



Australia's First Official use of U.S. Design of Experiments (DOE): User Trials to Determine Best Enhanced Steyr Rifle Configuration

International Test & Evaluation Association Conference
Washington – August 2015

Contents

- Background: Design of Experiments – Assumed in U.S. audience
- Background on Defence Trial 933: Enhanced Steyr Rifle Ancillaries
- Use of DOE for overall trial design
- Conduct of Trial
- Example of Selecting Best Performing Ancillary: Reversionary Sight System
- Pugh Matrices: More Complex Selections
- Inclusiveness: Concern for Shooters of Different Ability & Gender
- Building Value for Money Argument: Enhanced Rifle Performance Against Baseline
- Validity Results of Test Deconstruction
- Lessons Learned: DOE Advantages & Disadvantages



Background on Defence Trial 933: Enhanced Steyr Rifle Ancillaries

LAND 125 Phase 3C is a project to deliver. dismantled close combatants & general combatants with an enhanced F88 Steyr rifle, including an improved ability to fit ancillaries to the rifle that enhance performance in a number of missions.

'The ancillaries suite comprises several elements that can be fitted by means of the STANAG 4694 rail system and provide surveillance, target acquisition and point engagement by day and night:

- *Enhanced Day Sight Capability.*
- *Backup/reversionary aiming capability – preference to be fitted concurrently with day sight.*
- *Weapon Support Capability (bipod) – fitted in place of grenade launcher or foregrip.*
- *Laser Aiming and Illumination Device (LAID) Capability – to be fitted concurrently with day sights and when in-line Image Intensifier is fitted to weapon.*
- *Laser Aiming and Illumination Range Finding Device (LAIRD) Capability – to be fitted concurrently with day sights and when in-line Thermal Imager or Image Intensifier is fitted to weapon.*
- *In-Line Image Intensification System (IIS) Capability – required to be fitted concurrently with Enhanced Day Sight (but not thermal imager).*
- *In-Line Thermal Imaging System (TIS) Capability – to be fitted concurrently with Enhanced Day Sight (but not Image Intensifier).*
- *Foregrip Capability.'*

(Thales Australia, 2015)



Background on Defence Trial 933: Preview T&E & Trial Objectives

When Preview T&E became mandatory in 2013 a mature enhanced rifle design had been developed & undergone significant contractor-led T&E, however:

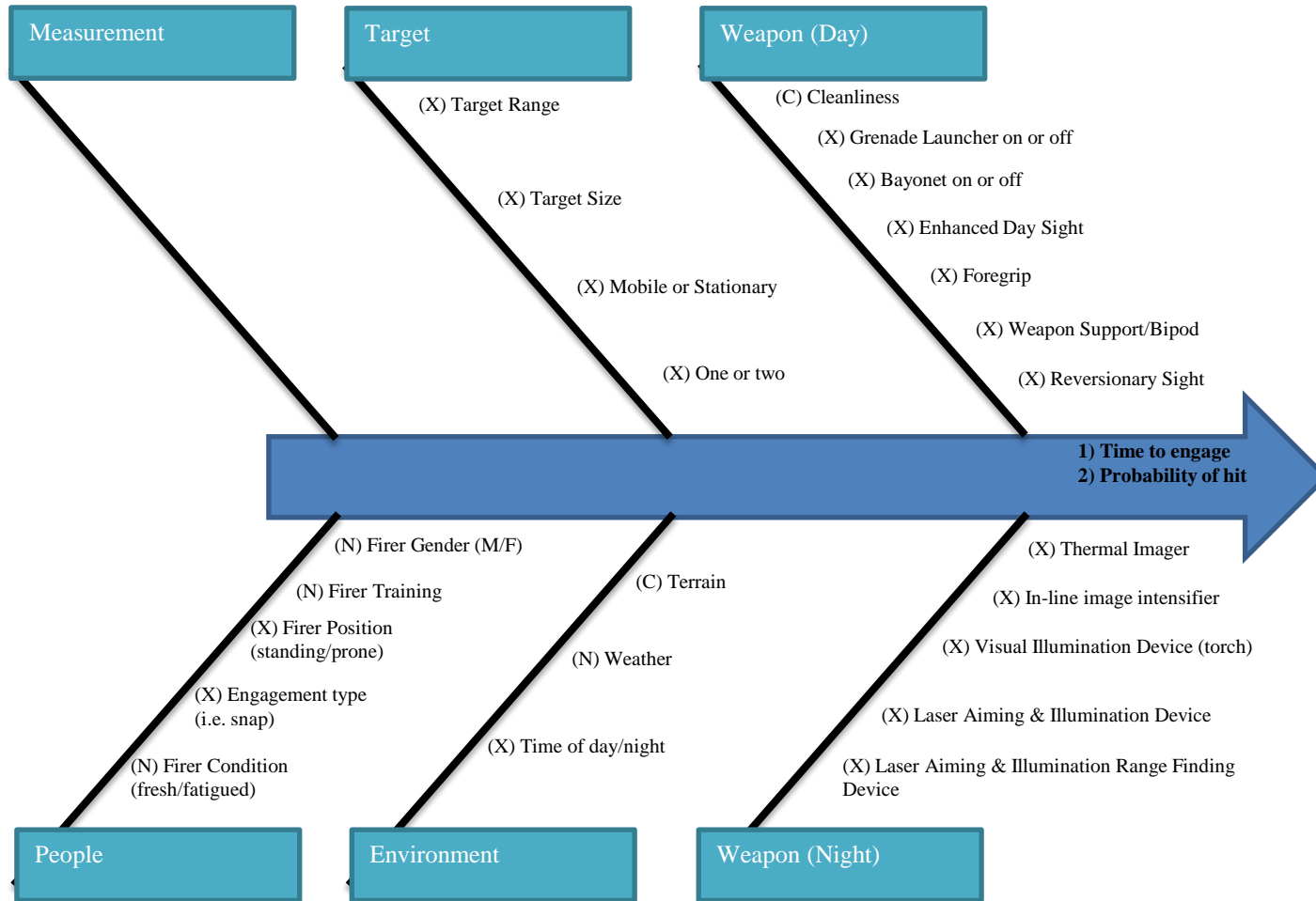
- not involving candidate ancillaries &
- not with representative users under the oversight of the operational test agency for Army.

Army, the Project Office, the Land Engineering Agency & ADTEO agreed to conduct a preview user trial, with the main purpose of '*selecting the preferred enhanced F88 ancillaries*' & secondary objectives, where possible, to:

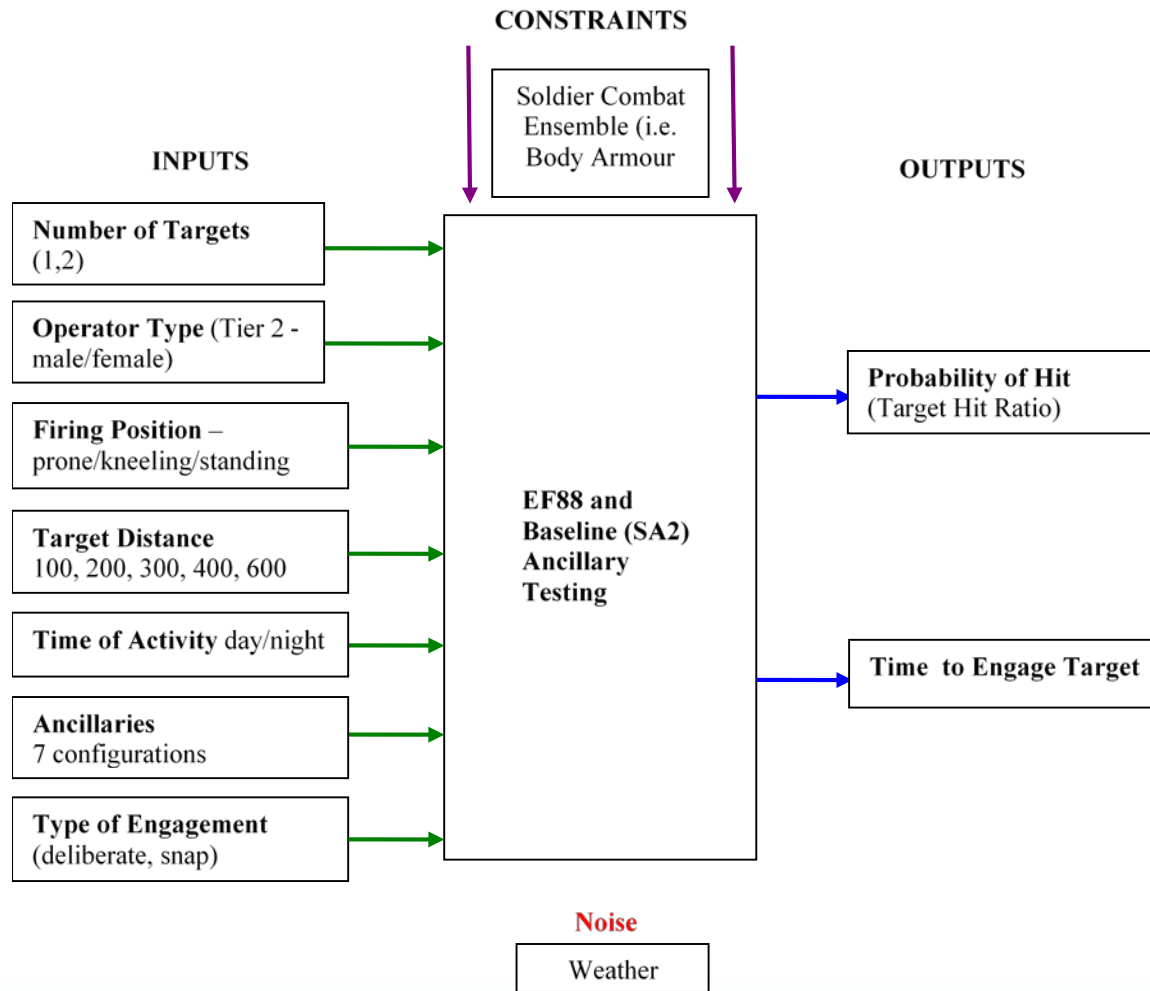
- Inform integration of the enhanced F88 mission system with the Soldier Combat System (i.e. the other kit a soldier must wear).
- Inform training considerations for introduction to service.
- Inform effectiveness of the collective employment of the enhanced F88 mission system.
- Compare enhanced F88 mission system performance against the current baseline F88 SA2.



Use of DOE for overall Trial Design



Use of DOE for overall Trial Design



Use of DOE for overall Trial Design

Problem Formulation - Test Parameters - Factor Table																No. Of factors = 16, All Combinations = 4,199,040	
Lbl	Factor Name	No. Lvl's	Dep. On	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6	Level 7	Level 8	Level 9	Level 10	Level 11	Level 12	Level 13	
A	Distance	5		100	200	300	400	600									
B	Shooter	2		M	F												
C	No of Targets	2		1	2												
D	Shooter Fire Position	3		Prone	Kneeling	Standing											
E	Time of Day	2		Day	Night												
F	Type of Engagement	2		Rapid	Deliberate												
G	Sights	3		Sight1	Sight2	Sight3											
H	LAIRD	2		Y	N												
I	Image Intensifier	3		II1	II2	II3											
J	Thermal	3		T1	T2	T3											
K	Foregrip	3		Grip1	Grip2	Grip3											
L	Torch	3		Torch1	Torch2	Torch3											
M	Bipod	3		Bipod1	Bipod2	Bipod3											
N	GLA	2		Y	N												
O	Bayonet	2		Y	N												
P	RSS	3		RSS1	RSS2	RSS3											



Use of DOE for overall Trial Design

Problem Formulation - Display Log - OA Test Plan

OA : 64 Tests

Deg of Freedom : 28
All Combinations : 4,199,040

Test No.	Dupl. Of Test	A : Distance	B : Shooter	C : No of Targets	D : Shooter Fire Position	E : Time of Day	F : Type of Engagement	G : Sights	H : LAIRD	I : Image Intensifier	J : Thermal	K : Foregrip	L : Torch	M : Bipod	N : GLA	O : Bayonet	P : RSS
1		100	M	1	Prone	Day	Rapid	Sight1	Y	II1	T1	Grip1	Torch1	Bipod1	Y	Y	RSS1
2		100	F	2	Prone	Night	Deliberate	Sight1	Y	II1	T1	Grip1	Torch1	Bipod1	Y	Y	RSS1
3		100	M	1	Prone	Night	Deliberate	Sight1	Y	II1	T3	Grip3	Torch3	Bipod3	N	N	RSS3
4		100	F	2	Prone	Day	Rapid	Sight1	Y	II1	T2	Grip2	Torch2	Bipod2	N	N	RSS2
5		100	M	2	Standing	Day	Deliberate	Sight3	N	II3	T1	Grip1	Torch3	Bipod1	Y	N	RSS2
6		100	F	1	Kneeling	Night	Rapid	Sight2	N	II2	T1	Grip1	Torch2	Bipod1	Y	N	RSS3
7		100	M	2	Kneeling	Night	Rapid	Sight2	N	II2	T2	Grip2	Torch1	Bipod3	N	Y	RSS1
8		100	F	1	Standing	Day	Deliberate	Sight3	N	II3	T3	Grip3	Torch1	Bipod2	N	Y	RSS1
9		300	F	2	Prone	Night	Deliberate	Sight3	N	II3	T2	Grip2	Torch2	Bipod1	N	Y	RSS1
10		300	M	1	Prone	Day	Rapid	Sight2	N	II2	T3	Grip3	Torch3	Bipod1	N	Y	RSS1
11		300	F	2	Prone	Day	Rapid	Sight2	N	II2	T1	Grip1	Torch1	Bipod2	Y	N	RSS2
12		300	M	1	Prone	Night	Deliberate	Sight3	N	II3	T1	Grip1	Torch1	Bipod3	Y	N	RSS3
13		300	F	1	Kneeling	Night	Rapid	Sight1	Y	II1	T3	Grip3	Torch1	Bipod1	N	N	RSS3
14		300	M	2	Standing	Day	Deliberate	Sight1	Y	II1	T2	Grip2	Torch1	Bipod1	N	N	RSS2
15		300	F	1	Standing	Day	Deliberate	Sight1	Y	II1	T1	Grip1	Torch2	Bipod2	Y	Y	RSS1
16		300	M	2	Kneeling	Night	Rapid	Sight1	Y	II1	T1	Grip1	Torch3	Bipod3	Y	Y	RSS1
17		200	M	1	Standing	Night	Deliberate	Sight2	N	II1	T1	Grip3	Torch2	Bipod1	Y	Y	RSS2



Use of DOE for overall Trial Design

These 64 test conditions (6400 rounds) should have been done in early part of the trial (i.e. first week) to **screen** for what factors were correlated & likely to be significant.

Test design team **balked** at complexity of full multifactor design - concerned data would not yield specific specification compliance outcomes (test integration tension)

To keep the preview T&E integrated, the team's experts in T&E of small arms reduced the test plan to a series of about **seven smaller multifactor tests** that were designed only to select each ancillary. These two or three factor test designs tested all combinations within that smaller multifactor test.

The assumptions in decomposing the T&E were:

- Gender was unlikely to be a significant factor in selecting ancillaries.
- Fitment of the grenade launcher assembly & bayonet was unlikely to be a significant factor in selecting ancillaries.
- All ancillaries would have independent effects from each other.
- Engagement conditions like range of target, shooter position & prevailing light conditions would be tailored to only those most important to each ancillary.



Use of DOE for overall Trial Design

Example test design for two-factor Reversionary Sight System (RSS) selection

Test No.	A : Distance & Engagement	B : RSS
1	100m Deliberate	Baseline
2	100m Deliberate	RSS1
3	100m Deliberate	RSS2
4	100m Deliberate	RSS3
5	100m Snap	Baseline
6	100m Snap	RSS1
7	100m Snap	RSS2
8	100m Snap	RSS3
9	300m Snap	Baseline
10	300m Snap	RSS1
11	300m Snap	RSS2
12	300m Snap	RSS3



Conduct of Trial

Several standard preview trial methodologies were included (without change by DOE):

- Integration & fitment exercise at Victoria Barracks Melbourne to confirm ancillaries fit to all versions of rifle with logical procedures, no safety concerns & no interference using all permutations of soldier combat ensemble & varying anthropometrics.
- Training evaluation exercise with experienced training soldiers first receiving train the trainer from manufacturers & then qualifying the representative operational soldiers.
- Collective user evaluation by small squads up to a platoon size working through representative operational scenarios that required them to use the enhanced rifle & ancillaries collectively (i.e. an ambush).

That is, selection of the best performing ancillaries based on more than performance. Indeed, at source selection board will include factors of a cost & contractual nature

Trial conducted in two phases with 3rd Brigade Townsville on High Range instrumented rifle range:

- Late 2014 – day ancillary evaluation
- Early 2015 day & night ancillary evaluation.



Example of Selecting Best Performing Ancillary: Reversionary Sight System

Analyse of two-factor test with DOE PRO XL© showed:

- significant variation between three engagement scenarios in both
 - PH ($F_{2,228} = 146.982$, $p < 0.000$) &
 - TE ($F_{1,152} = 42.037$, $p < 0.000$);
- significant variation between configurations in PH ($F_{3,228} = 74.114$, $p < 0.000$) but not TE ($F_{3,152} = 0.594$, $p = 0.620$),
 - mainly attributable to improved PH between the baseline rifle & new RSS without significant compromise in TE;
- no significant correlation between configuration and engagement type (i.e. AB) for either PH ($F_{6,228} = 0.138$, $p = 0.991$) or TE ($F_{3,152} = 0.815$, $p = 0.488$).

Individual F-test & t-test were calculated between configurations for each engagement type & results added to a **Pugh decision matrix**

(Reagan & Kiemele, 2008, pp. 237-242)



Example Selection

Requirements	weight ▼	Baseline	RSS1	RSS2	RSS3
Mean Probability of Hit, 100m deliberate	1	0.526	0.945	0.975	0.985
Standard deviation Probability of Hit, 100m deliberate	1	0.298	0.188	0.112	0.049
Mean Probability of Hit, 100m snap	1	0.539	0.955	0.915	0.945
Standard deviation Probability of Hit, 100m snap	1	0.299	0.06	0.23	0.1
Mean Probability of Hit, 300m snap	1	0.101	0.48	0.52	0.53
Standard deviation Probability of Hit, 300m snap	1	0.13	0.254	0.253	0.2
Mean time to engage, 100m snap (msec)	1	1818	1819	1670	1637
Standard Deviation time to engage, 100m snap (msec)	1	872	498	715	661
Mean time to engage, 300m snap (msec)	1	2226	2657	2629	2641
Standard deviation time to engage, 300m snap (msec)	1	812	766	812	969
	# Winners	2	4	0	4

Scoring

Basic Pugh:

-1 = worse than baseline
 0 = same as baseline
 +1 = better than baseline

Detailed Pugh:

-3 = far worse than baseline +3 = far better than baseline
 -2 = much worse than baseline +2 = much better than baseline
 -1 = worse than baseline +1 = better than baseline
 0 = same as baseline



Example Selection

Requirements	weight ▼	Baseline	RSS1	RSS2	RSS3
Mean Probability of Hit, 100m deliberate	1	0	3	3	3
Standard deviation Probability of Hit, 100m deliberate	1	0	1	2	3
Mean Probability of Hit, 100m snap	1	0	3	3	3
Standard deviation Probability of Hit, 100m snap	1	0	3	1	2
Mean Probability of Hit, 300m snap	1	0	3	3	3
Standard deviation Probability of Hit, 300m snap	1	0	-2	-2	-1
Mean time to engage, 100m snap (msec)	1	0	0	1	1
Standard Deviation time to engage, 100m snap (msec)	1	0	3	1	2
Mean time to engage, 300m snap (msec)	1	0	0	0	0
Standard deviation time to engage, 300m snap (msec)	1	0	0	0	0
Sum of Values		0	14	12	16

Best performing reversionary sight is RSS3 followed by RSS1, even when weightings adjusted to:

- give greater emphasis to longer range over shorter range,
- improved means over reduced variance, &
- to probability of hit over time to engage
- RSS1 = 43, RSS2 = 40, RSS3 = 48



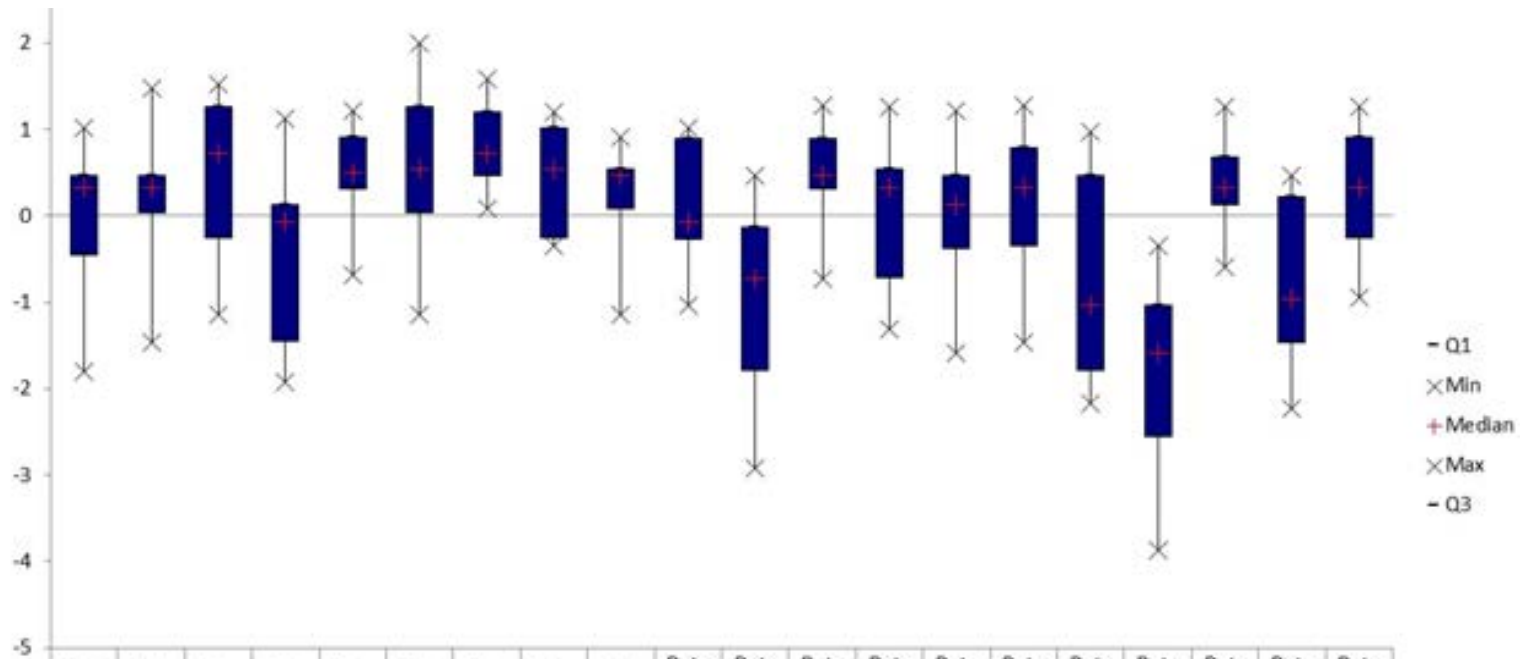
Pugh Matrices: More Complex Selections

Requirements	weight ▼	Baseline	EDS (A)	EDS (B)	EDS (C)
Stand Snap 100m					
Mean Probability of Hit	4				
Standard deviation Probability of Hit	2	0.215	0.127	0.156	0.147
MeanTime to Engage (msec)	2				
Standard deviation Time to Engage (msec)	1	527	666	598	584
Kneel Snap 200m					
Mean Probability of Hit	4				
Standard deviation Probability of Hit	2	0.225	0.194	0.273	0.201
MeanTime to Engage (msec)	2				
Standard deviation Time to Engage (msec)	1	803	1001	786	1011
Sit Rapid 200m					
Mean Probability of Hit	4				
Standard deviation Probability of Hit	2	0.261	0.224	0.279	0.150
MeanTime to Engage (msec)	2				
Standard deviation Time to Engage (msec)	1	631	715	746	580
Standing Rapid Realign L-R 100m					
Mean Probability of Hit	4				
Standard deviation Probability of Hit	2	0.229	0.125	0.127	0.238
MeanTime to Engage (msec)	2				
Standard deviation Time to Engage (msec)	1	291	291	300	363
Stand Moving Target 100m					
Mean Probability of Hit	4				
Standard deviation Probability of Hit	2	0.268	0.208	0.179	0.233
MeanTime to Engage (msec)	2				
Standard deviation Time to Engage (msec)	1	500	778	237	324
Prone Range Realign 100m					
Mean Probability of Hit	4				
Standard deviation Probability of Hit	2	0.139	0.115	0.153	0.134
MeanTime to Engage (msec)	2				
Standard deviation Time to Engage (msec)	1	391	435	534	527
Prone Range Realign 400m					
Mean Probability of Hit	4				
Standard deviation Probability of Hit	2	0.221	0.200	0.297	0.243
MeanTime to Engage (msec)	2				
Standard deviation Time to Engage (msec)	1	477	493	465	505
Prone Snap 100m					
Mean Probability of Hit	4				
Standard deviation Probability of Hit	2	0.105	0.000	0.045	0.045
MeanTime to Engage (msec)	2				
Standard deviation Time to Engage (msec)	1	511	483	427	430
Prone Snap 200m					
Mean Probability of Hit	4				
Standard deviation Probability of Hit	2	0.174	0.214	0.154	0.279
MeanTime to Engage (msec)	2				
Standard deviation Time to Engage (msec)	1	691	715	723	688
Prone Snap 300m					
Mean Probability of Hit	4				
Standard deviation Probability of Hit	2	0.302	0.252	0.223	0.268
MeanTime to Engage (msec)	2				
Standard deviation Time to Engage (msec)	1	909	766	940	795
Prone Snap 400m					
Mean Probability of Hit	4				
Standard deviation Probability of Hit	2	0.280	0.275	0.311	0.370
MeanTime to Engage (msec)	2				
Standard deviation Time to Engage (msec)	1	861	843	1013	990
# Winners	5		14	15	11

Requirements	weight ▼	Baseline	EDS (A)	EDS (B)	EDS (C)
Stand Snap 100m					
Mean Probability of Hit	4	0	1	3	1
Standard deviation Probability of Hit	2	0	3	1	1
MeanTime to Engage (msec)	2	0	1	1	2
Standard deviation Time to Engage (msec)	1	0	-1	-1	-1
Kneel Snap 200m					
Mean Probability of Hit	4	0	0	1	2
Standard deviation Probability of Hit	2	0	1	0	0
MeanTime to Engage (msec)	2	0	-1	-1	1
Standard deviation Time to Engage (msec)	1	0	-1	0	-1
Sit Rapid 200m					
Mean Probability of Hit	4	0	0	1	3
Standard deviation Probability of Hit	2	0	1	-1	3
MeanTime to Engage (msec)	2	0	0	1	-1
Standard deviation Time to Engage (msec)	1	0	-1	-1	1
Standing Rapid Realign L-R 100m					
Mean Probability of Hit	4	0	0	1	1
Standard deviation Probability of Hit	2	0	3	3	0
MeanTime to Engage (msec)	2	0	-1	0	-1
Standard deviation Time to Engage (msec)	1	0	0	0	-1
Stand Moving Target 100m					
Mean Probability of Hit	4	0	0	0	0
Standard deviation Probability of Hit	2	0	1	2	1
MeanTime to Engage (msec)	2	0	-1	2	1
Standard deviation Time to Engage (msec)	1	0	-3	3	2
Prone Range Realign 100m					
Mean Probability of Hit	4	0	0	0	0
Standard deviation Probability of Hit	2	0	1	-1	0
MeanTime to Engage (msec)	2	0	0	2	1
Standard deviation Time to Engage (msec)	1	0	0	-1	-1
Prone Range Realign 400m					
Mean Probability of Hit	4	0	1	1	2
Standard deviation Probability of Hit	2	0	1	-1	0
MeanTime to Engage (msec)	2	0	-2	-3	-2
Standard deviation Time to Engage (msec)	1	0	0	0	0
Prone Snap 100m					
Mean Probability of Hit	4	0	2	1	1
Standard deviation Probability of Hit	2	0	3	2	2
MeanTime to Engage (msec)	2	0	1	1	1
Standard deviation Time to Engage (msec)	1	0	0	1	1
Prone Snap 200m					
Mean Probability of Hit	4	0	-1	1	-1
Standard deviation Probability of Hit	2	0	-1	0	-3
MeanTime to Engage (msec)	2	0	1	1	2
Standard deviation Time to Engage (msec)	1	0	0	-1	1
Prone Snap 300m					
Mean Probability of Hit	4	0	0	2	1
Standard deviation Probability of Hit	2	0	1	2	1
MeanTime to Engage (msec)	2	0	1	1	2
Standard deviation Time to Engage (msec)	1	0	2	-1	1
Prone Snap 400m					
Mean Probability of Hit	4	0	2	2	1
Standard deviation Probability of Hit	2	0	0	-1	-2
MeanTime to Engage (msec)	2	0	0	-1	1
Standard deviation Time to Engage (msec)	1	0	1	-1	-2
Sum of Values	0	43	70	64	

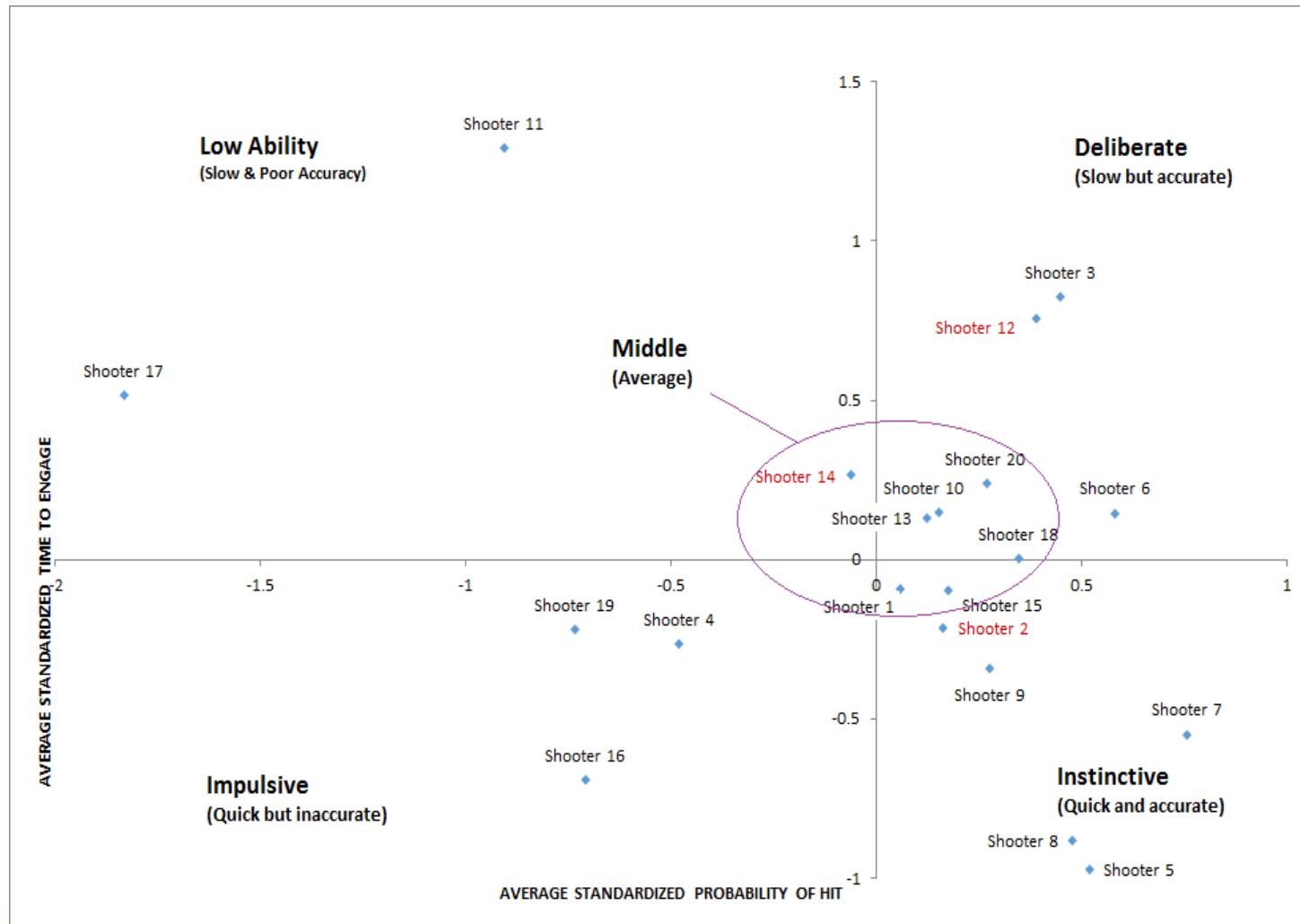


Inclusiveness: Concern for Shooters of Different Ability & Gender

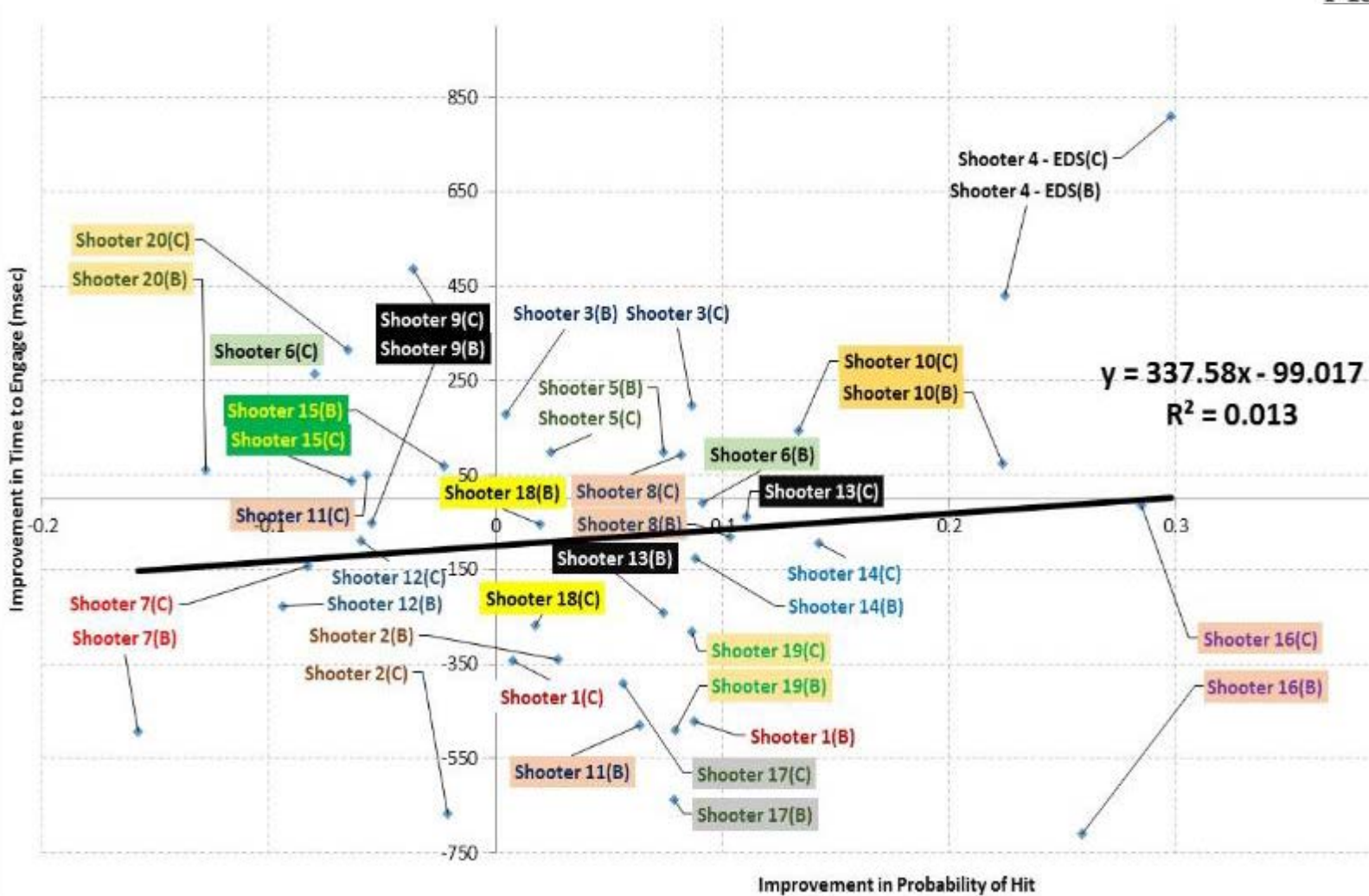


	Data Set #1	Data Set #2	Data Set #3	Data Set #4	Data Set #5	Data Set #6	Data Set #7	Data Set #8	Data Set #9	Data Set #10	Data Set #11	Data Set #12	Data Set #13	Data Set #14	Data Set #15	Data Set #16	Data Set #17	Data Set #18	Data Set #19	Data Set #20
- Q1	-0.458	0.0384	-0.256	-1.451	0.3192	0.0431	0.4657	-0.256	0.0816	-0.267	-1.783	0.3192	-0.707	-0.385	-0.353	-1.783	-2.552	0.1282	-1.459	-0.256
× Min	-1.795	-1.459	-1.136	-1.92	-0.686	-1.139	0.0816	-0.353	-1.136	-1.039	-2.924	-0.725	-1.306	-1.592	-1.459	-2.177	-3.864	-0.592	-2.238	-0.943
+ Median	0.3249	0.3249	0.7279	-0.074	0.4966	0.535	0.7279	0.535	0.4657	-0.069	-0.725	0.4657	0.3249	0.1282	0.3249	-1.031	-1.592	0.3249	-0.955	0.3249
× Max	1.0184	1.48	1.5282	1.1267	1.2144	1.998	1.5799	1.2082	0.917	1.0184	0.4657	1.2799	1.2606	1.2144	1.2776	0.9756	-0.353	1.2606	0.4657	1.2606
- Q3	0.4657	0.4657	1.2606	0.1288	0.917	1.2606	1.2082	1.0139	0.535	0.8947	-0.125	0.8947	0.5477	0.4657	0.7833	0.4657	-1.031	0.6735	0.2203	0.917

Inclusiveness: Different Ability & Gender



Inclusiveness: Different Ability & Gender



Inclusiveness: Different Ability & Gender

Five sub-groups of different shooter ability proposed & checked for improvement in both responses:

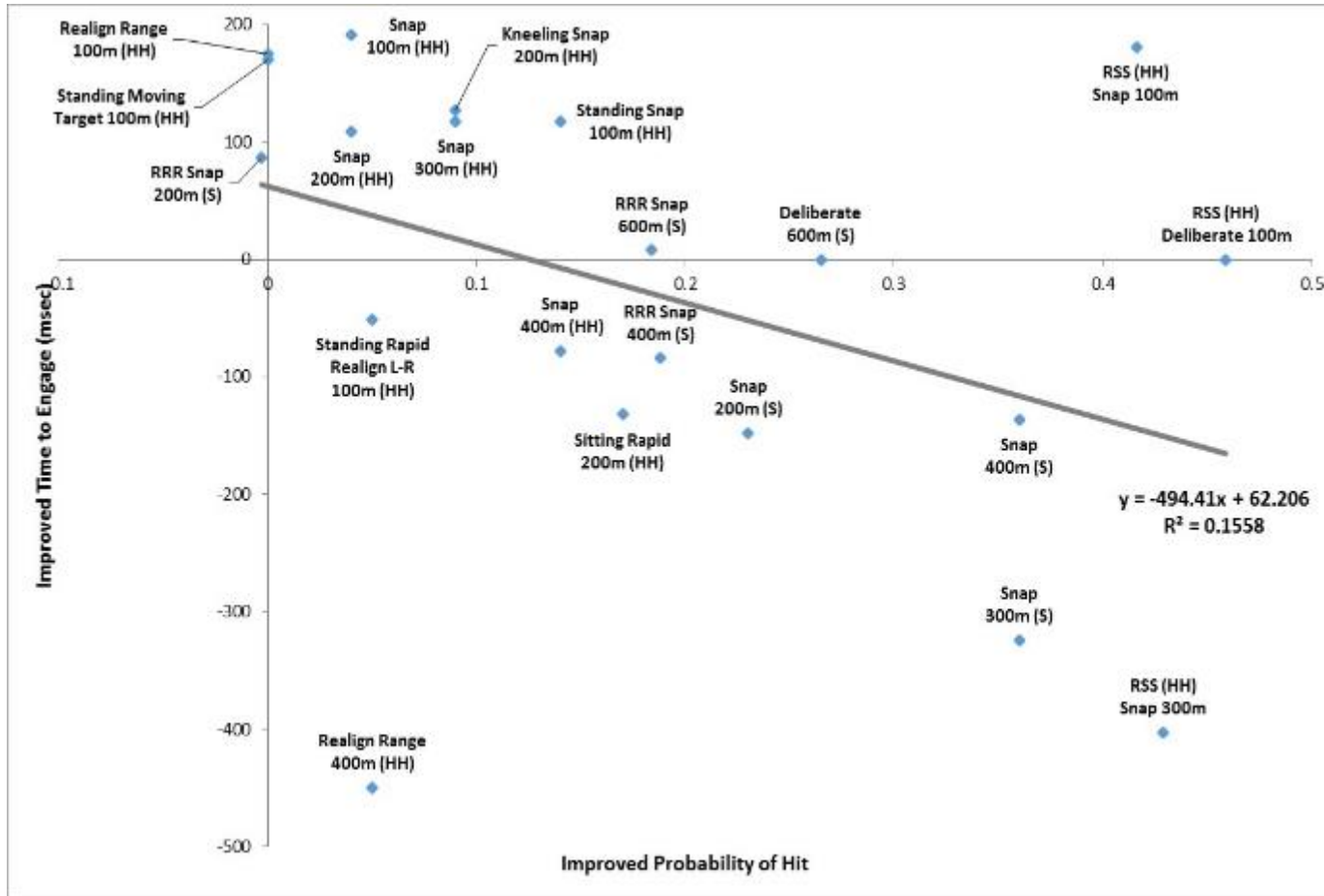
- Results mixed for *instinctively good* shooters and the *slow deliberate* shooters, with about half doing better & half doing worse, meaning there may be some negative reaction from shooters used to doing well in groups
- For *impulsive* shooters and the *middle average* shooters, all improve accuracy while most did so at a slight expense to timeliness.

Three female shooters participated:

- is 15%, representative of females in ADF
- each was of a different ability sub-group,
- due low numbers had to be statistically matched to males of similar ability
- generally no significant difference from males except for when using the EDS(C) configuration which was found to cause females to take significantly longer to engage than their matched male counterparts.



The Value for Money Argument: Enhanced Rifle Performance Against Baseline



Notes:

Prone unless stated

(HH) = handheld

(S) = supported

RRR = Rapid Random Realignment

RSS = Reversionary sights used,

L-R = left-right

Validity Results of Test Deconstruction

Grenade Launchers Assemblies (GLA) were fitted to four rifles throughout.

Such low numbers, mixed in **without dedicated factor screening** makes valid statistical analysis **difficult**

Using shooters of matched ability enabled two factor statistical comparison indicating there **were significant differences** in both PH & TE due to GLA.

Further examination by individual t-test & F-test **isolated** the significant GLA **effect to:**

- interacting with EDS(A) effect to negatively effect PH, &
- interact with EDS(C) effect to be likely to positively effect TE

For Gender, using shooters of matched ability enabled two factor statistical comparison indicating there **were** significant differences only in TE.

Further examination by individual t-test & F-test **isolated** significant effect on gender **to:**

- EDS(C) causing females to take longer TE



Lessons Learned: DOE Advantages & Disadvantages

- DOE requires both professional training and coaching soon after training so as to reinforce techniques – training has been available, looking at options for the coaching.
- Integrating developmental & operational T&E does bring tension & compromise, but will benefit project in contract & capability realisation
- Should have done a DOE-designed screen of significant factors (1 week, 6400 rounds) before finalising test plans
- Even some smaller three-factor tests would have used less runs if structured on rdExpert © before finalising tests.
- Constructing test files using DOE PRO XL© before trial would have greatly sped up analysis afterwards.
- Pugh Matrices are excellent tool for selection trials, easy to use, structured & logical
- Having baseline capability in the testing provides a great reference for ability & to keep the VFM perspective on the overall advantage, not just margins between new choices
- Inclusiveness is a valuable concept to bring into DOE teaching & analysis wherever diverse operators are involved
- There is nothing intrinsically stopping Australia following the US lead on DOE in Defence T&E:
 - extra complexity of test techniques is offset by the more professional outcomes & efficiency saving from integrated T&E.



Questions

