Paper ID: 048

Meta-Analysis of Multiple Simulation-Based Experiments

François Bernier, Defence R&D Canada (DRDC), Canada

David S. Alberts, Institute for Defense Analyses, U.S.A

Paul Pearce, Defence Science and Technology Laboratory (DSTL), U.K.

18th ICCRTS

C2 in Underdeveloped, Degraded and Denied Operational Environments

June 19-21, 2013 - IDA – Alexandria, Virginia, U.S.A.

Outline

- Introduction
- Meta-analysis
- Selecting simulation models and developing hypotheses
- Defining common independent and dependent variables
- Modeling effects
- SAS-085 example
- Summary

Introduction

- Simulations enable us to conduct more cost effective, less destructive, better controlled and more repeatable experiments
- Simulation-based experiments are commonplace but combining them into a meta-analysis is less frequent
- The Code of Best Practice: Campaigns of Experimentation (Alberts & Hayes, 2005) and related literature do not specifically discuss how the results of a series of experiments can be integrated into a set of findings and reflected in modifications to a conceptual model
- Other research fields (mainly human sciences) provide guidance in this regard

Meta-Analysis

- Meta-analysis is a method that combines the results of multiple experiments for identifying patterns, similarities and disagreement among the results
- The value of a meta-analysis exceeds the sum of values of each experiment taken individually
- Most meta-analyses are retrospective (past experiments) but some are **prospective** (designed before the results are known)
- Meta-analyses can be based on aggregated data (AD) or **individual participant data (IPD)**
- An existing simulation model can be (re)used at a low lost, thereby facilitating the conduct of a IPD prospective meta-analysis that require re-executing the model for investigating alternative hypotheses or getting better/more detailed data

Benefits of a Meta-Analysis

- **Generalization**: results applicable to the study space and in between contexts not explicitly tested by experiments
- **Cross-Platform Results**: control for heterogeneity among experiments
- **Increased Statistical Power**: more chance to detect an effect
- **Reduced Individual and Local Biases**: experimental errors and biases are expected to cancel each others, improving the quality of results
- **Promoted Synergies, Interactions, and Discussions among Researchers**: the approach is more likely to create fruitful interactions, foster highly critical thinking, help challenge assumptions, and support the generation of insights

Meta-Analysis Process

Conducting a Campaign of Experimentation (CoE) with a meta-analysis involves a few changes from a single experiment. Changes are related to:

- Selecting simulation models and developing hypotheses
- Defining common independent and dependent variables
- Modeling effects

Selecting Simulation Models and Developing Hypotheses

- Waterfall approach
  - 1: Establishing **the objectives** of the meta-analysis and identifying the **specific hypotheses** that will be explored
  - 2: Selecting among existing simulation models whose validity has been established
- Iterative approach
  - 1: Establishing the **general objectives** and **candidate hypotheses**
  - 2: Selecting among existing simulation models whose validity has been established
  - 3: Objectives and hypotheses are revisited and a further refinement is undertaken, including the addition of more hypotheses, based on the improved understanding as to the capabilities of the available simulation models

Defining Common Independent and Dependent Variables

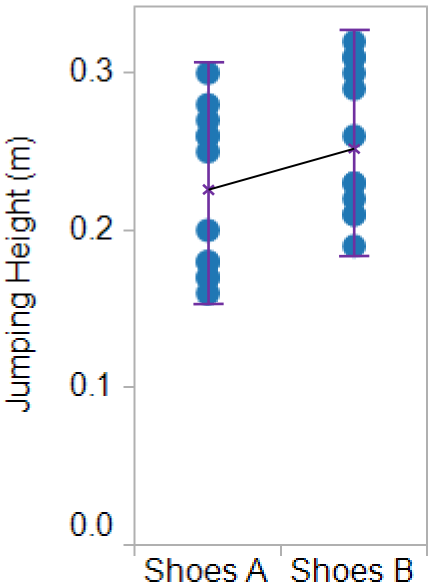
- A two step process
  - Deciding which dependent variables are needed to test the hypotheses and which independent variables are appropriate for determining their effect on the dependent variables
  - Determine how each experimental platform will capture these variables
- Similarity is often required in measures across the experimental platforms, but this is not always feasible or sometimes even desirable
- Two ways to facilitate the task
  - Relying on theories and definitions (e.g., Situational Awareness from scientific literature, C2 Approach from the NEC C2 Maturity Model)
  - Considering if variability is preferable to uniformity (e.g., Endeavor Spaces)

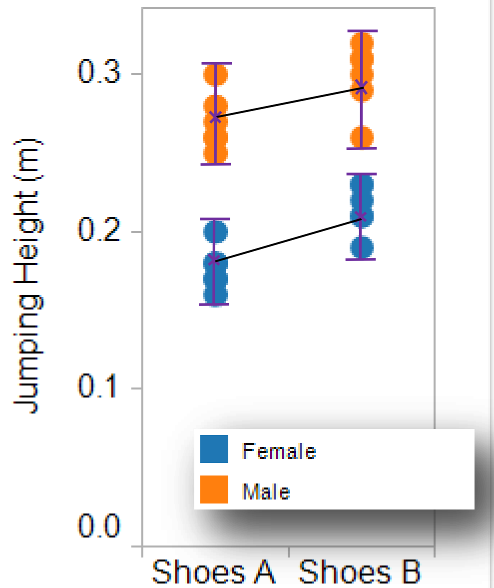
Modeling Effects

- A statistical model is required when at least one variable is probabilistic
- A statistical model establishes relationships between and among the variables of interest, the validity of which is important for the hypotheses under test.
- The Linear Mixed Model plays an important role in the analysis
  - The treatment(s) is(are) usually (a) fixed effect(s)
  - The variable experiment/simulation model is usually a random effect
- A fixed effect limits the findings to the values tested while a random effect assume that the levels tested are a sample of the whole population
- The experiment is a block which is a group of similar experimental units
  - The model captures the variance between and within blocks → better estimate of the impact of the treatments.

Modeling Effects

Are shoes B better for jumping than A?

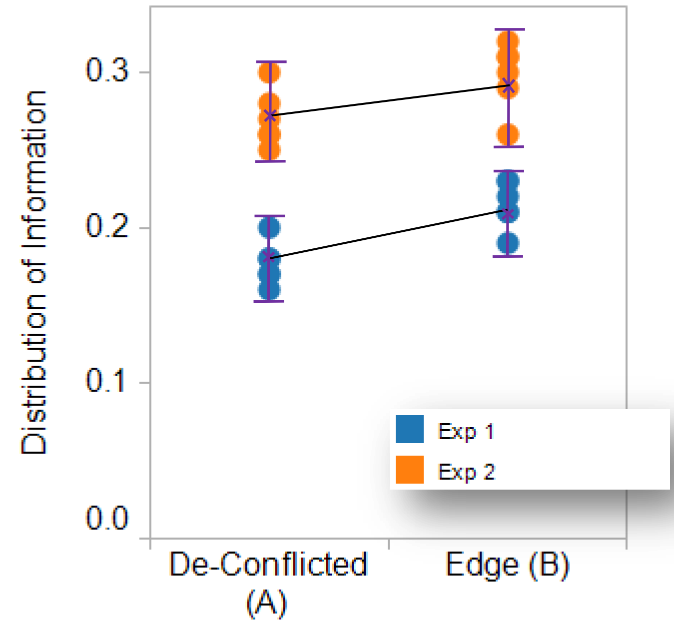




Hard so say visually
A stat test says no

Visually: more likely
A stat test says yes

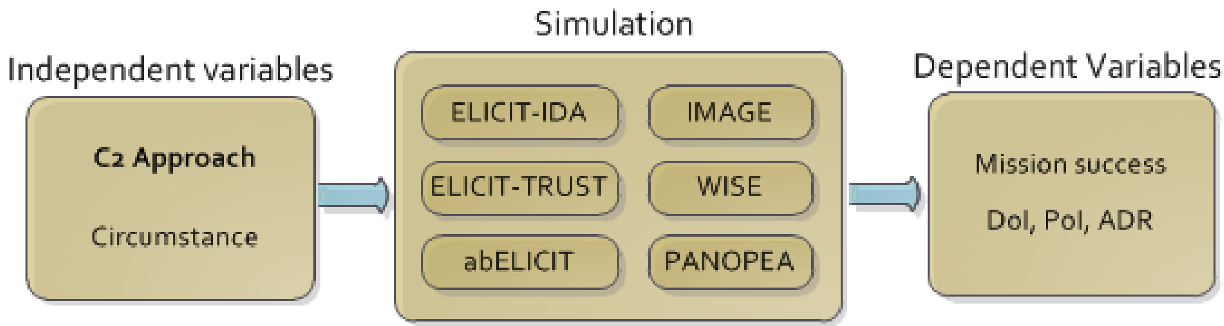
Is the C2 Approach B better than A?



An Example: SAS-085 C2 Agility and Requisite Maturity

- The SAS-085 NATO Research Task Group on Command and Control (C2) Agility and Requisite Maturity was created with the objective of improving the understanding of the importance of C2 agility for NATO and its member nations
- Several experiments were designed, conducted, and analysed separately for studying C2 Agility-related concepts
- SAS-085 developed a Campaign of Experimentation (CoE) aiming at providing a more complete, robust, and generalizable set of findings
- Five NATO member nations, namely USA, Portugal, Canada, United-Kingdom, and Italy jointly conceived a CoE and conducted a meta-analysis using multiple experimental platforms

SAS-085: Defining Common Independent and Dependent Variables



- The CoE included six experiments (each with an experimental platform)
- *C2 Approach* is the treatment (fixed effect)
- *Experiment* is a blocking variable (random effect)
- *Circumstance* is a random variable specific to each experiment

SAS-085: Defining Common Independent and dependent Variables

- Verification were made on the similarity of the C2 Approaches implemented across the experiments
- Not all experiments implements all of the C2 Approaches

	ELICIT-IDA (USA)	ELICIT-TRUST (USA)	abELICIT (Portugal)	IMAGE (Canada)	WISE (UK)	PANOPEA (Italy)
Conflicted		X		X		
De-Conflicted	X	X		X	X	X
Coordinated	X	X	X	X		
Collaborative	X	X	X	X	X	X
Edge	X	X	X			X

SAS-085: Endeavour Space and Circumstances

- The Endeavor Space was populated by circumstances/mission challenges
- Purpose:
  - Calculating an agility score
  - Reproducing the natural variability found in the real world and then improve external validity of the meta-analysis
  - Reducing the probability of selecting only circumstances that would be systematically detrimental or beneficial to some C2 Approaches
- A total of 231 different circumstances were created for the Campaign of Experimentation, far more than any previous single experiment

SAS-085: Endeavour Space and Circumstances

	ELICIT-IDA	ELICIT-TRUST	abELICIT	IMAGE	WISE	PANOPEA
Self	Network damage (3)	Message/Drop rates (3)	Infostructure degradation (2)	Latency (3)	Bandwidth efficiency (2)	
		Trust (3)	Agent performance (3)	Missing org (2)		Ship decision-making capability (2)
		Selfishness (3)	Organisation disruption (2)			Intelligence DM capability (2)
Environment	Challenge (4)		Key info. available (3)	Number of rebels (3)	Comm. link quality (2)	Number of pirates (2)
	Noise in information (3)			Crisis severity (3)		Weather condition (2)
	Cognitive complexity (3)					Misleading information (2)
#CiC	108	27	6	54	4	32

SAS-085: Dependent Variables

SAS 085
C2 AGILITY AND REQUISITE MATURITY

- Some hypotheses were related to measuring the location in the C2 Approach Space (ADR, PoI, DoI)
- Because of the large number of possible measures, it was decided that having diversified measures would capture more perspectives of the characteristics of these dimensions

Experiment	ADR	PoI	DoI
ELICIT-IDA	Amount of individual with decision rights divided by total number of individuals.	Scaled square root of number of information related transactions (post, pulls, shares).	Average percent of factoids received by each individual.
ELICIT-TRUST	Amount of individual with decision rights divided by total number of individuals.	Average number of links used.	Average percent of factoids received by each individual.
abELICIT	Amount of individual with decision rights divided by total number of individuals.	Average network reach of each individual.	Average information accessed by each individual.
IMAGE	Number of decisions allocated to the collective divided by the total number of possible decisions.	Sum of all co-conducted activities between organizations divided by the sun of all conducted activities.	Normalised difference between all variables values known by all individuals and the ground truth.
WISE	1-Betweeness Centrality	Mean of the (normalised value of Sociometric status) + (1-Bavelas-Leavitt centrality) + Inverse path length + Clustering score / 4	Mean HQ SA scores + (1-Eigenvector Centrality)).
PANOPEA	All the information taken directly by frigates and helos.	Total number of communications among actors divided by number of alerts from intelligence	Average number successful received alerts against the total number of sent alerts.

SAS-085: Results – Agility Score

- Agility Score is measured by the proportion of the endeavor space in which a collective is successful
- An agility score was calculated for each C2 Approach and experiment
- Since some values are missing, the average value was not calculated as the arithmetic means but as the least squares means

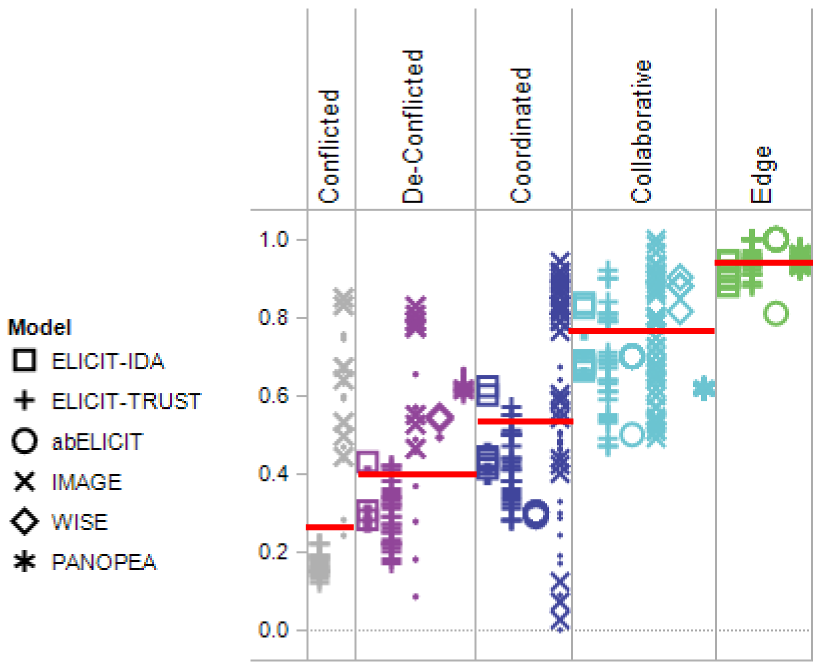
C2 Approach	ELICIT-IDA	ELICIT-TRUST	abELICIT	IMAGE	WISE	PANOPEA	LS-Mean
Conflicted		0.04		0.39			0.09 (0.10)
De-Conflicted	0.06	0.06		0.50	0.21	0.13	0.14 (0.09)
Coordinated	0.10	0.06	0.02	0.54			0.20 (0.09)
Collaborative	0.26	0.18	0.13	0.89	0.42	0.47	0.39 (0.09)
Edge	0.55	0.46	0.33			0.63	0.59 (0.09)

$$\neq \frac{0.04 + 0.39}{2}$$

SAS-085: Results – Allocation of Decision Rights

What is the average value of the Allocation of Decisions Rights (ADR) for each C2 Approach?

Individual values of ADR



Average (LS-Means) values of ADR

C2 Approach	ADR
Conflicted	-0.05 (0.13)
De-Conflicted	0.10 (0.12)
Coordinated	0.41 (0.12)
Collaborative	0.50 (0.12)
Edge	1.08 (0.12)

Summary

- The methodology presented may provide guidance for applying the principles of meta-analysis to the context of simulation-based experiments
- As the pool of simulation models reaches a significant size, there is growing potential for applying the methodology
- Statistical analysis and experimental design are complex fields and it is likely that better methods exist and were not introduced in this paper
- Although there are many challenges to overcome with combining multiple experiments/simulation models in a meta-analysis, the benefits should exceed the drawbacks
- Three papers (#015, #034, #066) on this experiment are presented in this conference

SAS 085

C2 AGILITY AND REQUISITE MATURITY

DRDC | RDDC

SCIENCE, TECHNOLOGY AND KNOWLEDGE FOR CANADA'S DEFENCE AND SECURITY

SCIENCE, TECHNOLOGIE ET SAVOIR POUR LA DÉFENSE ET LA SÉCURITÉ DU CANADA

IDA

[dstl]