Always a Step Ahead
ARDEC
ARMAMENTS

Application of Statistical Methods and Designed Experiments to Development of Technical Requirements

Eli Golden, GStat

UNPARALLELED
COMMITMENT
& SOLUTIONS

U.S. ARMY
U.S. ARMY RDECOM

U.S. ARMY ARMAMENT
RESEARCH, DEVELOPMENT
& ENGINEERING CENTER

Overview

- ARDEC is using statistically designed and analyzed surveys to enrich technical development of new munitions by ensuring documented user requirements are genuinely reflective of customer needs.

- Outline
  - Program Background
  - Voice of the Customer as Data
  - Data in Defense Acquisition
  - Survey Methodology
  - Discrete Choice Modeling
  - Results and Conclusions
  - Data Quality Challenges/Best Practices
  - Future Work
  - Acknowledgements
  - References

Program Background

- Combat Engineering and Infantry brigades use line charges to breach obstacles by uncovering, destroying, or otherwise disabling subsurface explosive threats.

- Army inventory currently contains variety of breaching munitions with different capabilities scaled towards **threat type** (anti-personnel v. anti-vehicle), **mode of deployment** (man-portable v. vehicle-borne), and **level of enemy contact** (combat v. post-combat)

- 2015 – Maneuver Center of Excellence (MCOE) drafts Capability Development Document (CDD) to address growing gaps between current capability and emerging challenges. Provides both *objective* and *threshold* requirements for various Key Performance Parameters (KPPs) and Key System Attributes (KSAs): **Portability**, **Lethality**, **Ground Disturbance**, **Stand-Off Distance**, **Time to Employ**, **Initiation System**, **Emplacement Accuracy**, **Threat Defeat Mechanism**

- Advanced Breaching and Demolition Technology (ABDT) team initiated comprehensive effort to develop, conduct, and analyze Voice of Customer surveys, over growing concerns that not all performance thresholds could be met simultaneously.

- Several rounds of surveys conducted by the team targeting experienced combat engineers and infantrymen. Surveys were designed, issued, and analyzed in an iterative fashion such that lessons learned could be implemented in successive recurrences in order to improve the quality and integrity of future observations.

An Assault Breacher Vehicle fires a mine clearing line charge during operation Rawhide, March 14. ABV's from 1st Combat Engineer Battalion launched MCLCs to breach a path into a city used by enemy insurgents to smuggle weapons, drugs and improvised explosive device making material. (U.S. Marine Corps photo by Cpl. John McCall, March 28, 2011/released)

“Voice of the Customer” as Data

- Making products better ≠ making *better products*.
- Technical organizations have traditionally minimal stake in requirements development
- Traditional customer satisfaction surveys commonly address what is wrong with current product but unable to identify consumer preferences with regard to specific attributes ← **reactive**
- Surveys designed like scientific experiments can provide a strong market research tool that enables technical teams to gain insight for actual customer preferences. ← **proactive**
- Can assess latitude between existing design margins and actual user expectations.
- Human perceptions = *latent* information (not explicitly stated, recognized, or intended). Trying to capture the representative hierarchy of user preferences simply by asking each respondent likely a futile task. A well designed survey will extract latent information indirectly and create *manifest*, actionable information.

Statistics in Defense Acquisition Lifecycle

Material Solution Analysis Phase (MSA)

- Requirements Analysis: *Conjoint Analysis DOE*
- Trade Analysis: *Comparison DOE*
- Past Systems Research: *Data Mining*
- Capabilities Assessment & Decision Analysis: *Probabilistic Risk Assessment, Stochastic Simulation, Bootstrapping*
- *Reliability Growth Planning*

Technology Maturation & Risk Reduction Phase (TM&RR)

- Pre-prototype product performance characterization: *DOE for Modeling & Simulation (DACE)*
- *Factor Screening Design of Experiments (DOE's) & Dimensional Analysis*
- *Statistical Power & Sample Size Analysis*
- *Design for Reliability*
- *Characterization DOE*

Engineering & Manufacturing Development Phase (EM&D)

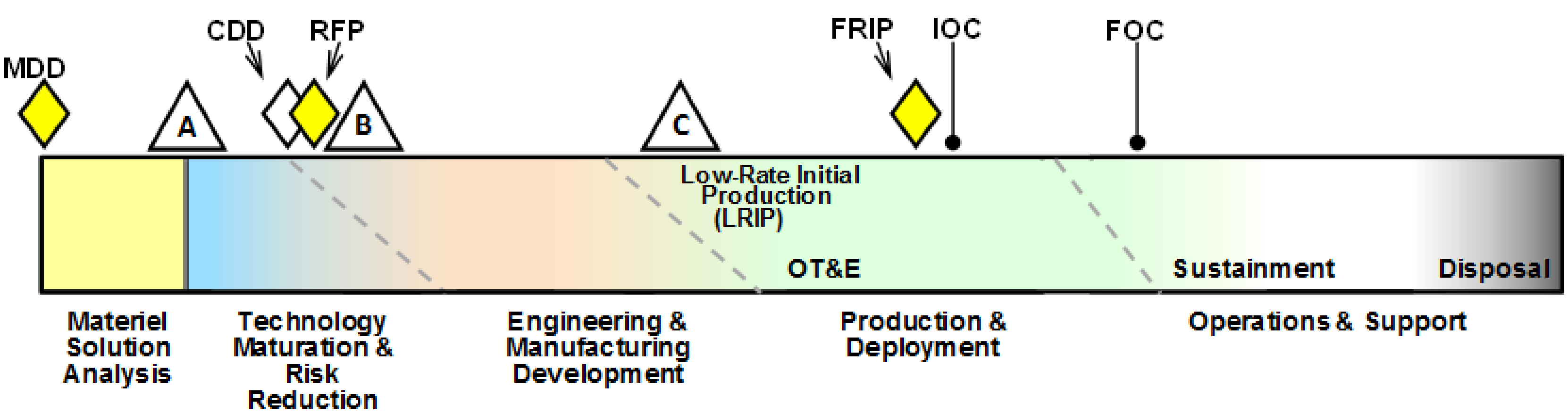
- Product and Process Improvement and Optimization: *Factorial DOE and Response Surface Methods (RSM)*
- Lean Manufacturing, DFSS, TOC
- Simulation-Based Robustness Sensitivity & Tolerance Analysis
- *Statistical Process Control (SPC) & Process Capability Analysis*
- *DOE for Evolutionary Operations*
- *TEMP OT&E 'Statistical Measures of Merit'*

Production & Deployment Phase (P&D)

- *Sampling Analysis*
- *Tolerance Refinement Analysis*
- Manufacturing: *EVOPS, SPC & SQC*
- Production Process Optimization: *Discrete-Event Simulation*
- Failure Investigations: *Multivariate Analysis, Predictive Analytics, Probabilistic Risk Assessment*

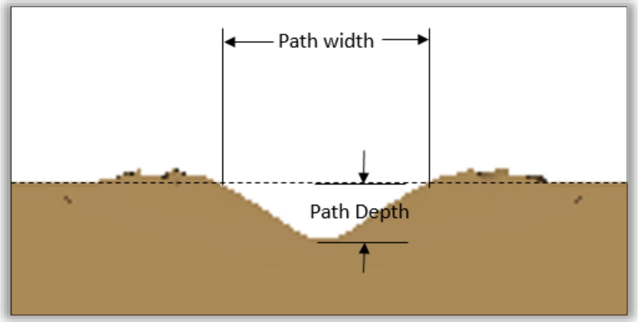
Operations & Support (O&S)

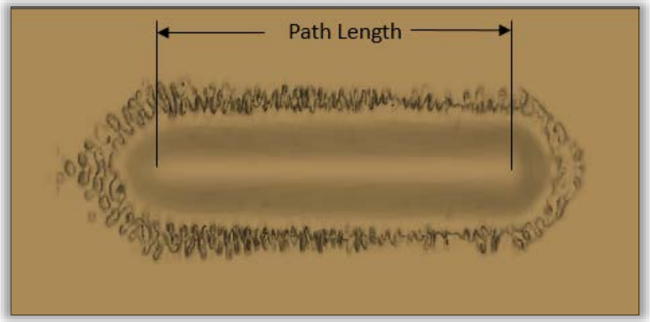
- *DOE for V&V*
- Environmental Impact Studies & Modeling
- Stockpile Prediction and Tracking: *Data Mining & Predictive Analytics*



Discrete Choice Design Methodology

1. System parameterized in terms of key attributes, including features that both add value (e.g. “stand-off distance”) and take it away (e.g. “system weight”). small list (2-4) of discrete levels for each attribute are chosen to be representative of realistic end-product traits.





Attribute 1	Attribute 2	Attribute 3	Attribute 4	Attribute 5
L1	L1	L1	L1	L1
L2	L2	L2	L2	L2
L3	L3	L3	L3	
		L4		

Methodology (cont'd.)

2. Pairs of proposed systems created from randomly assigned attribute levels presented to the subject one at a time. Subject chooses a, b, or c. Prior to finalizing survey design, sets are reviewed to ensure the comparisons are non-trivial. Instructions and diagrams included to ensure understanding of task.

	Attribute 1	Attribute 2	Attribute 3	Attribute 4	Attribute 5
a	L2	L3	L1	L3	L1
b	L3	L3	L2	L1	L1
c	*Neither option outperforms my current equipment*				

3. Survey returned. Data entered, cleaned, and processed, parsimonious model constructed that predicts perceived utility. Mathematical assumptions validated.

4. Report key findings to team. Incorporate information into body of knowledge. Use lessons learned to make targeted adjustments to survey. Issue new survey to larger set of respondents. Repeat process with new survey.

*Other data collected includes demographics, Lickert-style climate questions, free-form text entry.

Choice Model Form

- For a given respondent n, let the net utility for a given product with attribute set x_i be expressed by:

$$U_{ni}(s, x_i) = v_{ni}(s, x_i) + \varepsilon_{ni}(s, x_i)$$

- $v_{ni}(s, x_i) =$ deterministic function of both product (x_i) and demographic (s) attributes indicative of “representative” tastes

- $\varepsilon_{ni}(s, x_i) =$ random function of independent individually distributed values for each attribute set, reflecting all uncontrolled factors that influence subject n’s choice

- Probability that the subject n will choose feature set i over all other possible alternatives (J) is given by the conditional logit model:

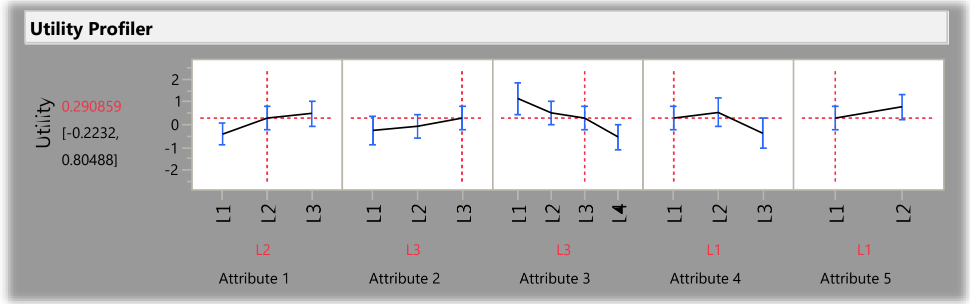
$$P_{ni} = \frac{e^{v_{ni}(s,x_i)}}{\sum_{j=1}^{J} e^{v_{nj}(s,x_j)}}$$

- Maximum Likelihood estimates for model coefficients solved numerically using Iteratively Reweighted Least Squares (IRLS) procedure with Newton-Raphson Algorithm

Results

- Predicts psychological tradeoffs that respondents make when evaluating several attributes together, which may or may not be apparent to respondent themselves

- Software output contains:
  - Utility Profiler: dynamic visualization of prediction formula with 95% confidence intervals
  - Model coefficient estimates
  - Model diagnostics and details
  - Effect tests for significance of each specified model term



Parameter Estimates

Term	Estimate	Std Error
Attribute 1[L1]	-0.53988397	0.1130172960
Attribute 1[L2]	0.16556564	0.1176142004
Attribute 2[L1]	-0.24048109	0.1522296264
Attribute 2[L2]	-0.05955953	0.1084071442
Attribute 3[L1]	0.79564001	0.1963657008
Attribute 3[L2]	0.16342993	0.1212106805
Attribute 3[L3]	-0.06607361	0.1942312141
Attribute 4[L1]	0.14110296	0.1725773519
Attribute 4[L2]	0.38669818	0.1116311905
Attribute 5[L1]	-0.24977632	0.0899203180
No Choice Indicator	-1.58887820	0.1553625830

AICc	999.33849
BIC	1046.7276
-2*LogLikelihood	976.8679
-2*Firth LogLikelihood	927.72706

Converged in Gradient

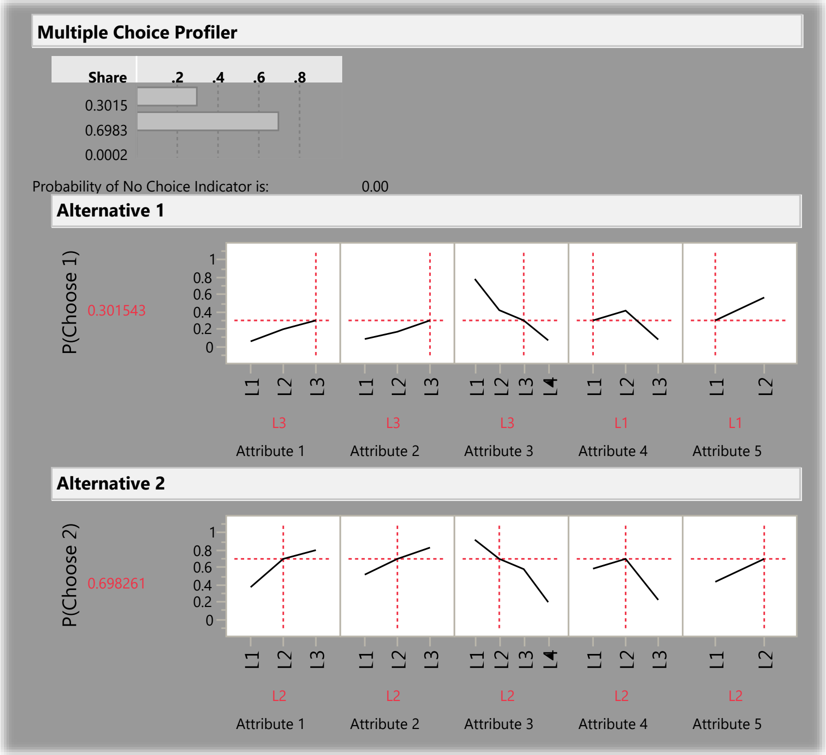
Firth Bias-Adjusted Estimates

Likelihood Ratio Tests

Source	L-R ChiSquare	DF	Prob>ChiSq
Attribute 1	24.832	2	<.0001*
Attribute 2	4.552	2	0.1027
Attribute 3	54.784	3	<.0001*
Attribute 4	31.167	2	<.0001*
Attribute 5	7.810	1	0.0052*
No Choice Indicator	144.461	1	<.0001*

Next Steps: Multiple Choice Predictions

- Utility profiler enables us to visualize estimates of perceived utility for given set attributes and demographics.
- Math model does not automatically solve problem simply by existing.
- Conditional logit model easily lends itself to computing relative probabilities for multiple choices, facilitating comparison of different hypothetical feature sets.
- Used to demonstrate that a new feature set could show increased value to the soldier over baseline system even without simultaneously meeting every performance threshold in the draft CDD.



Next Steps: Prioritizing Attributes

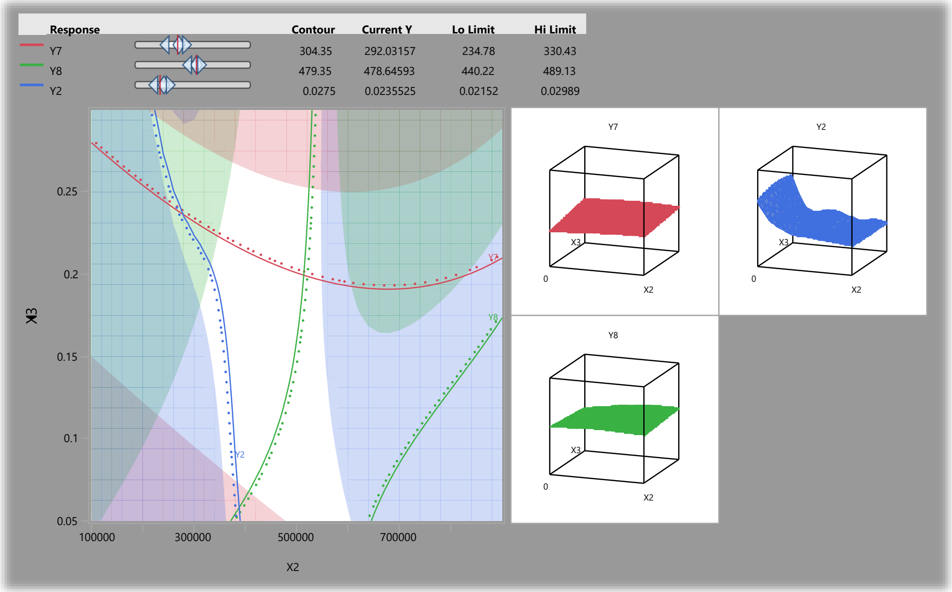
- Various techniques exist to compute attribute importance indices using model estimates

Variable Importance: Independent Uniform Inputs

Summary Report

Column	Main Effect	Total Effect	.2 .4 .6 .8
Attribute 3	0.452	0.47	
Attribute 4	0.176	0.194	
Attribute 1	0.15	0.168	
Attribute 2	0.081	0.098	
Attribute 5	0.059	0.076	

- Helps team prioritize technical objectives during product development process, enabling:
  - Technology downselect
  - Optimization of multiple competing functional responses (e.g. longer path ⇒ higher weight)
  - Efficient allocation of programmatic resources



Piloting

- No guarantee of success with only one survey.
- Lessons learned pertain as much to survey process as they do to product itself.
- Credible prediction models can be used to train the discrete choice design algorithm for the next revision.

Choice Design

Attributes

Model

DOE Model Controls

Prior Specification

☐ Ignore prior specifications. Generate the Utility Neutral design.

Prior Mean

Effect	Prior Mean
Attribute 1 1	0.000
Attribute 1 2	0.000
Attribute 2 1	0.000
Attribute 2 2	0.000
Attribute 3 1	0.000
Attribute 3 2	0.000
Attribute 3 3	0.000
Attribute 4 1	0.000
Attribute 4 2	0.000
Attribute 5	0.000

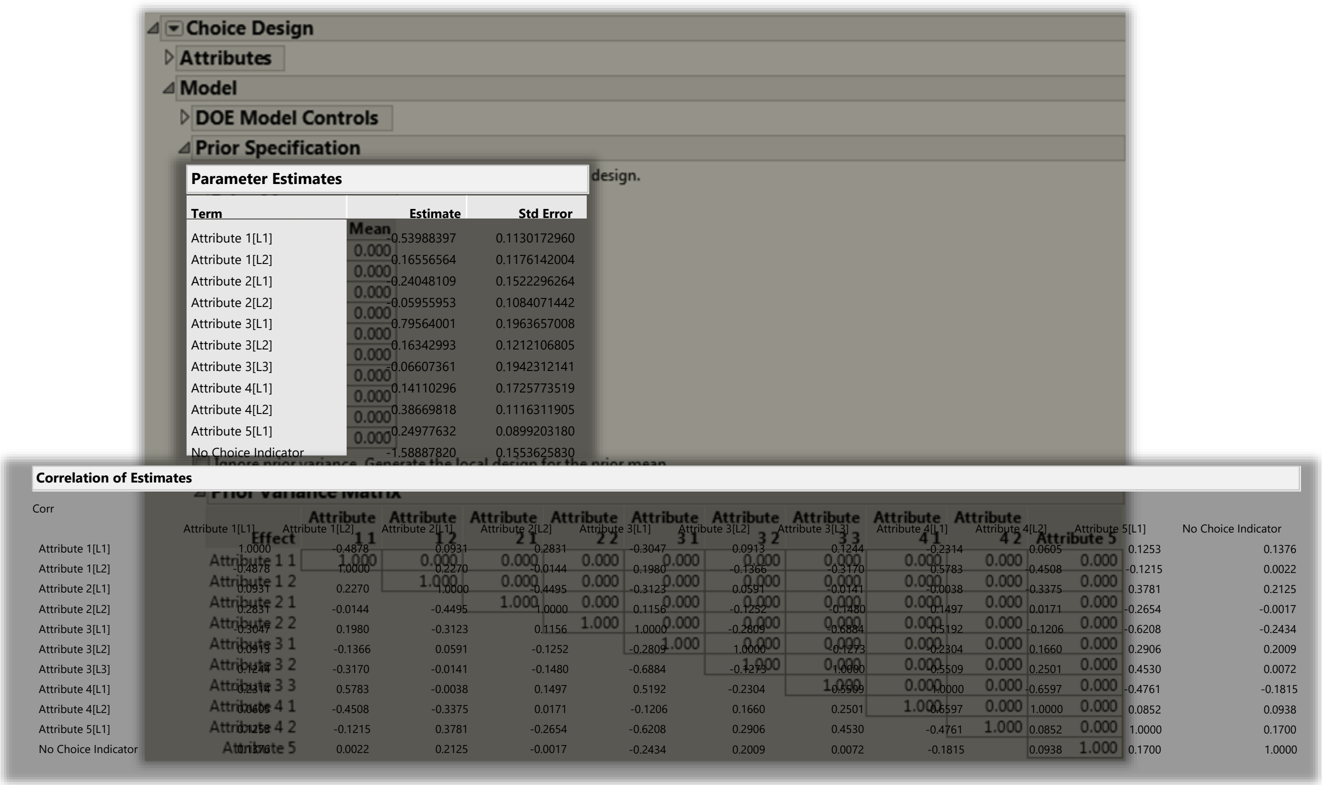
☐ Ignore prior variance. Generate the local design for the prior mean.

Prior Variance Matrix

Effect	Attribute 1 1	Attribute 1 2	Attribute 2 1	Attribute 2 2	Attribute 3 1	Attribute 3 2	Attribute 3 3	Attribute 4 1	Attribute 4 2	Attribute 5
Attribute 1 1	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Attribute 1 2		1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Attribute 2 1			1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Attribute 2 2				1.000	0.000	0.000	0.000	0.000	0.000	0.000
Attribute 3 1					1.000	0.000	0.000	0.000	0.000	0.000
Attribute 3 2						1.000	0.000	0.000	0.000	0.000
Attribute 3 3							1.000	0.000	0.000	0.000
Attribute 4 1								1.000	0.000	0.000
Attribute 4 2									1.000	0.000
Attribute 5										1.000

Piloting

- Coefficient estimates and covariance matrices computed from utility-neutral pilot studies can be used to create more effective and efficient question sets for follow-on surveys using local or Bayesian D-optimal design criteria.



Parameter Estimates

Term	Estimate	Std Error
Attribute 1[L1]	-0.53988397	0.1130172960
Attribute 1[L2]	0.16556564	0.1176142004
Attribute 2[L1]	-0.24048109	0.1522296264
Attribute 2[L2]	-0.05955953	0.1084071442
Attribute 3[L1]	0.79564001	0.1963657008
Attribute 3[L2]	0.16342993	0.1212106805
Attribute 3[L3]	-0.06607361	0.1942312141
Attribute 4[L1]	0.14110296	0.1725773519
Attribute 4[L2]	0.38669818	0.1116311905
Attribute 5[L1]	0.24977632	0.0899203180
No Choice Indicator	-1.58887820	0.1553625830

Correlation of Estimates

Corr	Attribute 1[L1]	Attribute 1[L2]	Attribute 2[L1]	Attribute 2[L2]	Attribute 3[L1]	Attribute 3[L2]	Attribute 3[L3]	Attribute 4[L1]	Attribute 4[L2]	Attribute 5[L1]	No Choice Indicator
Attribute 1[L1]	1.0000	-0.4878	0.0931	0.2831	-0.3047	0.0913	0.1244	-0.2314	0.0605	0.1253	0.1376
Attribute 1[L2]	-0.4878	1.0000	0.2270	-0.0144	0.1980	-0.1366	-0.3170	0.5783	-0.4508	-0.1215	0.0022
Attribute 2[L1]	0.0931	0.2270	1.0000	-0.4495	-0.3123	0.0591	-0.0141	-0.0038	-0.3375	0.3781	0.2125
Attribute 2[L2]	0.2831	-0.0144	-0.4495	1.0000	0.1156	-0.1252	-0.1480	0.1497	0.0171	-0.2654	-0.0017
Attribute 3[L1]	-0.3047	0.1980	-0.3123	0.1156	1.0000	-0.2809	-0.6884	0.5192	-0.1206	-0.6208	-0.2434
Attribute 3[L2]	0.0913	-0.1366	0.0591	-0.1252	-0.2809	1.0000	-0.1273	-0.2304	0.1660	0.2906	0.2009
Attribute 3[L3]	0.1244	-0.3170	-0.0141	-0.1480	-0.6884	-0.1273	1.0000	-0.5509	0.2501	0.4530	0.0072
Attribute 4[L1]	-0.2314	0.5783	-0.0038	0.1497	0.5192	-0.2304	-0.5509	1.0000	-0.6597	-0.4761	-0.1815
Attribute 4[L2]	0.0605	-0.4508	-0.3375	0.0171	-0.1206	0.1660	0.2501	-0.6597	1.0000	0.0852	0.0938
Attribute 5[L1]	0.1253	-0.1215	0.3781	-0.2654	-0.6208	0.2906	0.4530	-0.4761	0.0852	1.0000	0.1700
No Choice Indicator	0.1376	0.0022	0.2125	-0.0017	-0.2434	0.2009	0.0072	-0.1815	0.0938	0.1700	1.0000

Data Quality Challenges

- Several unforeseen data quality issues. Many initially arbitrary decisions not without consequence.

× Choices that are too difficult are ignored or answered arbitrarily. Choices that are too easy are uninformative.

× Unconventional answers cannot be included in math models.

× Uncontrolled milieu: surveys provided to complete over weekend, leaving subjects responsible for controlling test environment.

× Language used may not be common between occupational specialties, leading to survey bias.

× Other systemic errors: recall bias, data entry error/illegible subject handwriting, opportunistic sampling strategy.

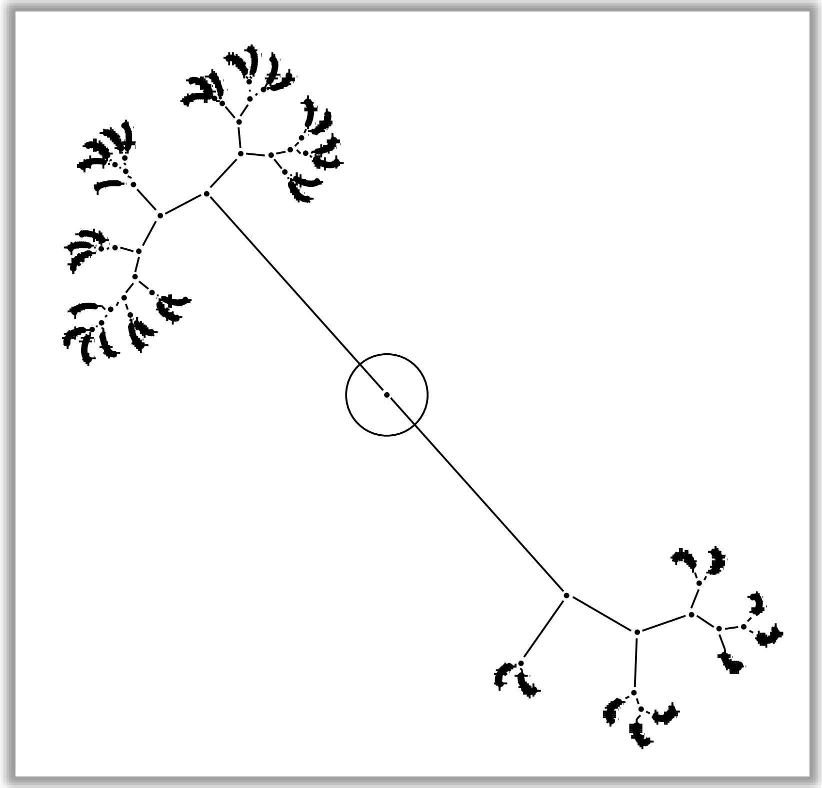
Best Practices

- ✓ Balance question difficulty to maximize information received.
- ✓ Provide instructions and background information that are explicit and unequivocal.
- ✓ Utilize large samples in lieu of focus-grouping. Filter responses using demographic information to minimize bias of opportunistic sampling.
- ✓ Limit number of questions to minimize fatigue; Randomize order to mitigate *impact*.
- ✓ Refrain from disruptive visuals. Do not draw subject attention away from content.
- ✓ Design surveys that are impartial to all target demographics.
- ✓ Conduct post-survey interviews when possible [or use free-form questions] to aid interpretation.
- ✓ Allowing "none of the above" more informative than forcing a choice.

Future Work

- Digitally-authenticated, web-based survey platform
  - Currently under development in partnership with ARDEC's Military Web Applications and Software Solutions (MWASS) group.
  - Once deployed, will solve many data quality challenges:
    - fully automated. streamlined data collection and entry.
    - deliver wider pool of respondents → segmentation opportunities
    - larger database → models that are both more accurate and more precise.

- Text-based methods for exploration of open ended response questions
  - singular value decomposition
  - hierarchical clustering



These efforts will further ARDEC insight into customer needs and will contribute great value to an already valuable tool.

Acknowledgments

- Advanced Breaching and Demolition Technologies (ABDT) Integrated Product Team (IPT)
  - Stephen Roberts, US Army ARDEC
  - Juan Rodriguez, US Army ARDEC
  - Brian Black, Maneuver Support Center of Excellence (MSCOE)

- MSCOE Combat Engineer School, Fort Leonard Wood, MO

References

- McFadden, D. (1974), “Conditional Logit Analysis of Qualitative Choice Behavior,” in P. Zarembka, ed., *Frontiers in Econometrics*, pp. 105–142.

- Firth, D. (1993), “Bias Reduction of Maximum Likelihood Estimates,” Biometrika 80:1, 27–38.

- Nelder, J.A. and Wedderburn, R.W.M. (1972), “Generalized Linear Models,” *Journal of the Royal Statistical Society*, Series A, 135, 370–384.

- Kessels, R., Jones, B., Goos, P. (2011), “Bayesian Optimal Designs for Discrete Choice Experiments with Partial Profiles,” *Journal of Choice Modelling*, 4(3), 52-74.

- Train, K.E. (2009), *Discrete Choice Methods and Simulation*, Cambridge University Press.