C3X: Correlation, Causation and Controlled eXperimentation for C2

Dr. Paul J. Hiniker
Sr. Ops Research
September 14, 2004

The Problem

How does one test causal hypotheses on C2 effectiveness against empirical evidence?

Causation

- All we observe are covariations.
  (David Hume, 1740)

- The causal interpretation of a simple(or partial) correlation depends upon

  - the presence of a compatible causal hypothesis

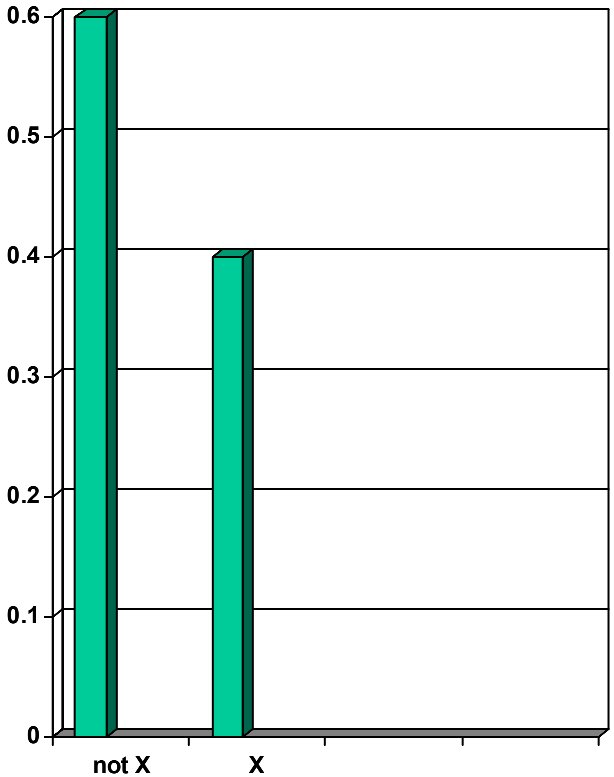
  - and the absence of a plausible rival hypothesis to explain the correlation on other grounds.

  (Herb Simon, 1957)

Causal Hypotheses & Correlation

Hypothesis: Fire engines prevent fire damage.

X---→Y $Y_X < Y_{not\ X}$ $\Phi_{XY} < 0$



	not X	X	
Y	.30	.20	.50
not Y	.20	.30	.50
	.50	.50	N=100

X = Fire engines = Fleet of 4+

Y =Percent of fires w/ damage > $500k(Expected)

40% < 60%

$$\Phi_{XY} = P_{XY} - P_X P_Y / \sqrt{(P_X Q_X P_Y Q_Y)} = -.20$$

Decomposing Correlations with Controls: Incendiary Fire Engines

X = Fire Engines Sent

Y = Fire Damage at Site

C = Size of Fire >1,000 ft

X ▸ Y

C → X, C → Y

	not X	X	
Y	.2	.3	.50
notY	.3	.2	.50
	.50	.50	100

$\Phi xy = .20$

notC

	notX	X	
Y	.2	.0	.20
notY	.6	.2	.80
	.50	.50	50

$\Phi xy\text{-}c = -.25$

C

	notX	X	
Y	.2	.6	.80
notY	.0	.2	.20
	.50	.50	50

$\Phi_{XYC} = -.25$

	notC	C	
Y	.1	.4	.50
notY	.4	.1	.50
	.50	.50	100

$\Phi_{YC} = .60$

	notC	C	
X	.1	.4	.50
notX	.4	.1	.50
	.50	.50	100

$\Phi_{XC} = .60$

Yule's (Covariance) Theorem for Dichotomous Attributes

$$\Phi_{XY} =$$

$$\Phi_{XY\text{-}C} P_{\text{-}C} \sqrt{(P_{X\text{-}C} Q_{X\text{-}C} P_{Y\text{-}C} Q_{Y\text{-}C} / P_X Q_X P_Y Q_Y)} + \Phi_{XYC} P_C \sqrt{(P_{XC} Q_{XC} P_{YC} Q_{YC} / P_X Q_X P_Y Q_Y)}$$

$$+ \Phi_{YC} \Phi_{XC}$$

- **For any two attributes, X and Y, and a third control attribute, C , the universal covariance can be decomposed into**

  a weighted average of the covariances within control subgroups , and, in addition

  a term involving the product of the covariances between Y and C and C and X.

**** N.B. In treating causality we assume, of course, that X and C are antecedent to Y.***

Controlled Experiment: Fire Engines Prevent Fire Damage

H: $X \overset{C}{\longrightarrow} Y$

	not X	X = TREATMENT	
EFFECT= Y	.32	.18	.50
not Y	.18	.32	.50
	.50	.50	N = 100 TRIALS

Random assignment of fleets(x) vs. few engines (not x)to fires:

- ***So*** $\Phi_{XC} = 0$.
- Since $\Phi_{XC} = 0$,

Experiment $\Phi_{XY} = (w' \Phi_{XY\text{-}C} + w \Phi_{XYC}) / 2 = \Phi_{XY.C}$, for $\forall$c, thus ruling out rival explanations.

- Experiment $\Phi_{XY} = -.28$
- So less fire damage is due to more fire engines on site.

Causal Modeling with Non-Experimental Data

- So to prevent spurious correlation, conduct of a controlled experiment guarantees $ø_{cx} = 0$ and ensures a valid test of a causal hypothesis.

- However, for non-experimental causal modeling,with one or more independent variables, one must verify that the residual error terms of all the variables are uncorrelated:

$$r_{Uy\,Uxi} = r_{Uxi\,Uxj} = 0, \text{ for all } X_i.$$

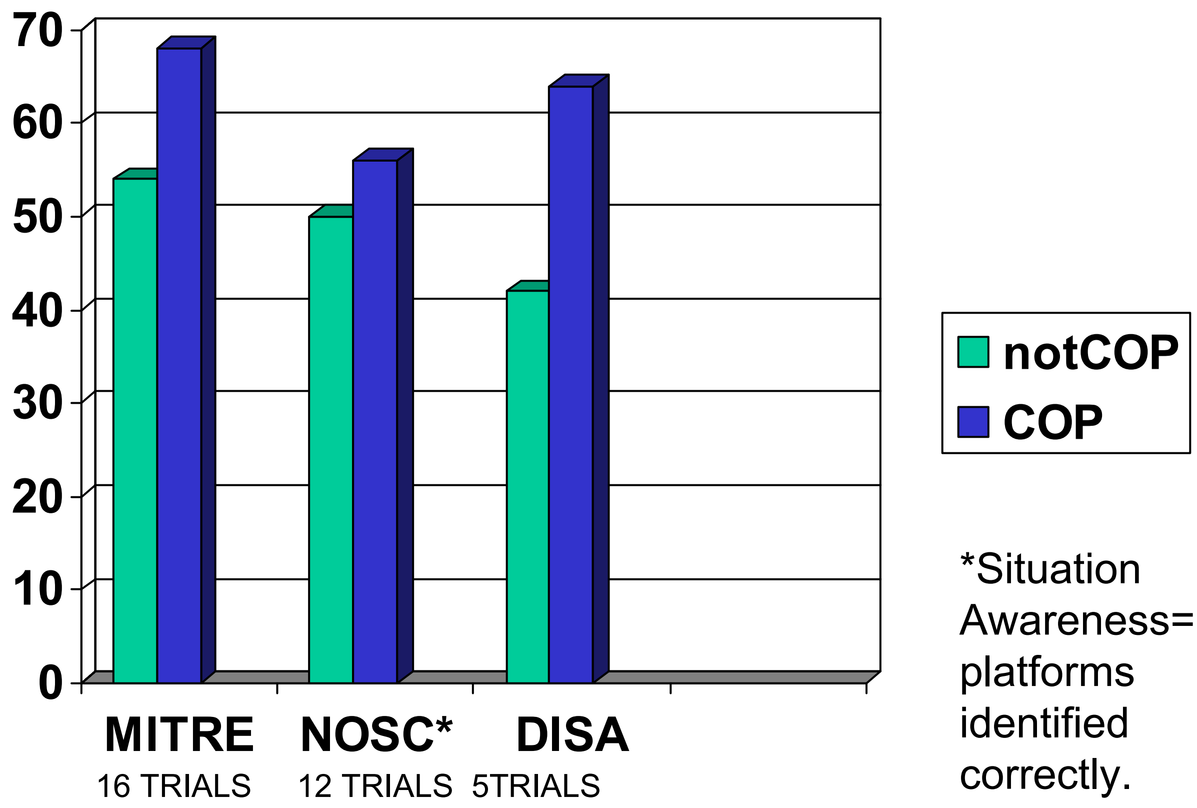
Otherwise, there could exist some extraneous variable(s), C_i, affecting both Y and X_i, hence forming part of u_y and u_{xi}, which would then be correlated; this would spuriously contribute to the correlations implied by the model.

- Thus simple correlation can neither prove nor disprove a causal hypothesis.

Controlled Experiments in C2

H: Use of a shared Common Operational Picture by a combat team(X) causes improved combat effectiveness(Y, in % platforms lost that are Red).

H: X ⟶ Y



*Situation Awareness= % platforms identified correctly.

Some Causal Hypotheses on NCW

- A basic assumption underlying most technological acquisitions for defense is the belief that the acquired capability will cause improved military effectiveness; therefore, controlled experimentation should be an integral part of the acquisition process.

- Net Centric Warfare (NCW) doctrine clearly includes such assumptions and several specific causal hypotheses such as the following:

H: Increased Shared Situation Awareness and Collaborative Planning by a distributed combat team causes increased decision loop speed and increased combat effectiveness.

- Such causal hypotheses warrant experimental testing.

Joint Interoperability, Assured Security, Best Value, Customer Success

Global Net-Centric Solutions -- The Warfighter's Edge