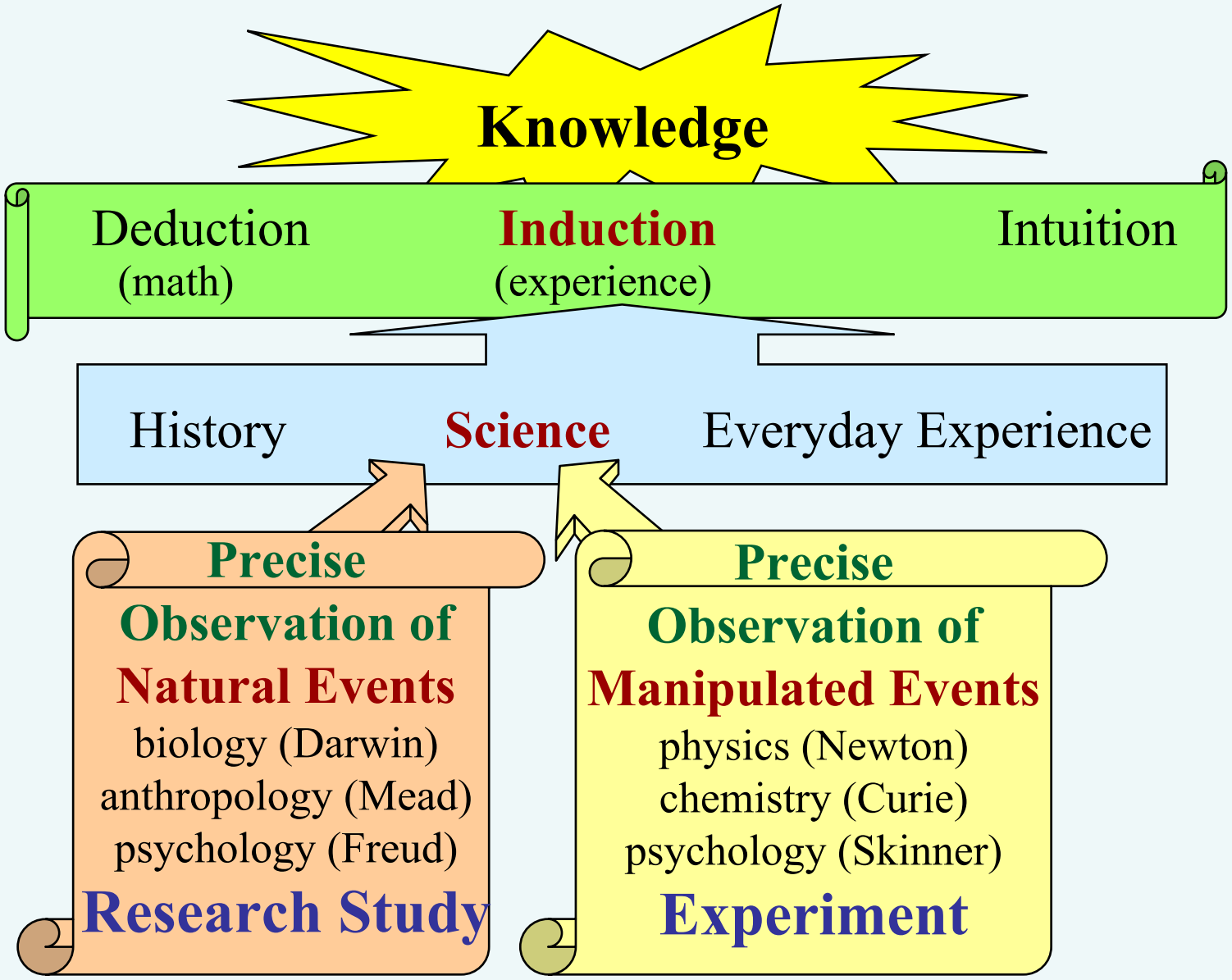
Able to empirically justify the value of new capability recommendations

Can maximize information from individual Experiments and accumulate rigor in Experiment Campaign

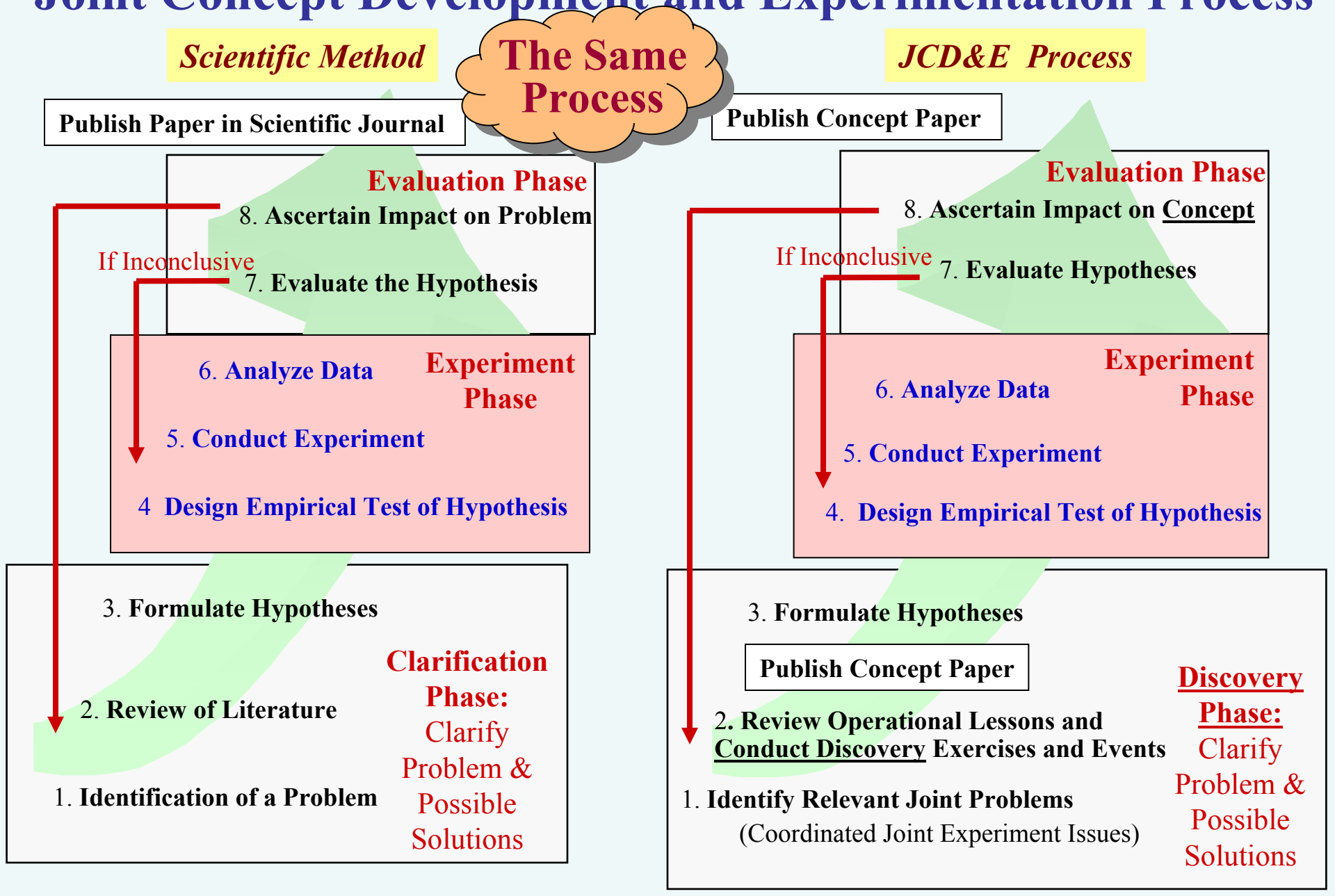
…using multiple experiment venues and continuous simulation in model-exercise-model paradigm

Scientific Method And Experimentation

Taxonomy of Sources of Knowledge



Scientific Method and the Joint Concept Development and Experimentation Process



Why Experiment?

Transformation is about--

"***changing something***" → **to increase *"Effectiveness"*** effectiveness/efficiency

To know what to change, you need to know the

"cause" of the intended output (effect).

Experimentation is the preferred technique to **determine "causes and effect"**

…in order to develop and validate—

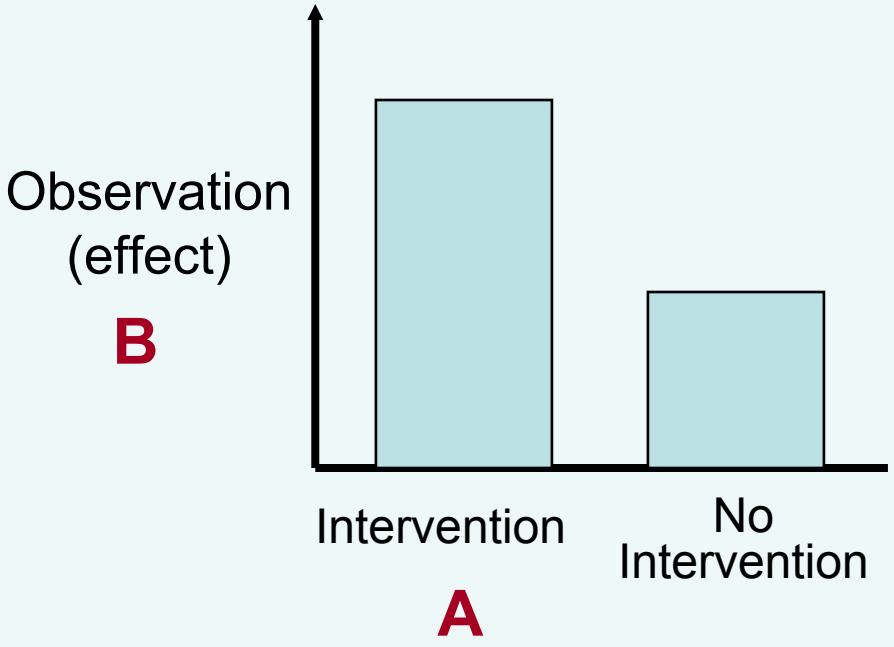
new **Joint Warfighting Capability (cause)**
that will **increase Warfighting Effectiveness (effect).**

What is an Experiment?

Simplest Experiment: (If A, Then B)

Intervene → Observe

→ Observe



Intervene → Observe

Occasionally:

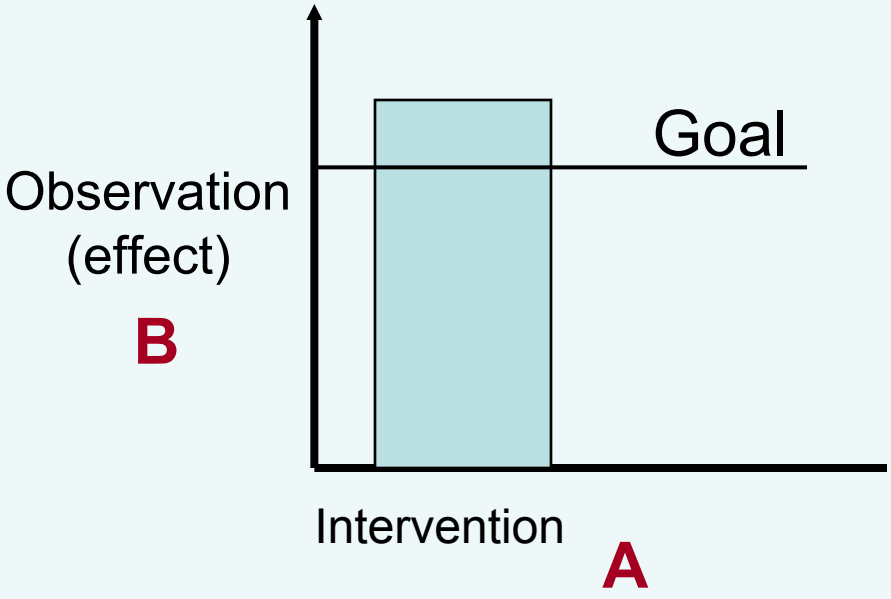
No side-by-side comparison required:

- ***Historical Goal or Baseline***

"If I sail west, I will reach the East (India)"

- ***Explicit criterion***

"Capability must deploy entire JTF within X days"



Useful Definition of Experiment

35 different definitions at "WWW. One-Look Dictionary Search"

Common Themes:

A test done in order to learn something or to discover whether something works or is true (Cambridge Advanced Learning Dictionary). An operation carried out under controlled conditions in order to discover an unknown effect or law, to test or establish a hypothesis, or to illustrate a known law (Merriam-Webster Dictionary)

Experiment –

"To explore the effects of manipulating a variable."

Shadish, Cook, & Campbell,. Experimental and Quasi-Experimental Designs for Generalized Causal Inference p. 507

Warfighting Experiment — To examine the effects of varying proposed warfighting capabilities or conditions.

Joint Warfighting Experiment

— To examine the effects of varying proposed joint warfighting capabilities or conditions.

What Will an Experiment Do for You?

A = proposed solution
B = operational problem to be overcome
C = another possible solution

- Does A affect B?
  - Is A important for solving B?
- How much A is necessary to solve B?
  - How much of B is alleviated by A?
- What is the best way to do A to solve B?
- Is C also necessary for A to work?
- Is A more important than C to solve B?

Not either-or,
…need both…
experiments
and **experience**

Analysis of historical data and the use of military experts is critical to understanding the real problem and proposing potential solutions.

But Experts & History …

- Sometimes produces contradictory implications
- May not include future environment characteristics
- **Can not quantify potential effects of new solutions**
- **Can not resolve "cause and effect" retrospectively**

Non-Laboratory Experiments Methods

William R. Shadish, Thomas D. Cook and **Donald T. Campbell**,. **Experimental and Quasi-Experimental Designs for Generalized Causal Inference** (Houghton Mifflin Co; **2002**)

Thomas D. Cook and **Donald T. Campbell. Quasi-Experimentation: Design and Analysis Issues for Field Settings** Rand McNally, **1979**)

Donald T. Campbell and Julian Stanley. **Experimental and Quasi-Experimental Designs for Research** (Rand McNally, **1963**)

Experiment rigor requirements based on 40 years of writing about non-laboratory experiment requirements.
Adapted ideas and terminology for joint warfighting experiments

Apply traditional scientific principles to Joint Experimentation in innovative ways

The **Logic** of Warfighting Experiments

2

3

4

5

21

Experiment Hypotheses

"educated guesses of what might happen"

Useful:

- Help to clarify what experiment is about
- Identify logical thread of the experiment
- Guide experiment design and data collection

Nothing magic:

2

If _____; then _____.

proposed solution(s) → **problem to be overcome**

independent variable → dependent variable

potential cause → possible effect

Sea Basing → Rapid deployment
Collaboration → Adaptive planning
Global Cell → Inter-theater coordination
Robust ISR → Deny sanctuaries

Experiment Hypothesis "avoidance?"

False Concerns about use of Hypotheses:

- **Too general**…one over the world…not helpful.
- **Not justified,** derived from theory; no war-fighting theory of war fighting.
- **Too constrictive**…detrimental to discovery.
- **Not appropriate for messy military field experiments**…only useful in laboratory experiments.
- **Don't have enough information** to formulate hypothesis
- **Demand too rigorous data and analysis** to reject/accept

"You know enough to construct hypothesis!!"

All you need…is some idea of…

…**problems you are trying to overcome**

…or missions or tasks attempting to conduct

and

…the **tools or capabilities you are proposing** to attempt to solve the problem or execute the task.

Pre-conditions for experiment

1. **Problem** identification:
   Know what problem or **task** needs to be overcome or improved.
2. Proposed **Solution** identification:
   Have a proposed solution (capability) for solving the problem…for accomplishing the task.
3. Relevant **conditions**:
   Knowledge of **conditions** impacting ability of solution to solve problems eg should work in some scenarios/situations but not others.
4. Assessment (**standards**):
   1. What **metrics** would indicate that the **solution did (or did not) solve problem.**
      (MOE/MOP—decrease time, decrease required resources, increase sustainment flow, decrease campaign time, increase number of options, etc)

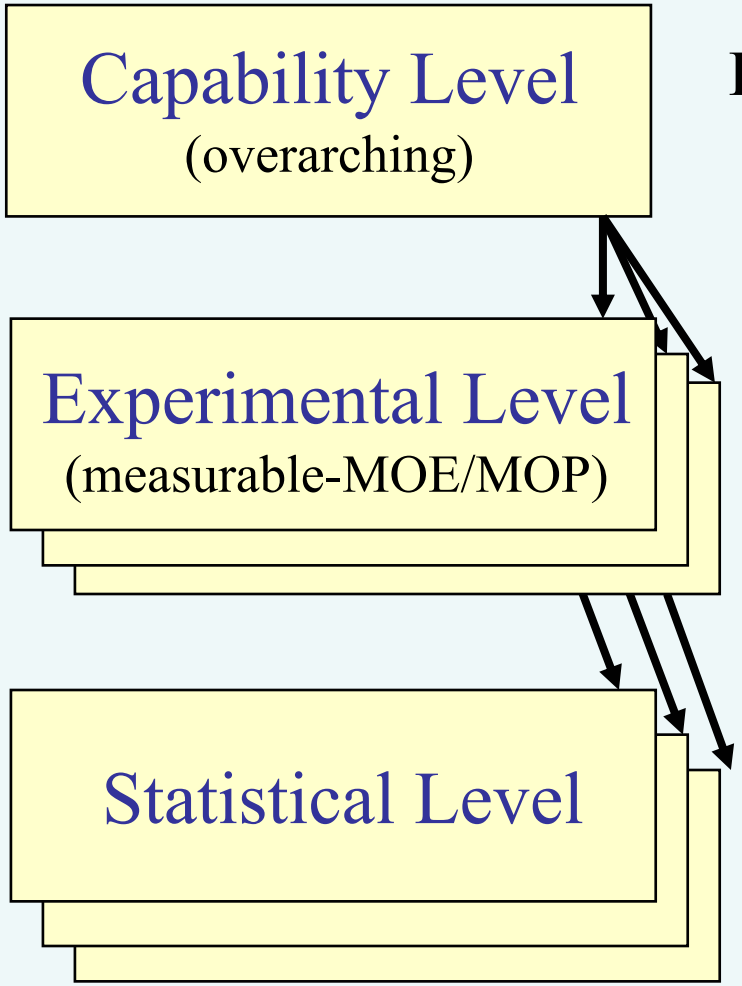
Once you have a proposed solution….hypotheses are easy!

Hypothesis:

If Proposed Solution; Then Problem to be Overcome

…trial Conditions… MOE: Metrics

Different Levels of Hypotheses



If Robust ISR is employed…;
then the threat will have no sanctuaries...

If the Advanced XX System is employed…;
then threat will be continuously tracked.

MOE/MOP

Ho: T ≥ YY
Ha: T < YY

Logic of hypothesis resolution

A **B**

If *proposed solution*: then *problem to be overcome (effect)*.

Logic of hypothesis resolution

3

1. **Did <u>A</u> occur?**
2. **Did <u>B</u> occur?**
3. **Was <u>B</u> due to <u>A</u> ?**

Internal Validity of an experiment

4 Four Requirements for Good (valid) Experiment

If New Capability (**A**) ; **Then** Effect (**B**) .

Requirement	Evidence for Validity	Threat to Validity
1 ability to **use new capability**	**A occurred**	**Asset did not work or was not used**
2 ability to **detect change** in effect	**B changed as A changed**	**Too much noise, can not detect any change**
3 ability to **isolate reason** for change	**A alone caused B**	**Alternate explanations of change available**
4 ability to **relate results** to actual operations	**Change in B due to A is expected in actual operations**	**Observed change may not be applicable**

5 *Five COMPONENTS of any EXPERIMENT*

TREATMENT (A)

Possible Cause (A)
Independent Variable
Examples
- new sensor
- new C2 process
- new JTF organization

Joint Capabilities

EFFECT (B)

Possible Effect (B)
Dependent Variable
Measure of Performance (MOP)
Examples
- targets detected or not
- time from sensor to shooter
- percent objectives met

TRIAL

EXPERIMENTAL UNIT

Smallest Unit Assigned to Treatment
Examples
- sensor operator
- sensor management cell
- Joint Task Force

ANALYSIS

Document CHANGE in B
Examples
- Outcome (B) compared to:
  - different treatments
  - different conditions

21 Threats to a Good Warfighting Experiment

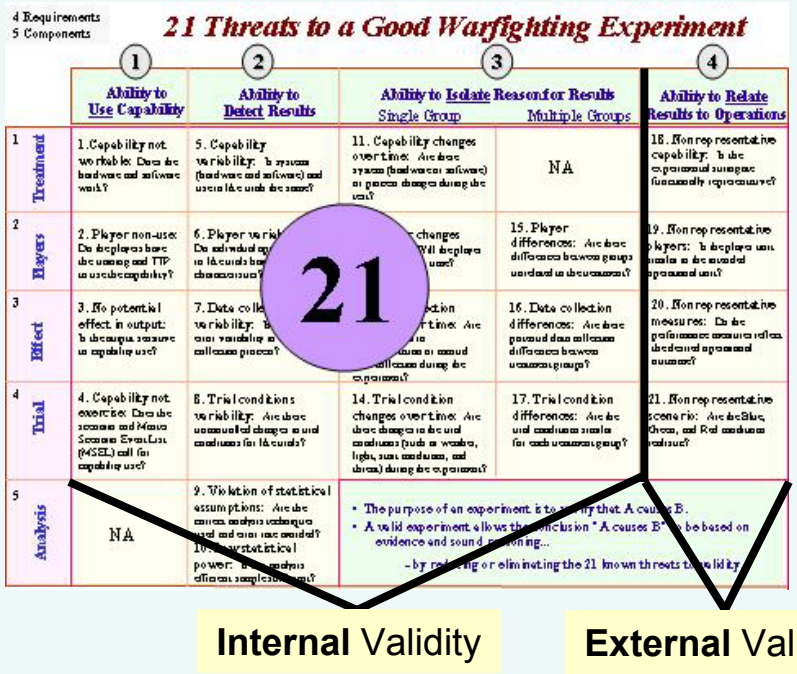
	1 Ability to Use Capability	2 Ability to Detect Results	3 Ability to Isolate Reason for Results Single Group	3 Multiple Groups	4 Ability to Relate Results to Operations
1 Treatment	**1.Capability not workable:** Does the hardware and software work?	**5. Capability variability:** Is systems (hardware and software) and use in like trials the same?	**11. Capability changes over time:** Are there system (hardware or software) or process changes during the test?	NA	**18. Nonrepresentative capability:** Is the experimental surrogate functionally representative?
2 Players	**2. Player non-use:** Do the players have the training and TTP to use the capability?	**6. Player variab** Do individual op in like trials hav characteristics?	**changes** Will the player time?	**15. Player differences:** Are there differences between groups unrelated to the treatment?	**19. Nonrepresentative players:** Is the player unit similar to the intended operational unit?
3 Effect	**3. No potential effect in output:** Is the output sensitive to capability use?	**7. Data colle** **variability:** Is error variability in collection process?	**ction** **r time:** Are s in tation or manual collection during the experiment?	**16. Data collection differences:** Are there potential data collection differences between treatment groups?	**20. Nonrepresentative measures:** Do the performance measures reflect the desired operational outcome?
4 Trial	**4. Capability not exercise:** Does the scenario and Master Scenario Event List (MSEL) call for capability use?	**8. Trial conditions variability:** Are there uncontrolled changes in trial conditions for like trials?	**14. Trial condition changes over time:** Are there changes in the trial conditions (such as weather, light, start conditions, and threat) during the experiment?	**17. Trial condition differences:** Are the trial conditions similar for each treatment group?	**21. Nonrepresentative scenario:** Are the Blue, Green, and Red conditions realistic?
5 Analysis	NA	**9. Low statistical power:** Is the analysis efficient sample sufficient? **10. Violation of statistical assumptions:** Are the correct analysis techniques used and error rate avoided?	• **The purpose of an experiment is to verify that A causes B.** • **A valid experiment allows the conclusion "A causes B" to be based on evidence and sound reasoning…** **- by reducing or eliminating the 21 known threats to validity.**		

21

Understanding 4 Experiment Requirement provides insights into Experiment Design TRADEOFFS

All Experiments are tradeoffs: -can not eliminate all threats to validity

The 100% valid Experiment does not exist



Tip balance according to decision requirements

- **Emphasize internal validity**
  - -Expect small effect
  - -Important to determine that A, and not C, caused effect
  - -Constructive and human-in-the-loop experiments
- **Emphasize external validity**
  - -Expect large effect
  - -Less important to address exactly "why"
    - -Verify effect will occur in actual operations
  - -Wargame and field experiments

A valid experiment is a balance between - -

- Internal validity - - precision and control
- External validity - - representativeness and realism

- Example: increasing repetitions for precision, also increases scenario familiarity thus decreasing realism

Valid experiment provides sufficient validity to support the pending decision

Rigorous Experiment Campaigns Require Multiple Methods to Meet the Four Requirements

Requirements	Strengths: Inherent Ability to Meet Requirements			
Employ Capability	+++	++	+	
Detect Change in Effect	+	+ + +	+ +	+
Isolate Reason for Effect		+ + +	+	+
Relate Results to Operations	+		+ +	+ + +

Capitalize on Strengths: Use combination for most Rigorous Conclusions

Analytic Wargame Experiments

Human Planners with Simulated Forces

Faster than Real-Time Constructive Simulations

Constructive Experiments

Simulated Forces

Faster than Real-Time Constructive Simulations

Human-in –the-Loop Experiments

Human Planners with Simulated Forces

Real-Time Interactive Simulations

Field Experiments

Actual Forces in Field Environment

Field Exercises & Real-Time Interactive Simulations

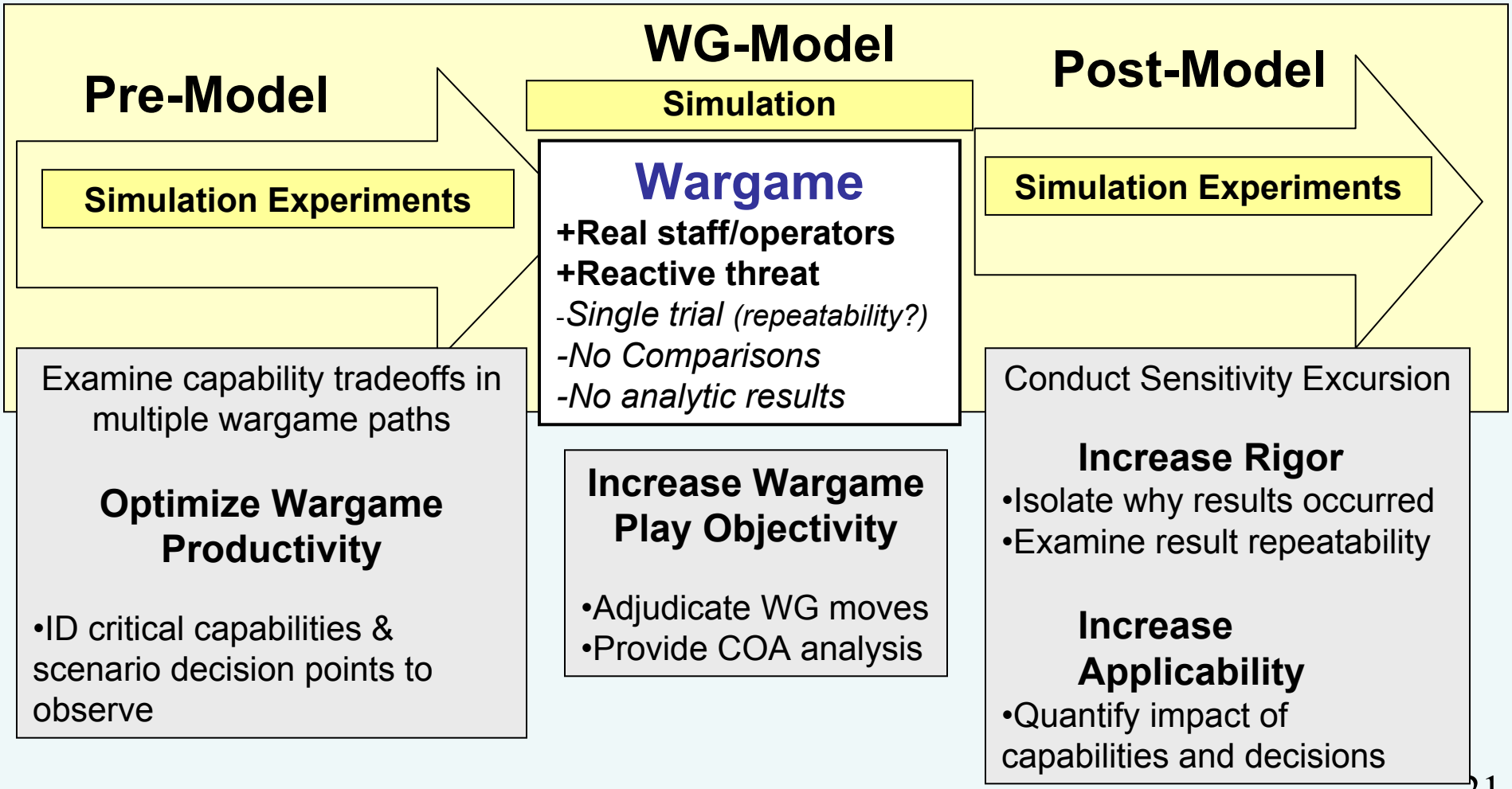
Modeling and Simulation Support Required

Combining Simulation Experiments and Wargames

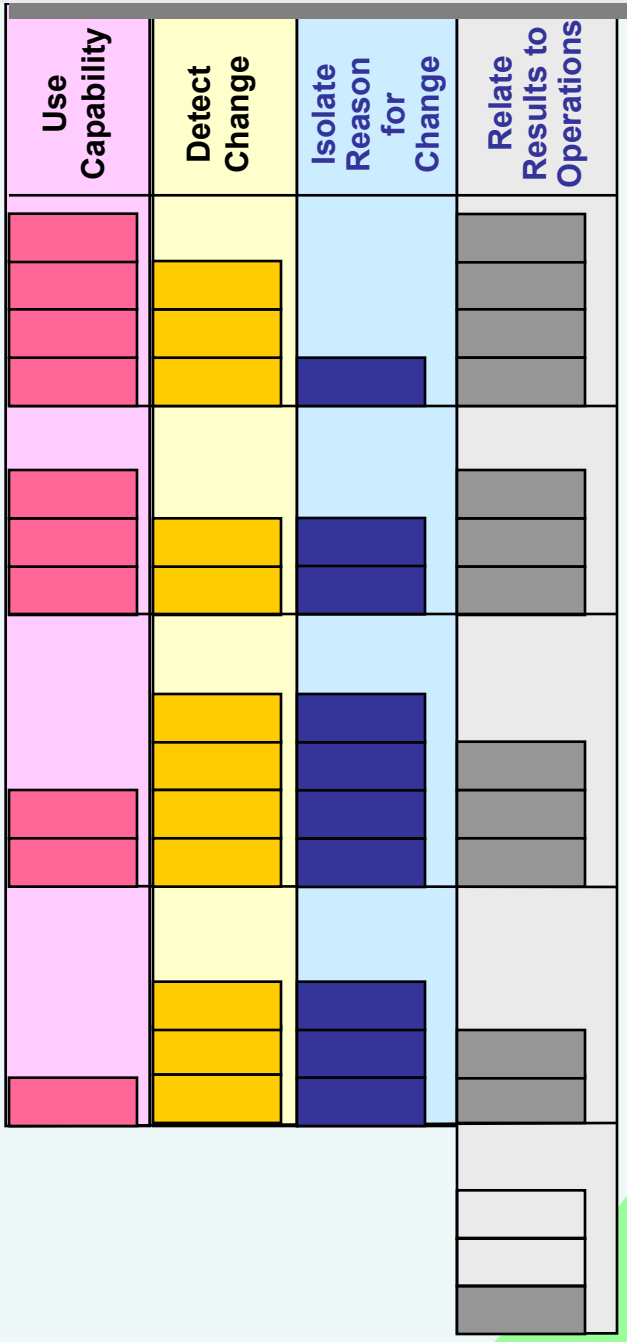
Improve Both

(combining the best of expert discussion and analytic model)

Model-Wargame-Model Paradigm



Emphasizing Exp Requirements During Concept and Prototype Campaign



Capability Implementation in Joint Force

Prototype Validation Experiments
- Demonstrate applicability to Combatant Commander's mission
- Examine predicted effectiveness in joint operational force
- Embed experiments within JTF exercises or training events

Joint Operational Capability

Prototype Refinement Experiments
- Investigate incorporation of latest HW & SW improvements
- Examine interoperability with existing fielded systems, and develop detailed tactics, techniques, and procedures (TTP)

Good Prototype

Concept Assessment Experiments
- Examine robustness across different scenarios and threat conditions
- Compare to other alternatives or baselines to quantify gains in effectiveness

Robust Concept

Concept Refinement Experiments
- Investigate optimal integration of piece-parts into most effective comprehensive solution
- Examine tradeoffs and synergistic effects between alternative combinations

Good Approach

Concept Discovery Events
- Describe future operational problem and propose solutions in coherent framework
- Operational lessons learned, military history, industry and academia workshops, conferences, & wargames

New Ideas

Sorting Through Terminology

A = New Sensor
B = Detections

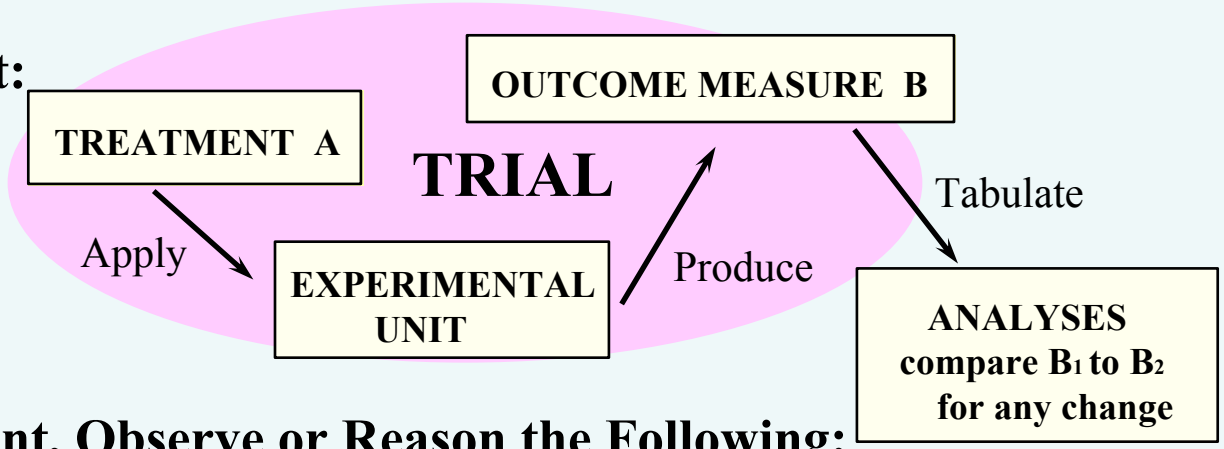
Event	Goal Stimulating Event	Purpose of Event
Train	**Practice on A** to get B.	Operation to assist entity in acquiring ability to do A.
Demonstrate	**Show how A works** to produce B.	Operation to show/explain how A works.
Test	**Determine if A works** (produces B). •How effective is A? •Can operator/unit do A?	**Operation to confirm the quality of A.**
Experiment	**Determine if A solves B.** •Is A related to B? •How much does A affect B?	**Operation to discover a causal relationship between B and something else (A).**

Experiment Logic: "2, 3, 4, 5, 21"
to support Capability Development

Formulate Hypothesis: "A will affect B" (expectation)

Conduct Experiment:

Executing 5 Components



As a Result of Experiment, Observe or Reason the Following:

Meeting the 4 Requirements (eliminating 21 threats)

- A was employed
- B changed as A changed
- A alone probably caused change in B
- Change in B occurred in typical unit and realistic scenario

Conclusion Based on Evidence and Sound Reasoning:

"A will cause B in actual operations"

Capability Development… to develop the right Capabilities (cause) to increase Joint Warfighting Effectiveness (effect)

Other Sources of Information

Richard A. Kass. The Logic of Warfighting Experiments. (in draft for comments)

Guide for Understanding and Implementing Defense Experimentation. The Technical Cooperation Program (TTCP), to be published in August 2005

David S. Alberts and Richard E. Hays. Campaigns of Experiments: Pathways to Innovation and Transformation. DoD Command and Control Research Program (CCRP), March 2005

David S. Alberts. Code of Best Practice for Experimentation. DoD Command and Control Research Program (CCRP), July 2002

NATO Code of Best Practices for C2 Assessment. DoD Command and Control Research Program (CCRP), 1998.

William R. Shadish, Thomas D. Cook, and Donald T. Campbell. Experimental and Quasi-Experimental Designs for Generalized Causal Inference (Houghton Mifflin Co; 2002)

“Take-Away’s”

Experimentation is uniquely suited to Capability Development

Develop Capabilities to cause increased effectiveness
… and design experiments to assess causality

Logic of Experimentation is not difficult:

2, 3, 4, 5, 21

Can apply principles of science and achieve robust defensible results in Experiments

Able to empirically justify the value of new capability recommendations

Can maximize information from individual Experiments and accumulate rigor in Experiment Campaign

…using multiple experiment venues and continuous simulation in model-exercise-model paradigm

Discussion