National Defense Industrial Association

21st Annual National Test and Evaluation Forum

Track 2: Government Approaches to Accelerate T&E and Fielding

"Experimental Design for Test and Evaluation"

Thursday, 10 March, 0930-0950

Dr Steve "Flash" Gordon, 407-482-1423, Steve.Gordon@gtri.gatech.edu

Overview

- Warfighting and Test and Evaluation Trends
  - Opportunities for Design of Experiments (DOE)
- Set Conditions for Successful DOE
- Example System Under Test
- Use DOE to Evaluate Outputs Versus Inputs

The Ultimate Customer is the Enemy, Who Must be Sure the Best Trained and Ready Forces with the Best Systems and Procedures are Poised to Respond/Fight.

Warfighting Trends for Military Operations

Precision, Speed, Improved Systems Capabilities

- Dominate: Knowledge, Speed, Precision, Lethality
- Plan and Adapt to Changing Circumstances
- Seize Fleeting Opportunities (Often in Urban Areas)
- Increasing Use of Precision Weapons
- Increasing Use of Inflight Tasking / Re-Tasking
- Growing Importance of Time-Sensitive Targeting
- More Targets Engaged Per Aircraft Sortie

Statistics in Recent Ops

Iraqi Freedom
- 68%+ of Weapons Deliveries were Precision
- Targets Received by Pilots Inflight 80% of Time
- Over 800 Time Sensitive Targets (TST)

Enduring Freedom
- 64% of Weapons Deliveries were Precision
- Targets Received by Pilots Inflight 80% of Time
- Target Approval Remained a Delicate Process
- Coordination for Tracking and Killing Time-Sensitive Target was not Smooth
- Target Approval Sometimes Slowed Campaign

Allied Force
- 35% of Weapons Deliveries were Precision

Desert Storm
- 9% of Weapons Deliveries were Precision
- On Average, Four Aircraft Needed to Service One Target

Train As We Fight!

Test and Evaluation (T&E) Trends

Test Through Spirals, Learn, Improve Systems

> "Preparing for the future will require us to think differently and develop the kinds of forces and capabilities that can adapt quickly to new challenges and to unexpected circumstances. An ability to adapt will be critical in a world where surprise and uncertainty are the defining characteristics of our new security environment."
>
> Donald Rumsfeld, SECDEF

- DEPSECDEF: Efficiency, Flexibility, Creativity, Innovation
- Goal: More Effective and Suitable Systems to Operational Combat Forces More Rapidly and at Less Cost to Taxpayer
- Evolutionary Acquisition Strategy: Spiral Development
- Extend T&E Over the Entire Life Cycle: Use Test to Learn
- Focus Testing on Missions and Mission Accomplishment

Overview

- Warfighting and Test and Evaluation Trends
  - Opportunities for Design of Experiments (DOE)
- Set Conditions for Successful DOE
- Look at an Example System Under Test
- Use DOE to Evaluate Outputs Versus Inputs

Set the Conditions for Successful DOE Evaluations

- Organize Legacy Data From DT&E, Run Charts, etc
- Understand the Process: Get Operator Feedback
- Evaluate if Process is in Control: Use Statistical Methods
- Reference Concept of Operations and Mission Description
- Diagram the Process
  - Input – Process – Output Diagram
  - Process Flow Diagram
  - Cause – Effect Diagram (Fishbone / Ishikawa Diagram)

…Continued on Next Slide

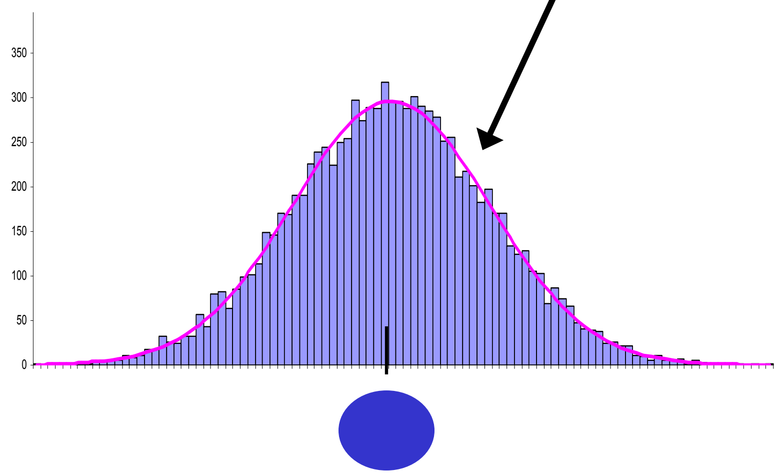
Set the Conditions for Successful DOE Evaluations

…Continued from Previous Slide

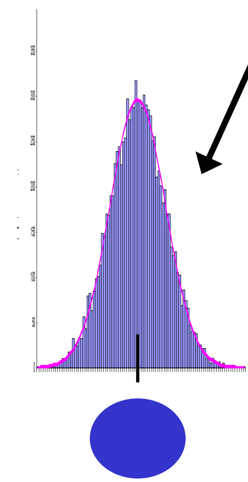
- Partition the Variables as Control, Noise, Experimental
- Write Standard Operating Procedures (SOP) for Process
- Verify Accurate Measurement System
- Screen Variables to Minimum Essential
- Determine the Inputs that Influence Location and Spread of the Output(s)
- Focus Initially: Understand Sources of Variation

Goal of DOE: Reduce Variation

Too Much Variation: Identify Source(s) and Reduce



Hit the Target
(or Very Near)
Consistently



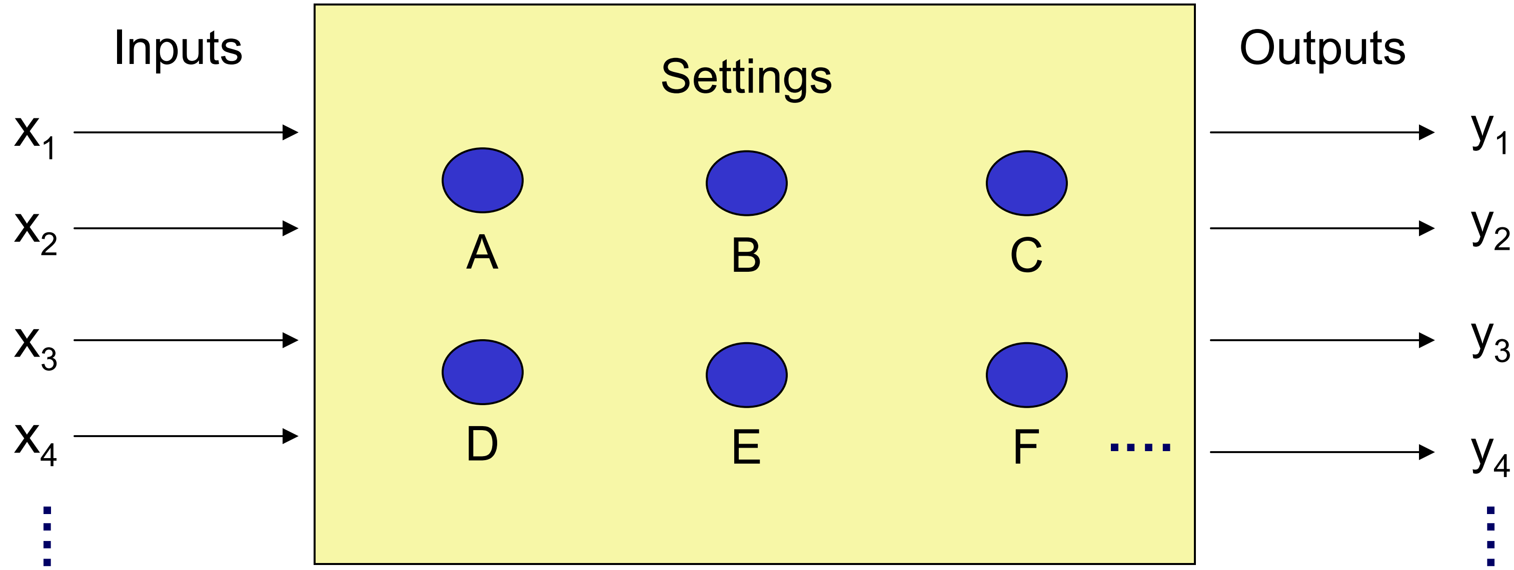
Variation in Output Variable(s) is Our Main Enemy

Assumption: Output Locations for Processes Can be Moved.

Overview

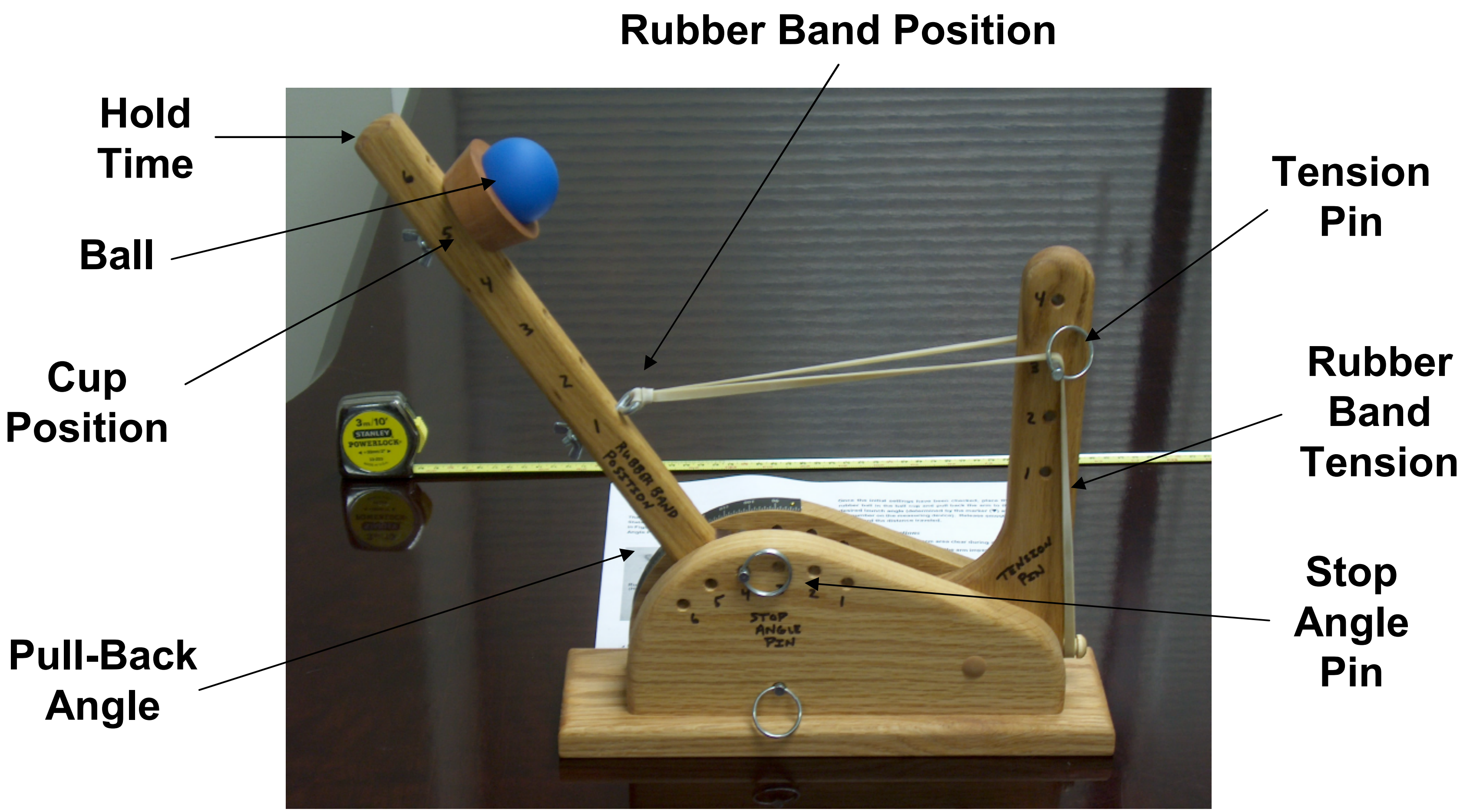
- Warfighting and Test and Evaluation Trends
  - Opportunities for Design of Experiments (DOE)
- Set Conditions for Successful DOE
- ➡ Look at an Example System Under Test
- Use DOE to Evaluate Outputs Versus Inputs

System Under Test (SUT)



- Stress the SUT to <u>At Least</u> the Limits of the Mission Profile
- Use T&E to Understand Risk and Technical Issue Areas
- Find Deficiencies Early and Reduce Life Cycle Costs
- Identify Ways to Increase Efficiency and Reduce Variability

Statapult Under Test



Statapult Courtesy of Air Academy Associates

Statapult Under Test

Rubber Bands Are Weapons

Florida boy accused of assault with rubber band; 13-year-old suspended 10 days after confrontation with teacher, *WKMG Local 6 News. (Feb 2005)*

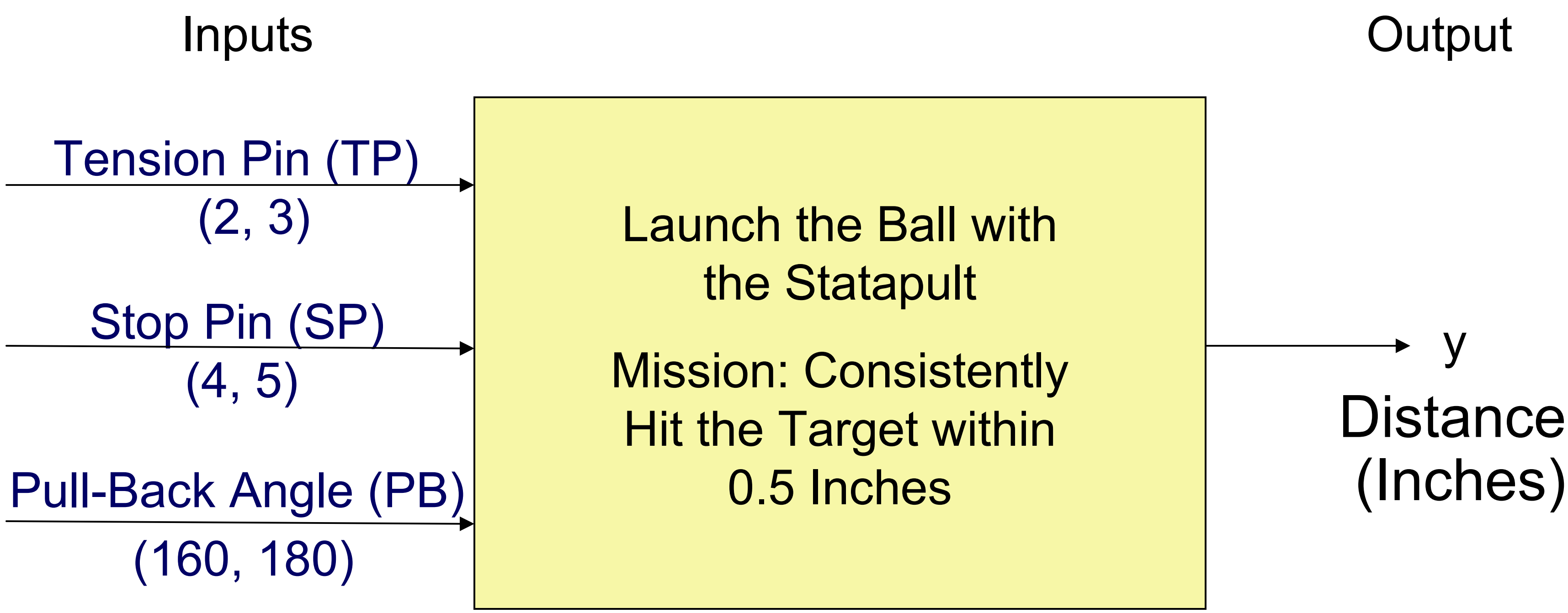
A 13-year-old student in Orange County, Florida, was suspended for 10 days and could be expelled from school over an alleged assault with a rubber band. Robert Gomez, a seventh-grader at Liberty Middle School, said he picked up a rubber band at school and slipped it on his wrist. Gomez said when his science teacher demanded the rubber band, the student said he tossed it on her desk. After the incident, Gomez received a 10-day suspension for threatening his teacher with what administrators say was a weapon. Other violations that also receive level 4 punishment include arson, assault and battery, bomb threats and explosives, according to the Code of Student Conduct. The district said a Level 4 offense includes the use of any object or instrument used to make a threat or inflict harm, including a rubber band.

Screening Design Result

Pull Back Angle…………………..Variable (PB)
Stop Pin………………………….....Variable (SP)
Tension Pin Height……………….Variable (TP)
Cup Height………………………..Constant/Controlled (5)
Rubber Band Position……………Constant/Controlled (2)
Rubber Band Temperature……...Noise
Ball Type…………………………..Constant/Controlled (Blue)
Ball Hold Time…………………….Constant/Controlled (SOP)
Ball Temperature………………….Noise
Operator…………………………...Controlled (SOP)
Room Temperature, Humidity…...Noise
Air Flow…………………………….Noise
Impact Surface……………………Constant/Controlled (SOP)…

Note: Robust Design Can Be Used to Understand and Reduce Variation Caused by Uncontrolled Noise Variables

Statapult Under Test



Question: How Can we Get the Most Information from this Evaluation with the Fewest Runs?

Statapult Under Test

Test Runs for Three Variables, Two Levels for Each

Presentation of Data in Multiple Tables is Typical in Some Disciplines: Not Design of Experiments

Table One
TP=2, SP=4

PB	Output Runs
160	
165	
175	
180	

……………….

Table Four +
TP=3, SP=5

PB	Output Runs
160	
165	
175	
180	

Classic Case of Completing Table After Table Does Not Deliver Maximum Knowledge with Minimal Runs

Statapult Under Test

Test Runs for Three Variables, Two Levels for Each

Regression Equations are an Improvement, but this Approach Below is Not a Designed Experiment:

TP	SP	PB	Average Distance
2	4	160	53.0 (5 runs)
2	4	165	58.5
2	5	170	46.5
3	5	170	62.0*
3	4	175	84.5
3	5	180	80.0
3	4	180	98.0

Regression Equation (Difficult to Determine Sensitivity)

$$\hat{Y} = 13.7391(TP) - 18.2175(SP) + 1.5783(PB) - 156.2826$$

*Confirmed at TP=3, SP=5, and PB=170 for Distance = 62.1521

Overview

- Warfighting and Test and Evaluation Trends
  - Opportunities for Design of Experiments (DOE)
- Set Conditions for Successful DOE
- Look at an Example System Under Test
- Use DOE to Evaluate Outputs Versus Inputs

Statapult Under Test

Test Runs for Three Variables, Two Levels for Each

- Code the Data to Normalize Units, Help Judge Importance
  - “Low” = “-1” or Just “-”; “High” = “+1” or Just “+”
- Create a Balanced Design that Delivers Orthogonally
- Look Explicitly for Interactions; Test for Nonlinearity

TP	SP	PB	TP*SP	TP*PB	SP*PB	TP*SP*PB
-1	-1	-1	+1	+1	+1	-1
-1	-1	+1	+1	-1	-1	+1
-1	+1	-1	-1	+1	-1	+1
-1	+1	+1	-1	-1	+1	-1
+1	-1	-1	-1	-1	+1	+1
+1	-1	+1	-1	+1	-1	-1
+1	+1	-1	+1	-1	-1	-1
+1	+1	+1	+1	+1	+1	+1

Each Column Adds to Zero; Dot Product of Any Two Columns is Zero

Statapult Under Test

Test Runs for Three Variables, Two Levels for Each

Set These & Test

TP	SP	PB	TP*SP	TP*PB	SP*PB	TP*SP*PB	Runs (5+)
-1	-1	-1	+1	+1	+1	-1	55,53.5,54,53,54
-1	-1	+1	+1	-1	-1	+1	80,80.5,80,79,79
-1	+1	-1	-1	+1	-1	+1	29,29,30,28,29
-1	+1	+1	-1	-1	+1	-1	60,60,60,60,60
+1	-1	-1	-1	-1	+1	+1	65,65.5,65,65,64.5
+1	-1	+1	-1	+1	-1	-1	100,98,100,99,100
+1	+1	-1	+1	-1	-1	-1	34,33,34,34,34
+1	+1	+1	+1	+1	+1	+1	80.5,80,80,80,80

Calculate Coefficients for All Variables in the Computation

$$\hat{Y} = 62.6 + 7*TP - 11.9*SP + 17.2*PB - .7*TP*SP + 3*TP*PB + 2.1*SP*PB + .8*TP*SP*PB$$

$$\hat{S} = .5 - .025*TP - .16*SP - .06*PB + .02*TP*SP + .14*TP*PB - .18*SP*PB - .02*TP*SP*PB$$

Model Confirms at TP=-1, SP=-1, and PB=0 for Y-hat Equals 66.5 inches.

Further Info About DOE

- Significant Underlying Theory in Statistics and Probability Distributions was Omitted in this Presentation
- Experimental Designs Vary by the Number of Factors, the Type and Number of Levels for Each Factor, and Purpose
- Good Rule-of Thumb Guidelines Can be Used to Determine the Number of Replications Needed
- Many DOE Details were Omitted:
  - Aliased Interactions
  - Randomized Replications
  - Model Confirmation
  - Nonlinearities
- DOE: Powerful Tool if Used Correctly

21st Annual National Test and Evaluation Forum

"Experimental Design for Test and Evaluation"

Questions?

Dr Steven C. Gordon
Georgia Tech Research Institute
Orlando Field Office Manager
3361 Rouse Road, Suite 210
Orlando, Fl 32817
(407) 482-1423x222
Cell (407) 963-2413
steve.gordon@gtri.gatech.edu

References

Understanding Industrial Designed Experiments,
Stephen R. Schmidt, Robert G. Launsby, Air Academy Press

Basic Statistics, Mark J. Kiemele, Stephen R. Schmidt,
Ronald J. Berdine, Air Academy Press

Design and Analysis of Experiments,
Douglas C. Montgomery, John Wiley and Sons

Lean Six Sigma, Michael L. George, McGraw-Hill

Air Academy Associates Short Courses and Course Notes
in "Basic Statistics" and "Design of Experiments"

Statapult Under Test

Even at Just 42 Test Shots, Output Appears to Approach a Normal/Gaussian Curve