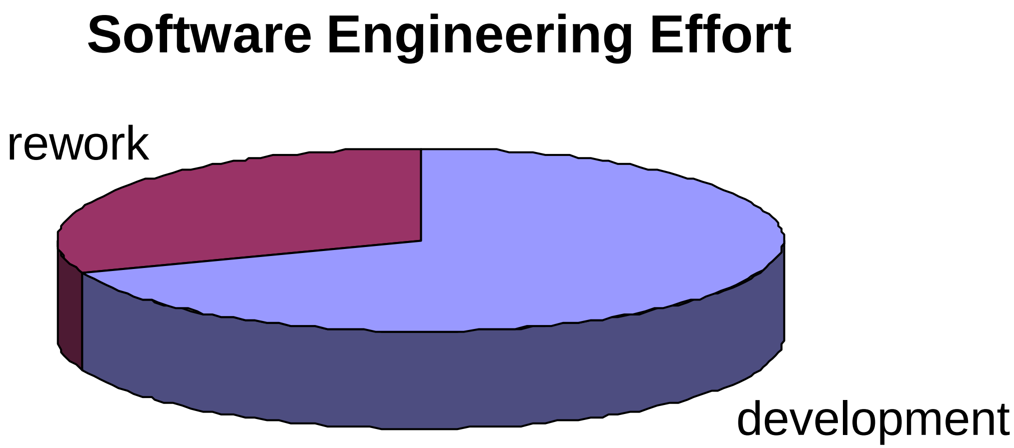
Statistical Tune Up of the Peer Review Engine

Tom Lienhard
Thomas_G_Lienhard@Raytheon.com

November, 2009

Overview

Project Definition: Reduce Rework by Reducing Defect Leakage
Currently, over 30% of the Software Engineering effort is consumed reworking products already deemed “fit-for-purpose”. A major contributor to this is defect leakage. Defect leakage is calculated as a percentage by summing the defects attributable to a specific phase that are detected in later phases divided by the total number of defects attributable to that phase. Defect leakage is a good indicator of the quality of the different phases of the software process. Defect leakage for the some software development phases is as high as 75%, where as our goal is set at 20%. Not catching and correcting defects at the earliest point in the process leads to cost and budget over-runs due to excessive rework. By investigating what types of defects go undetected during the various phases, corrections can be introduced into the process to help identify the top defect types.



Agenda – Six Sigma Toolbox Examples

Thought Process Map

Process Map

Failure Mode and Effect Analysis

What Was Learned So Far

Product Scorecard

What Was Learned So Far – Part 2

Improvement Goal

Distributional Characterization of Data

What Was Learned So Far – Part 3

DOE Conducted

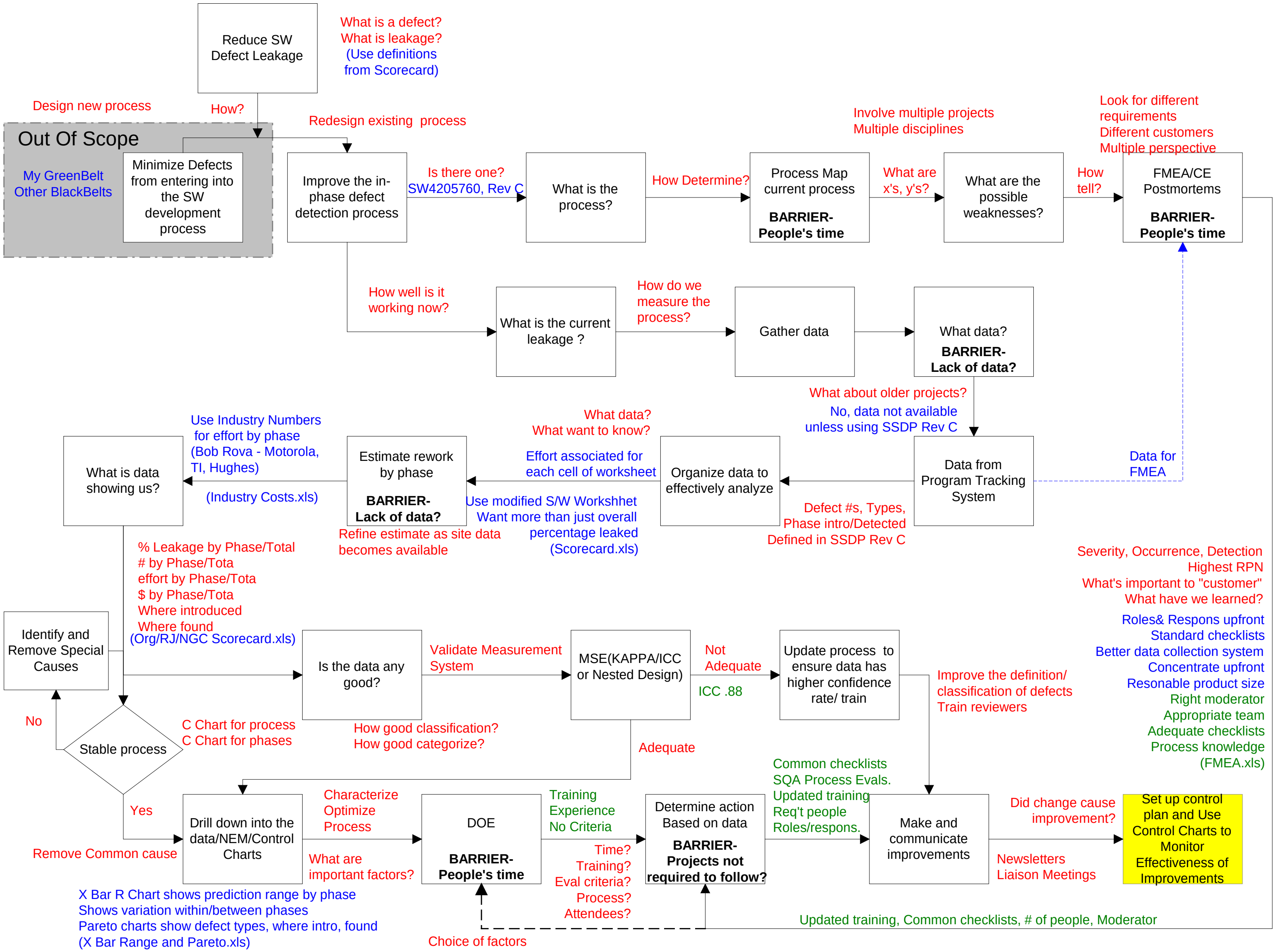
What Was Learned So Far – Part 4

Results

Conclusion

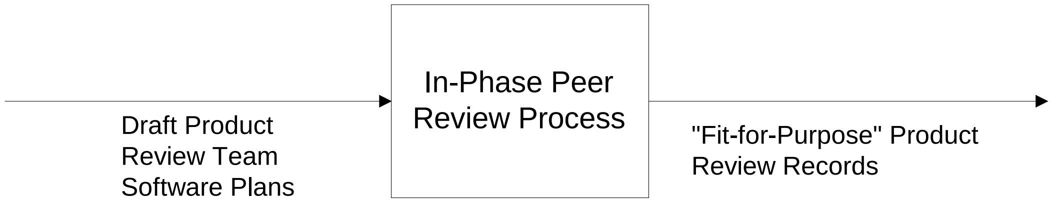
Thought Process Map – Where Are We Headed?

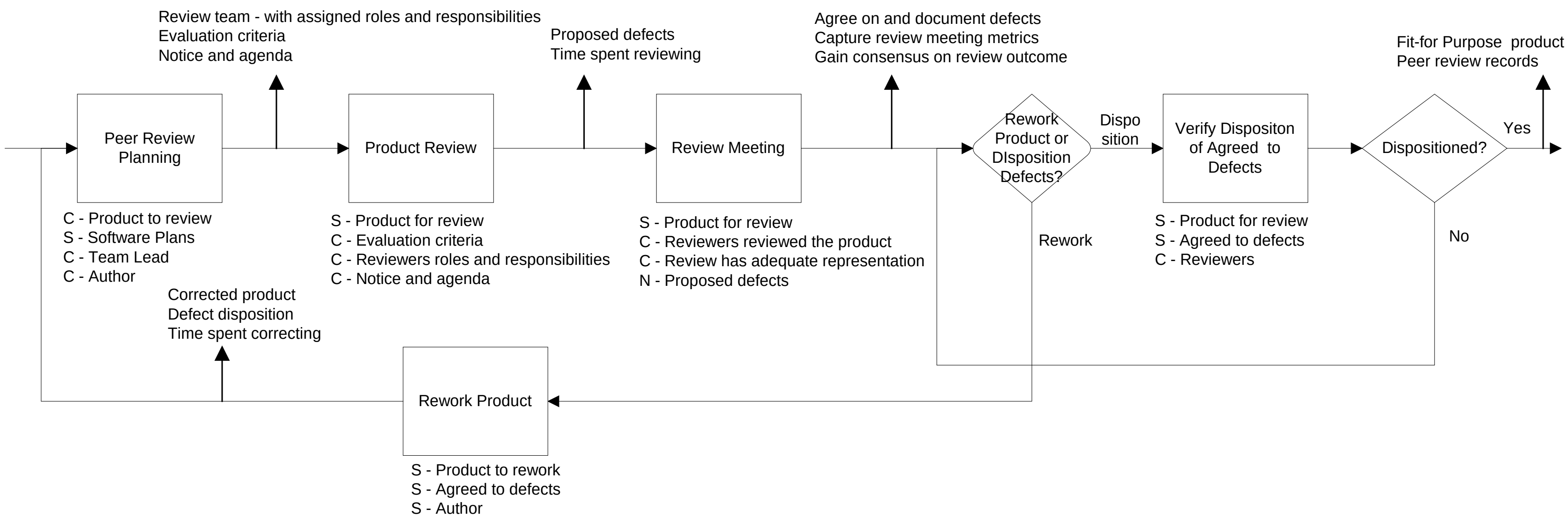
Raytheon



Green - Update Red - Question or expected result Blue - Answer or actual result **Underlined - Barrier**

Process Map – Walk the Process





X - Critical (statistical proven critical)
N - Noise (can't or choose not to control)
S - SOP (the standard way to do it)
C - Controllable (can be changed to see effect)

Failure Mode and Effect Analysis – How Can We Mess This Up?

Raytheon

Process or Product Name:	In-Phase Peer Review	Prepared by: Tom Lienhard & Team	Page ____ of ____
Responsible:	Tucson Software Engineering Process Group	FMEA Date (Orig) 23 Sept 99 (Rev) 29 Sept 99	

Process Step/Input	Potential Failure Mode	Potential Failure Effects	SEV	Potential Causes	OCC	Current Controls	DET	RPN	Actions Recommended
What is the process step/ Input under investigation?	In what ways does the process step go wrong?	What is the impact on the Key Output Variables (Customer Requirements) or internal requirements?	How Severe is the effect to the cusotmer?	What causes the process step to go wrong?	How often does cause or FM occur?	What are the existing controls and procedures (inspection and test) that prevent eith the cause or the Failure Mode? **Should include an SOP number.**	How well can you detect cause or FM?		What are the actions for reducing the occurrance of the Cause, or improving detection? **Should have actions only on high RPN's or easy fixes.**
Peer Review Planning	No Review Team identified upfront w/review package(roles & responsibilities)	Product not reviewed by appropriate disciplines	6	SW plans do not require this	1	SEPG/SQA and peer review of plans	1	6	None
			6	Lack of process awareness	6	Moderator to ensure review package complete (contains reviewers)	6	216	Re-train moderator and conduct process evaluations
	No product evaluation criteria identified with review package	Product not reviewed to customer and/or process requirements	9	SW plans do not require this	1	SEPG/SQA and peer review of plans	1	9	None
			9	Lack of process awareness	6	Moderator to ensure review package complete (contains review criteria)	6	324	Re-train moderator and conduct process evaluations
	No notice or agenda with review package	Team not able to give adequate review time	6	SW plans do not require this	1	SEPG/SQA and peer review of plans	1	6	None
			6	Lack of process awareness	6	Moderator to ensure review package complete (contains notice and agenda)	6	216	Re-train moderator and conduct process evaluations
Product Review	Product not reviewed	Defects not found	9	Adequate time not given to review	6	Moderator and SQD ensure adequate time was given to review product	3	162	None - cultural thing
			9	No or inadequate evaluation criteria	6	Moderator to ensure review package complete (contains evaluation criteria)	9	486	Create generic checklists for site (to highest level) and conduct process evaluations
			9	Inappropriate reviewers	6	Moderator to ensure appropriate reviewers	9	486	Update command media to identify required participants on notice & agenda and conduct process evluations
	Metrics not captured	Organization quantitative data incorrect/incomplete	3	SW plans do not require this	1	SEPG/SQA and peer review of plans	1	3	None
			3	Lack of process awareness	9	Moderator to ensure review process followed (metrics captured)	6	162	Re-train moderator and conduct process evaluations

Look, a legend!

Severity
9 - Defects go to customer
6 - Defects cause rework
3 - Data not collected
1 - No harm/no foul

Occurrence:
9 - Regular occurrence
6 - Occurs more than occasionally
3 - Occurs occasionally
1 - Rare occurrence

Detection:
9 - Nothing in place
6 - Based on individuals
3 - Check in place, usually works
1 - Check in place & working

Raytheon

What Was Learned, ….. So Far

Potential causes (factors) high RPN which kept showing up over and over on the FMEA:

- Inappropriate review team ("wrong" moderator, dominant, inexperienced, or yes-people made up the team)
- Lack of process awareness (both unintentional and deliberate)
- No or inadequate review criteria (review what is there not what is missing, biased review based on experience with phase)

Plan to minimize the occurrence and increase the detection:

- Update the process to highlight required participants, their roles and responsibilities on the Notice and Agenda
- Roll-out Peer Review training
- Have SQA perform peer review process evaluations
- Generate common evaluation criteria for all software products that can be used across the entire organization

Use what was learned about factors as an input into DOE

Raytheon

Product Scorecard

Number of defects identified by phase introduced and phase detected (Modified Software Worksheet from Product Scorecard)

Phase Detected

Phase Introduced

	Planning	Customer	Rqmts. Analysis	Design	Implementation	Test	Formal Test	Customer Before	TOTAL	Leaked
Planning	29	0	0	0	0	0	0	0	29	0
Customer	0	12	2	0	2	0	5	0	21	9
Rqmts. Analysis	0	0	61	14	29	26	71	1	202	141
Design	0	0	1	323	82	29	38	2	475	151
Implementation	0	0	1	5	220	43	44	10	323	97
Test	0	0	0	2	1	249	30	0	282	30
Formal Test	0	0	0	0	0	13	597	0	610	0
Customer Before	0	0	0	0	0	0	1	4	5	0
TOTAL	29	12	65	344	334	360	786	17	1947	428

Raytheon

Not All Defects Are Created Equal

Industry Standard* Cost to Detect and Correct Defects (in days)

Phase Detected

Phase Introduced

	Planning	Customer	Rqmts. Analysis	Design	Implementation	Test	Formal Test	Customer Before
Planning	0.07	0.1	0.19	0.17	0.67	1.23	1.4	0.54
Customer	0.02	0.15	0.7	1	2.03	2.7	3.23	3.07
Rqmts. Analysis	0	0.1	0.12	0.86	1.13	1.6	0.79	1.65
Design	0	0.11	0.13	0.13	0.1	0.8	3.13	2.64
Implementation	0	0	0.17	0.1	0.7	2.1	2.02	2.33
Test	0	0	0	0.08	0.03	0.08	0.15	0.16
Formal Test	0	0	0	0	0.04	0.18	0.25	0.58
Customer Before	0	0	0	0	0	0	2.7	3.4

* Motorola, Texas Instruments, Hughes Software Implementation of Six Sigma

Raytheon

Cost of Rework Due to Defects (in Days)

Phase Detected

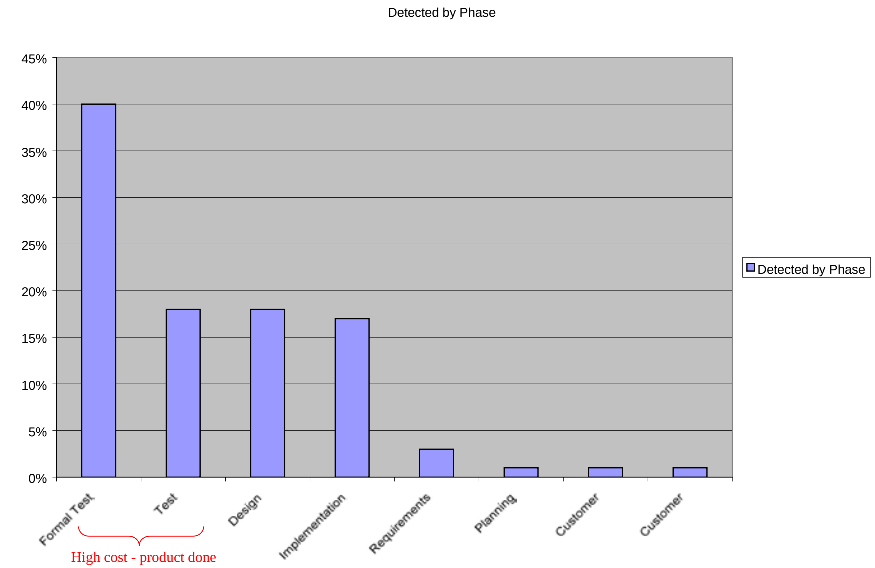
Phase Introduced

	Planning	Customer	Rqmts. Analysis	Design	Implementation	Test	Formal Test	Customer Before	TOTAL	Leaked
Planning	2.03	0	0	0	0	0	0	0	2.03	0
Customer	0	1.8	1.4	0	4.06	0	16.15	0	23.41	21.61
Rqmts. Analysis	0	0	7.32	12.04	32.77	41.6	56.09	0.79	150.61	143.29
Design	0	0	0.13	41.99	8.2	23.2	118.94	5.28	197.74	155.75
Implementation	0	0	0.17	0.5	154	90.3	88.88	23.3	357.15	203.15
Test	0	0	0	0.16	0.03	19.92	4.5	0	24.61	4.69
Formal Test	0	0	0	0	0	2.34	149.25	0	151.59	2.34
Customer Before	0	0	0	0	0	0	2.7	13.6	16.3	13.6
TOTAL	2.03	1.8	9.02	54.69	199.06	177.36	436.51	42.97	923.44	544.43

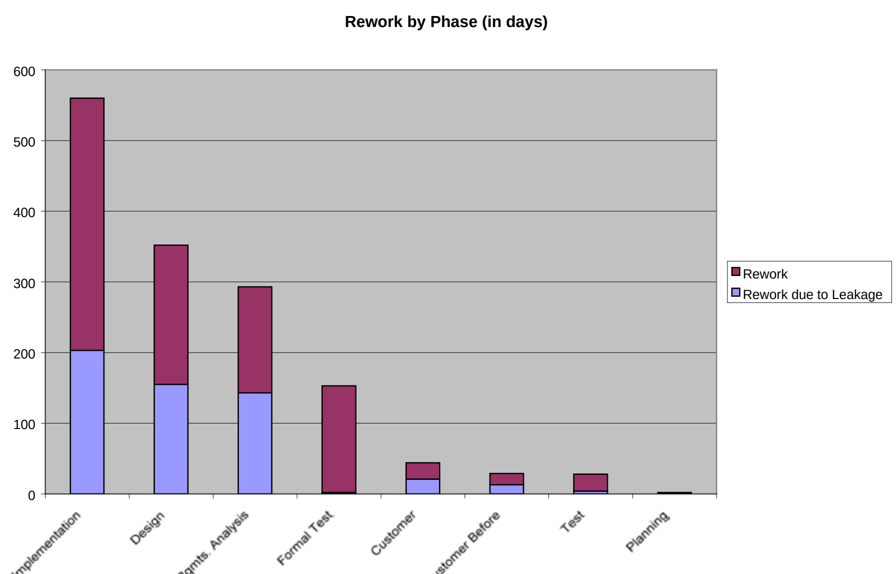
What Was Learned, ….. So Far (Part 2)

Using Pareto Charts We Know Where Defects Enter the Process . . .

And Where Those Defects Are Detected by the Process . . .



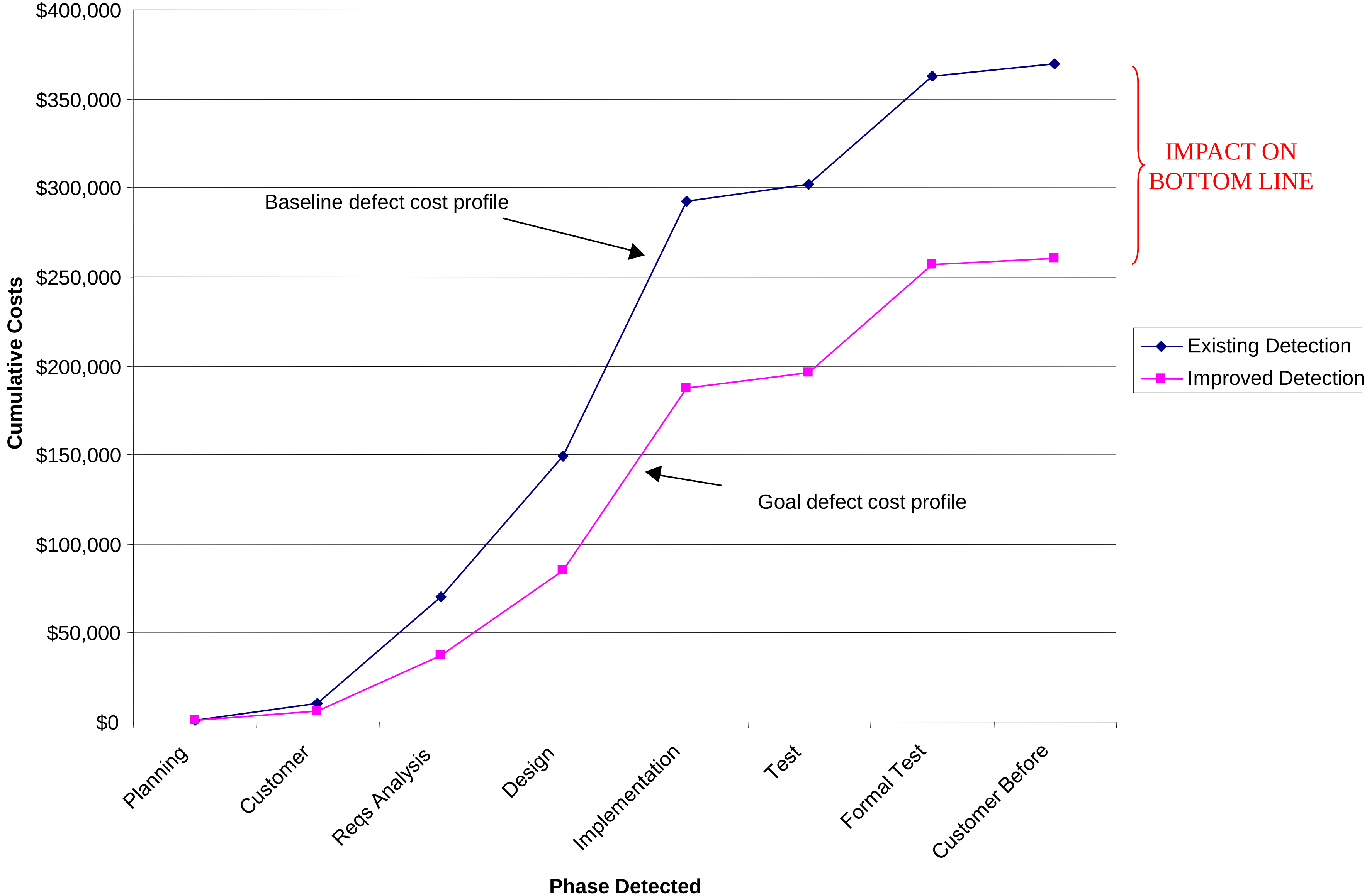
And Even Which Phases Are Hitting the Bottom Line



What We Learned…….

- Formal test introduced > 30% of defects
- Finding 58% of defects when product is done (i.e., testing)
- 3 Phases account for > 92% of rework due to leakage

Improvement Goal

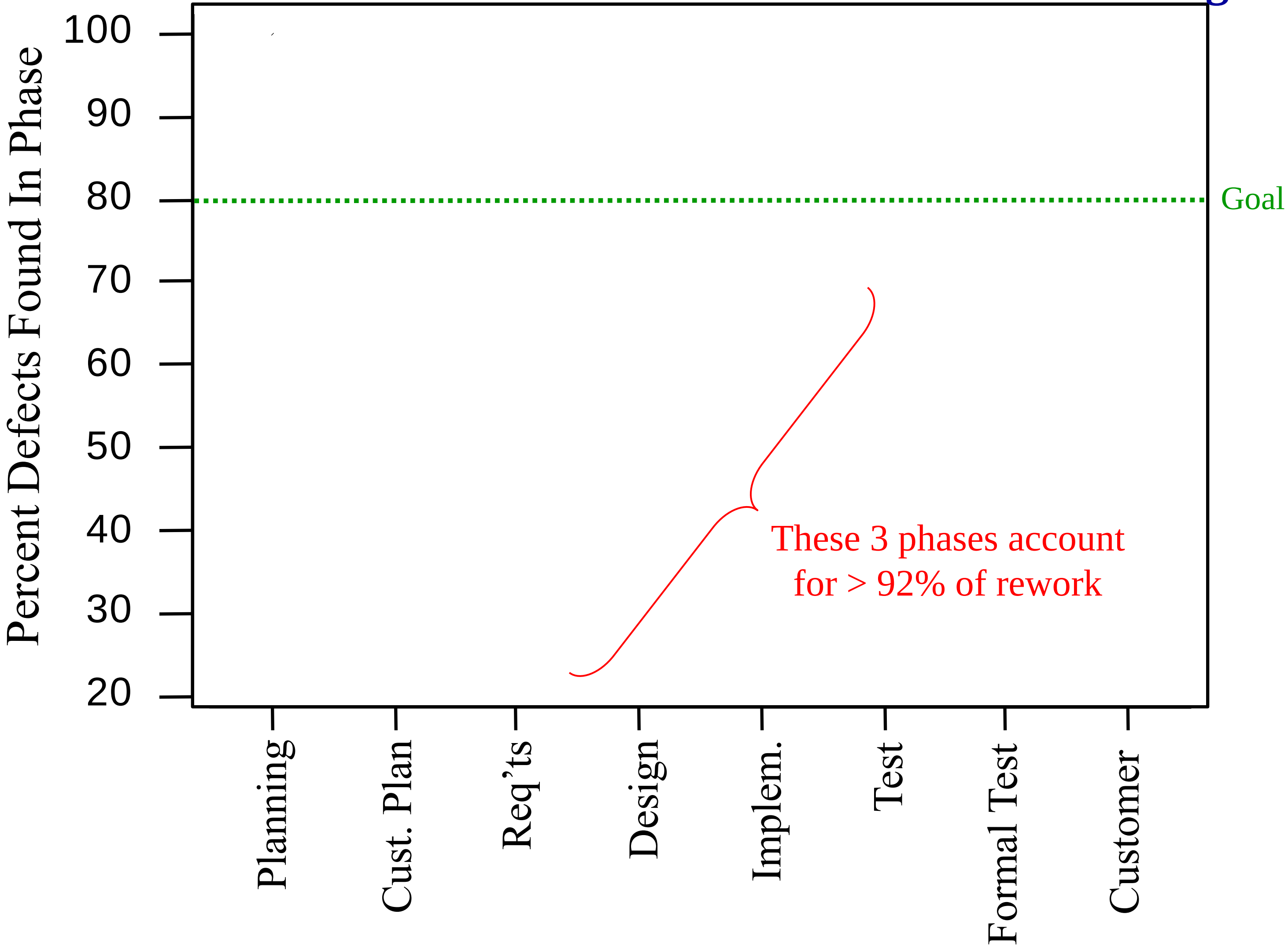


Same number of total defects introduced in the same phases

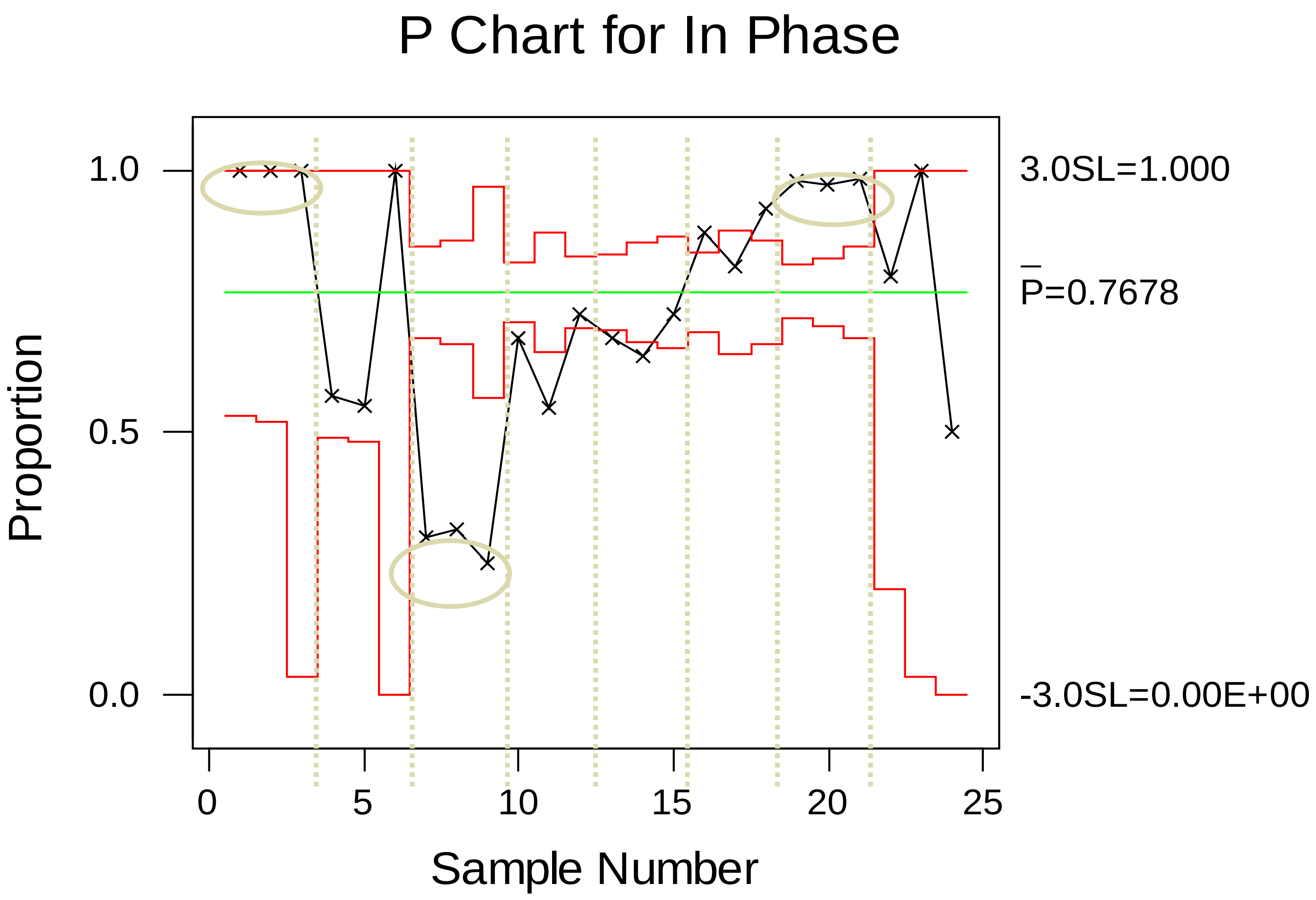
Distributional Characteristics of Data

Stable or Unstable?

Is the variation between or within subgroups?

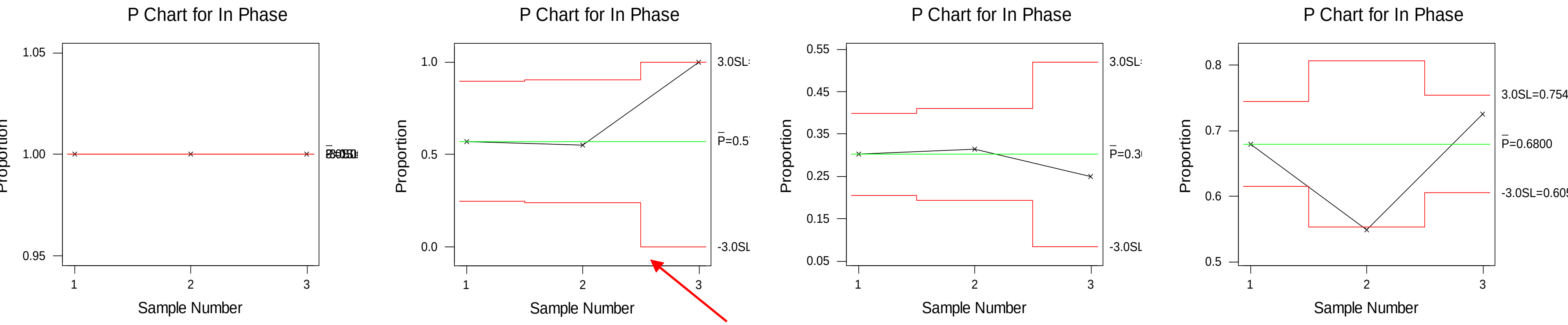


Stability of Entire Process…

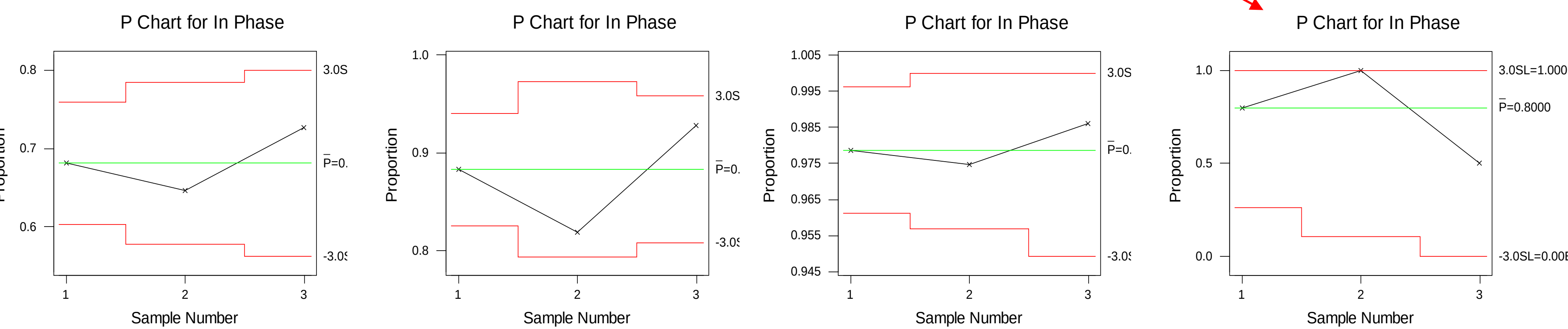


Raytheon

….Then applied to Each Phase



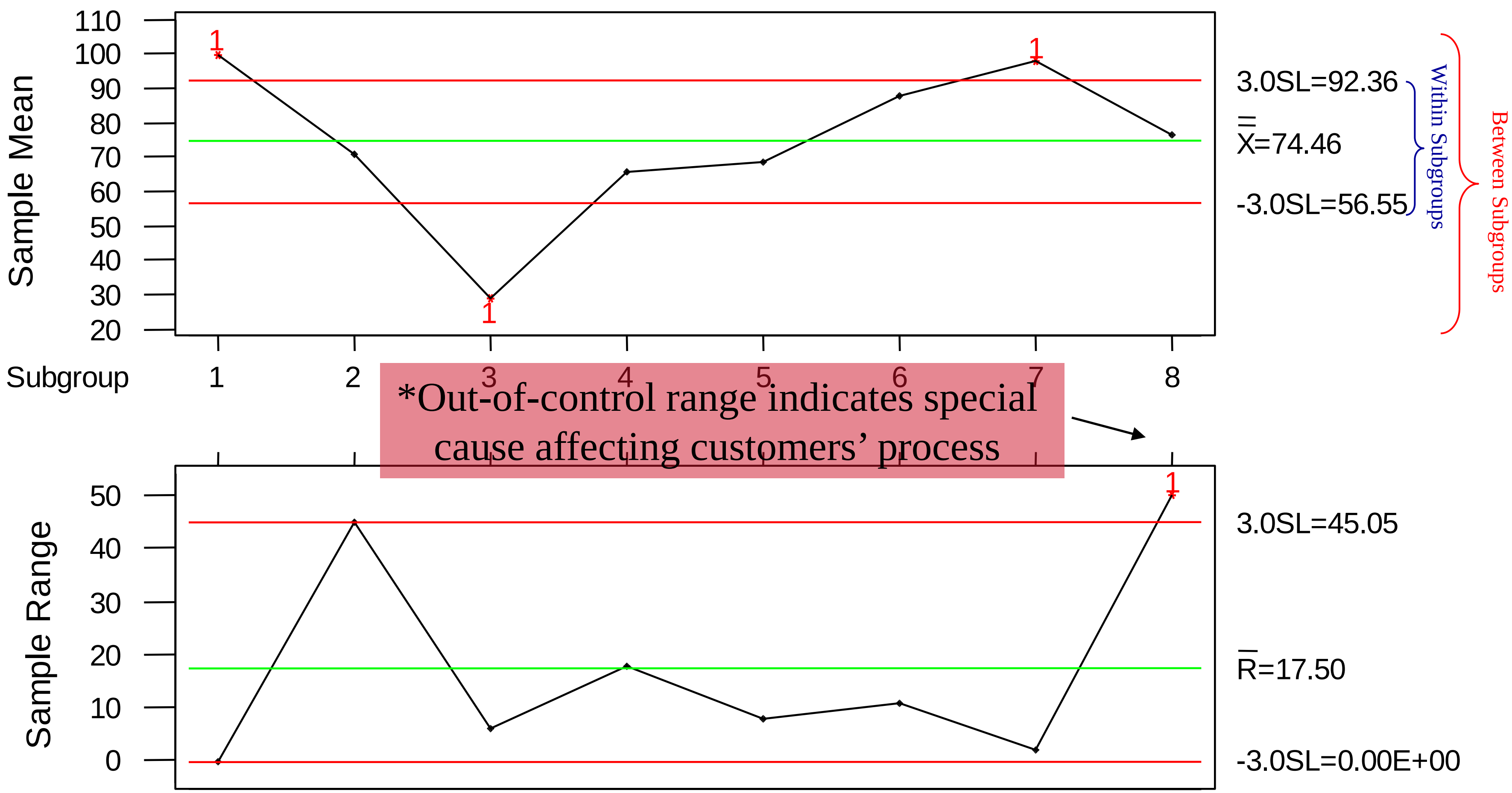
Predictable to be between 0% - 100%
Customer process, team
spun off to work this



Looks pretty stable within subgroups (projects)...

How About Between the Subgroups?

Displays averages “between” the subgroups



Displays range “within” the subgroups

ANOVA Confirms Our Suspicions

Analysis of Variance for percent

Source	DF	SS	MS	F	P
phase	7	10841.9583	1548.8512	8.701	0.000
project	16	2848.0000	178.0000		
Total	23	13689.9583			

Variance Components

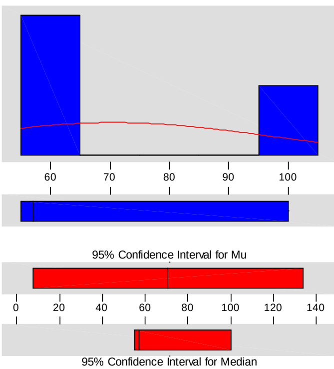
Source	Var Comp.	% of Total	StDev
phase	456.950	**71.97**	21.376
project	178.000	**28.03**	13.342
Total	634.950		25.198

Project variation Phase variation

But we sampled only 3 projects - what does the population look like?

Calculating Confidence Intervals...

Descriptive Statistics



Variable: Cust1

Anderson-Darling Normality Test

A-Squared:	0.446
P-Value:	0.081
Mean	70.6667
StDev	25.4231
Variance	646.333
Skewness	1.72000
Kurtosis	
N	3
Minimum	55.000
1st Quartile	55.000
Median	57.000
3rd Quartile	100.000
Maximum	100.000

95% Confidence Interval for Mu

7.512 133.821

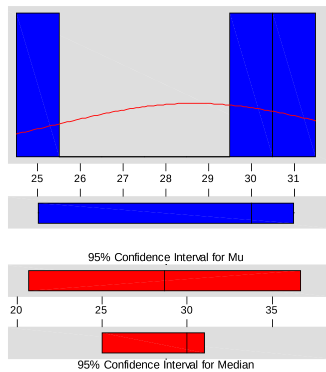
95% Confidence Interval for Sigma

13.237 159.777

95% Confidence Interval for Median

55.000 100.000

Descriptive Statistics



Variable: Req

Anderson-Darling Normality Te

A-Squared:	0.3
P-Value:	0.2
Mean	28.66
StDev	3.21
Variance	10.33
Skewness	-1.545
Kurtosis	
N	
Minimum	25.00
1st Quartile	25.00
Median	30.00
3rd Quartile	31.00
Maximum	31.00

95% Confidence Interval for M

20.6813 36.65

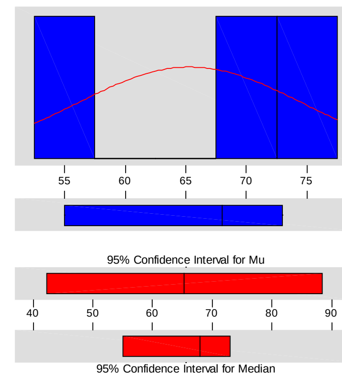
95% Confidence Interval for Sig

1.6737 20.20

95% Confidence Interval for Mec

25.0000 31.00

Descriptive Statistics



Variable: Design

Anderson-Darling Normality Test

A-Squared:	0.259
P-Value:	0.384
Mean	65.3333
StDev	9.2916
Variance	86.3333
Skewness	-1.18512
Kurtosis	
N	3
Minimum	55.0000
1st Quartile	55.0000
Median	68.0000
3rd Quartile	73.0000
Maximum	73.0000

95% Confidence Interval for Mu

42.2518 88.4149

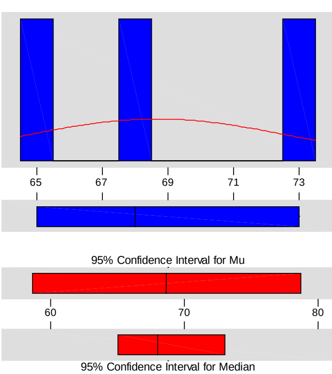
95% Confidence Interval for Sigma

4.8377 58.3951

95% Confidence Interval for Median

55.0000 73.0000

Descriptive Statistics



Variable: Implem

Anderson-Darling Normality Tes

A-Squared:	0.21
P-Value:	0.53
Mean	68.6667
StDev	4.041
Variance	16.3333
Skewness	0.72210
Kurtosis	
N	
Minimum	65.000
1st Quartile	65.000
Median	68.000
3rd Quartile	73.000
Maximum	73.000

95% Confidence Interval for Mu

58.6271 78.706

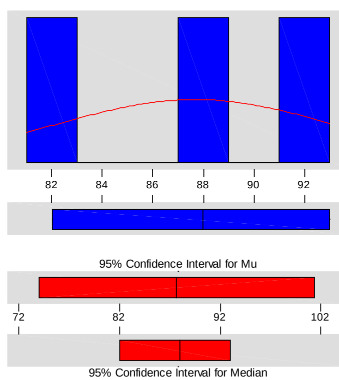
95% Confidence Interval for Sigm

2.1042 25.399

95% Confidence Interval for Medi

65.0000 73.000

Descriptive Statistics



Variable: test

Anderson-Darling Normality Te

A-Squared:	0.1
P-Value:	0.6
Mean	87.66
StDev	5.50
Variance	30.33
Skewness	-2.7E-
Kurtosis	
N	
Minimum	82.0
1st Quartile	82.0
Median	88.0
3rd Quartile	93.0
Maximum	93.0

95% Confidence Interval for M

73.985 101.3

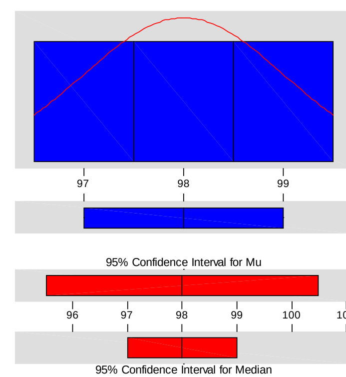
95% Confidence Interval for Sig

2.868 34.6

95% Confidence Interval for Mec

82.000 93.0

Descriptive Statistics



Variable: Form Tst

Anderson-Darling Normality Test

A-Squared:	0.189
P-Value:	0.631
Mean	98
StDev	1
Variance	1
Skewness	0
Kurtosis	
N	3
Minimum	97.000
1st Quartile	97.000
Median	98.000
3rd Quartile	99.000
Maximum	99.000

95% Confidence Interval for Mu

95.516 100.484

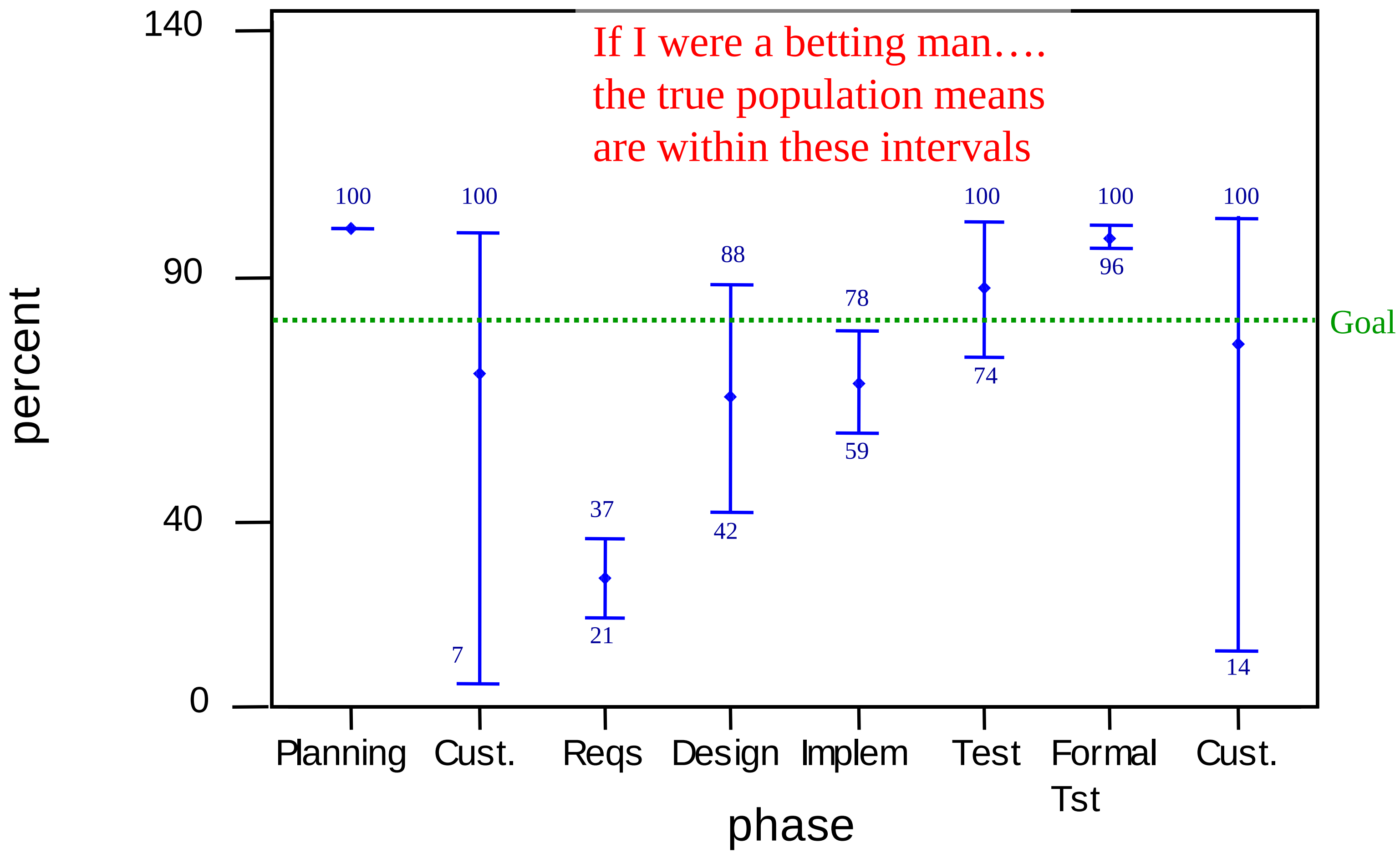
95% Confidence Interval for Sigma

0.521 6.285

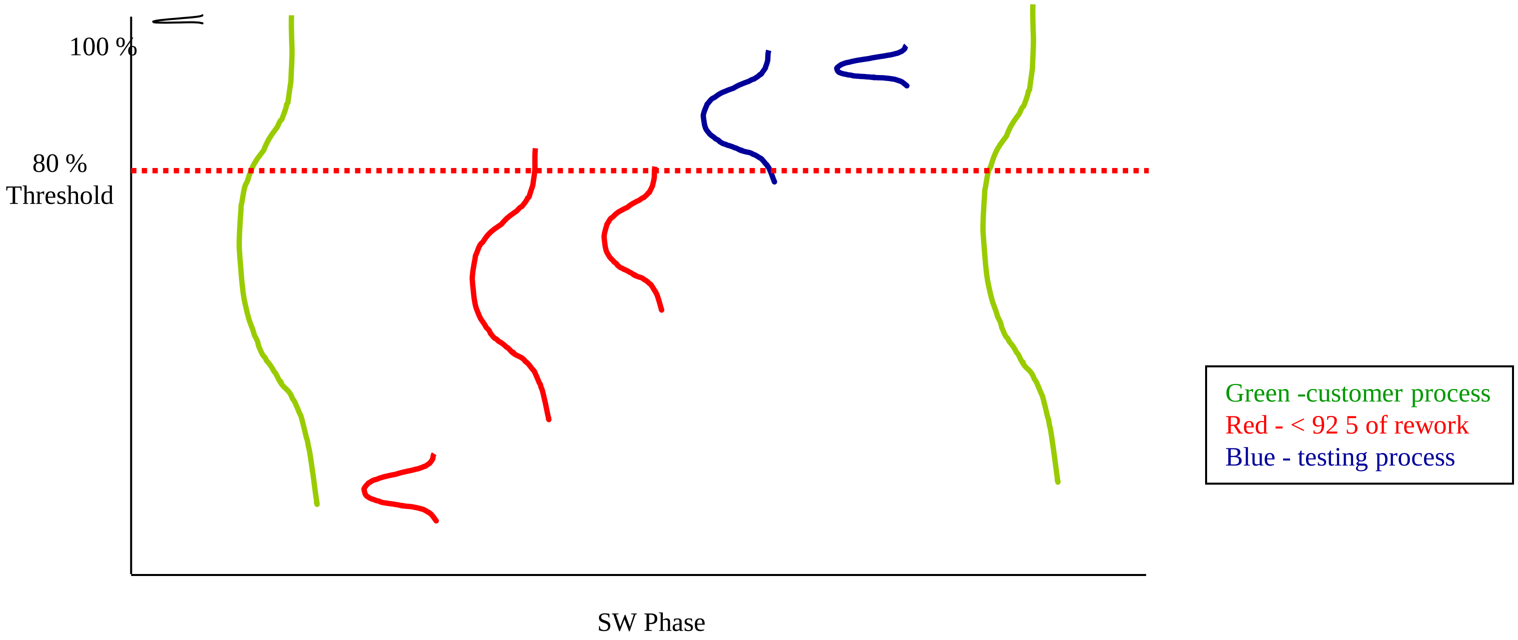
95% Confidence Interval for Median

97.000 99.000

...Gives Us Plausible Population Range



What Was Learned, ….. So Far (Part 3)



- Variation between the phases (72%) is greater than variation between projects (28%)
  - need to work largest source of variation - what changed between, what didn't, etc.
- If no action is taken 95% confident that
  - the Requirements Phase will find between 21% - 37% of defects in phase
  - the Design Phase will find between 42% - 88% of defects in phase
  - the Implementation Phase will find 59% - 78% of defects in phase

Design of Experiments (DOE)

Review Ada packages and C++ objects

Four Factors

experience	(<2 yrs	>2 yrs)
training	(No	Yes)
review criteria	(None	Checklist)
number of reviewers	(2	>2)

Block by Program

language, management style, schedule pressures

Sixteen Runs 2^{5-1} Half fraction

Resolution V Mains compounded w/4ways, 2 w/3ways

Response Variable

percentage of defects which match SEPG and project leads'

Limitations:

"chunks of code" reviewed were different

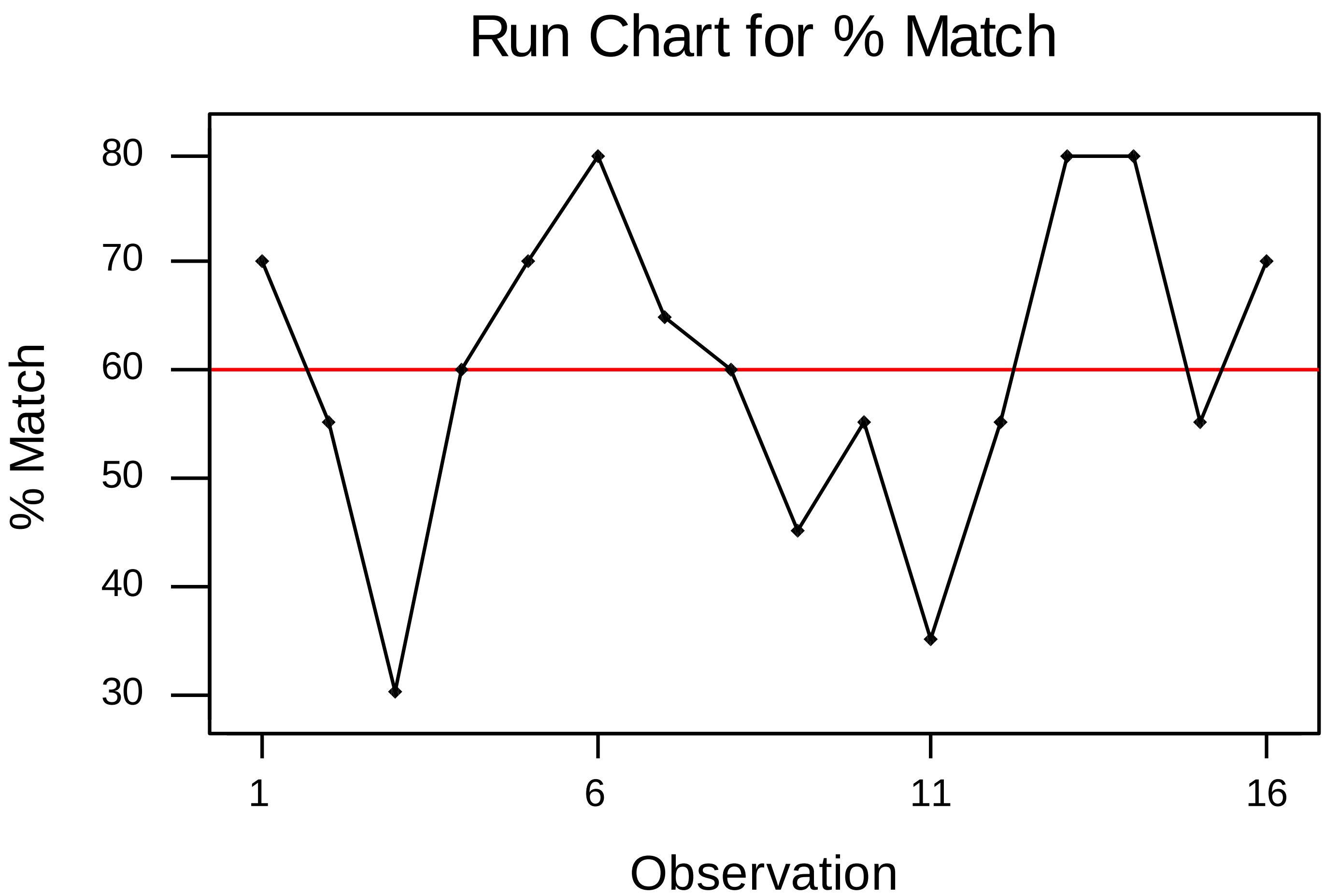
restrictions on randomization

hard to find "team" fulfilling factor levels

DOE Run Results

StdOrder	RunOrder	Program	Experience	Training	Criteria	Num People	% Match
1	1	-1	1	-1	-1	-1	70
5	2	-1	-1	-1	-1	1	55
3	3	-1	-1	-1	1	-1	30
7	4	-1	1	-1	1	1	60
2	5	-1	-1	1	-1	-1	70
6	6	-1	1	1	-1	1	80
4	7	-1	1	1	1	-1	65
8	8	-1	-1	1	1	1	60
9	9	1	-1	-1	-1	-1	45
13	10	1	1	-1	-1	1	55
15	11	1	-1	-1	1	1	35
11	12	1	1	-1	1	-1	55
14	13	1	-1	1	-1	1	80
10	14	1	1	1	-1	-1	80
12	15	1	-1	1	1	-1	55
16	16	1	1	1	1	1	70

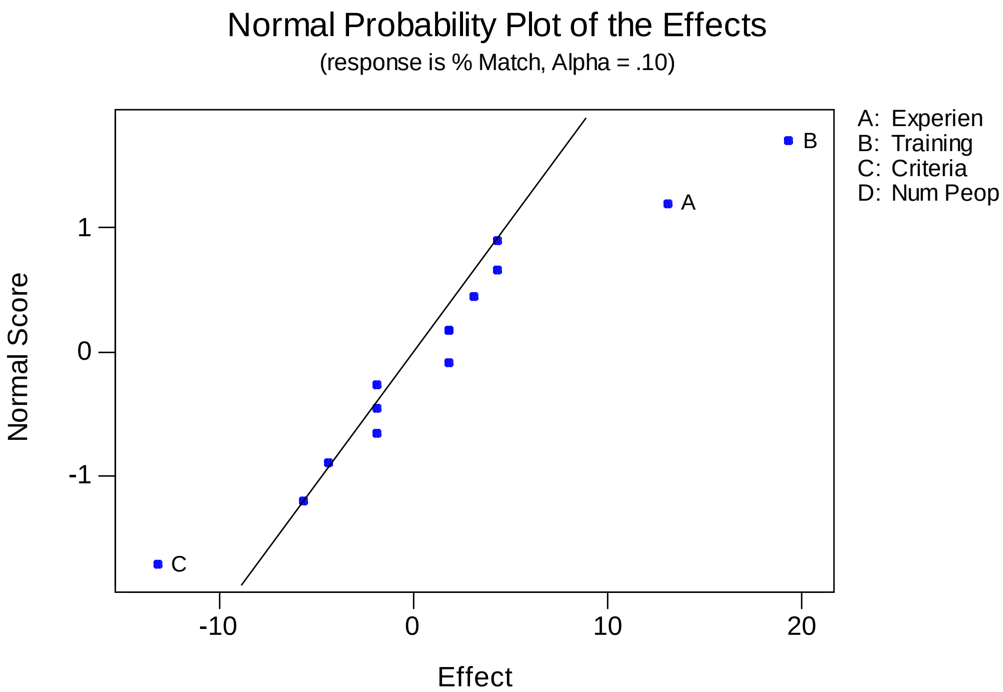
DOE Run Chart



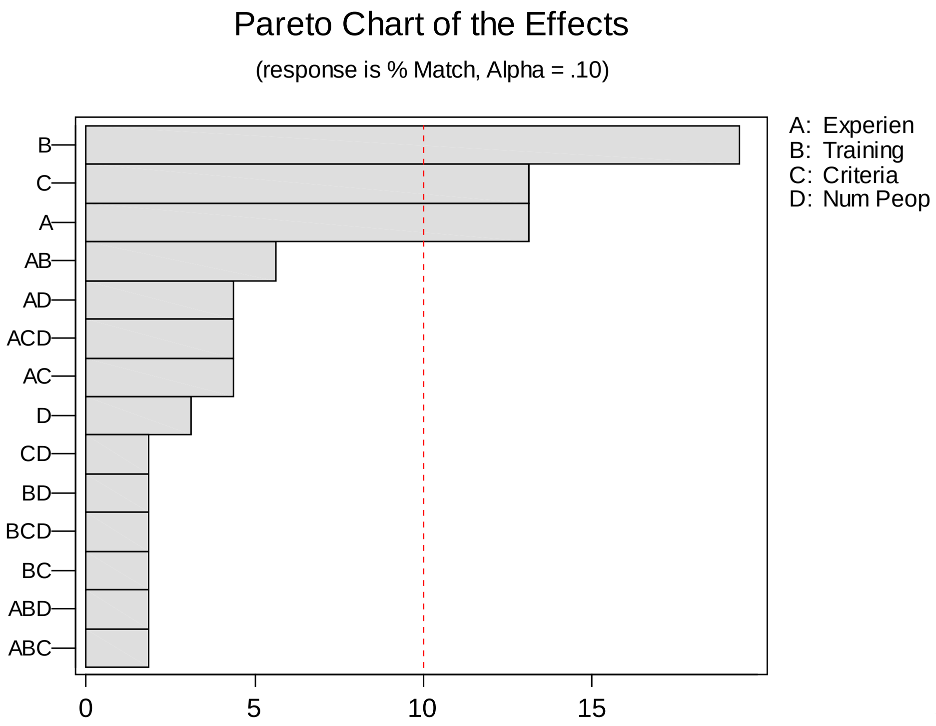
Sorting by Response

StdOrder	RunOrder	Program	Experience	Training	Criteria	Num People	% Match
6	6	-1	1	1	-1	1	80
10	14	1	1	1	-1	-1	80
14	13	1	-1	1	-1	1	80
1	1	-1	1	-1	-1	-1	70
2	5	-1	-1	1	-1	-1	70
16	16	1	1	1	1	1	70
4	7	-1	1	1	1	-1	65
7	4	-1	1	-1	1	1	60
8	8	-1	-1	1	1	1	60
5	2	-1	-1	-1	-1	1	55
13	10	1	1	-1	-1	1	55
11	12	1	1	-1	1	-1	55
12	15	1	-1	1	1	-1	55
9	9	1	-1	-1	-1	-1	45
15	11	1	-1	-1	1	1	35
3	3	-1	-1	-1	1	-1	30

Might have something here



Data looks pretty normal



Shows Training, Criteria, Experience as the influential factors

Fractional Factorial Fit

Estimated Effects and Coefficients for % (coded units)

Term	Effect	Coef
Constant		60.312
Program	-1.875	-0.938
Experien	13.125	6.562
Training	19.375	9.687
Criteria	-13.125	-6.562
Num Peop	3.125	1.562
Program*Experien	-1.875	-0.937
Program*Training	4.375	2.187
Program*Criteria	1.875	0.937
Program*Num Peop	-1.875	-0.937
Experien*Training	-5.625	-2.812
Experien*Criteria	4.375	2.188
Experien*Num Peop	-4.375	-2.187
Training*Criteria	-1.875	-0.938
Training*Num Peop	1.875	0.937
Criteria*Num Peop	1.875	0.937

Shows same thing
Training, Criteria, Experience
as the influential factors

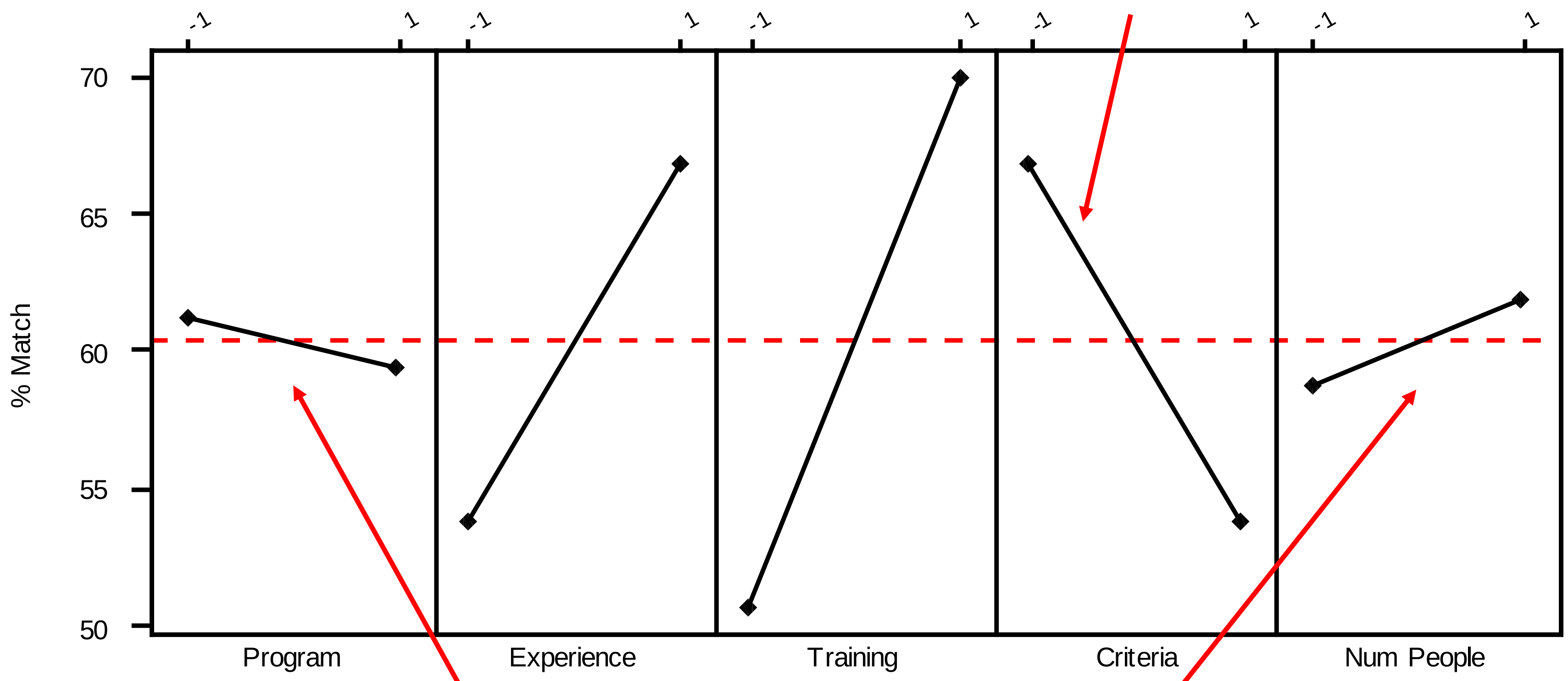
Analysis of Variance for % (coded units)

Source	DF	Seq SS	Adj SS	Adj MS	F
Main Effects	5	2932.8	2932.8	586.56	*
2-Way Interactions	10	440.6	440.6	44.06	*
Residual Error	0	0.0	0.0	0.00	
Total	15	3373.4			

Main Effects Plot (data means) for % Match

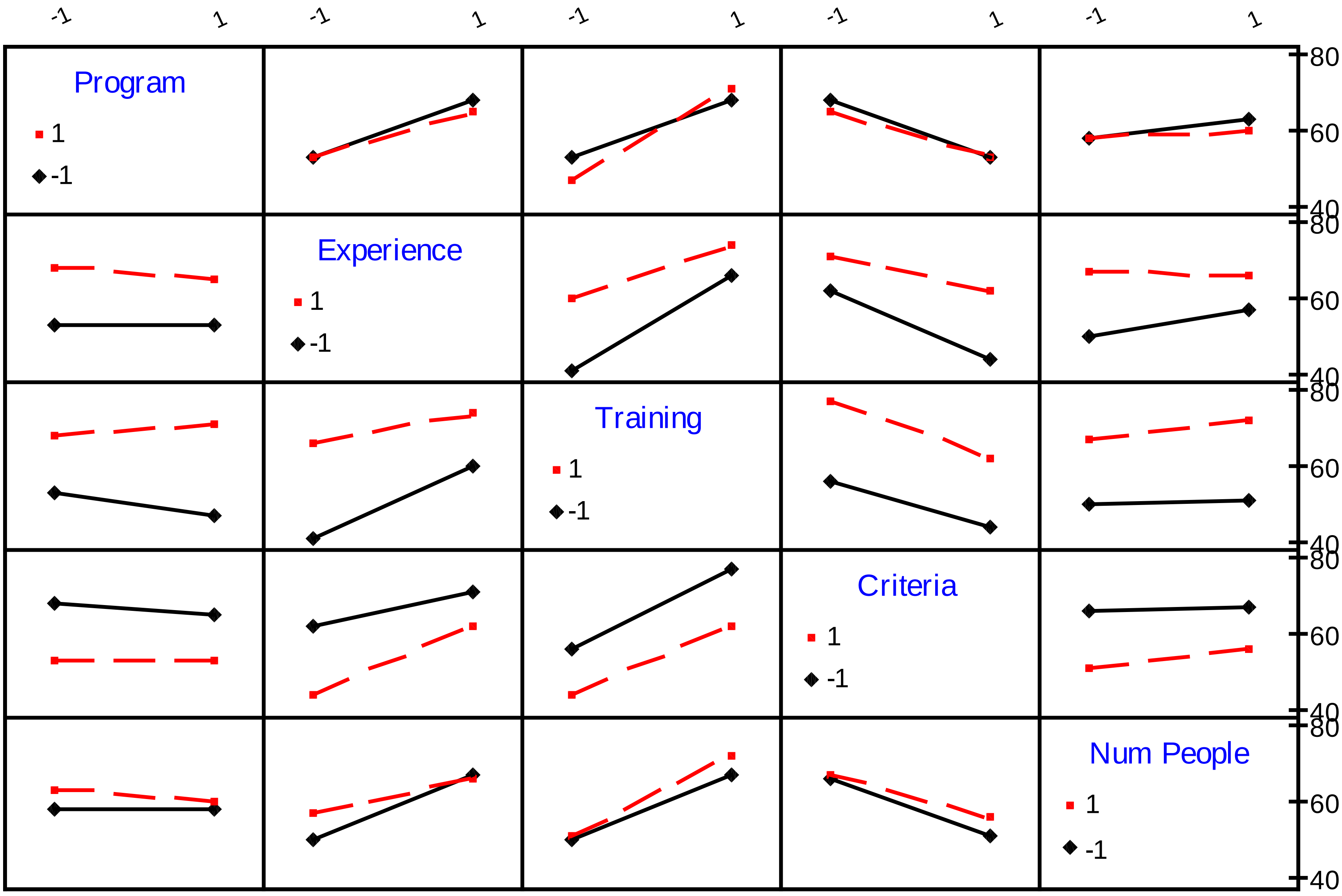
Same data but not we know whether to set high or low

Whoa, what do we have here?



Process pretty robust to program, number of reviewers

Interaction Plot (data means) for % Match



What Was Learned, ….. So Far (Part 4)

Experience was no brainier - more is better

Training was no brainier - have to have it!
  Phase dependent training, not just peer review training

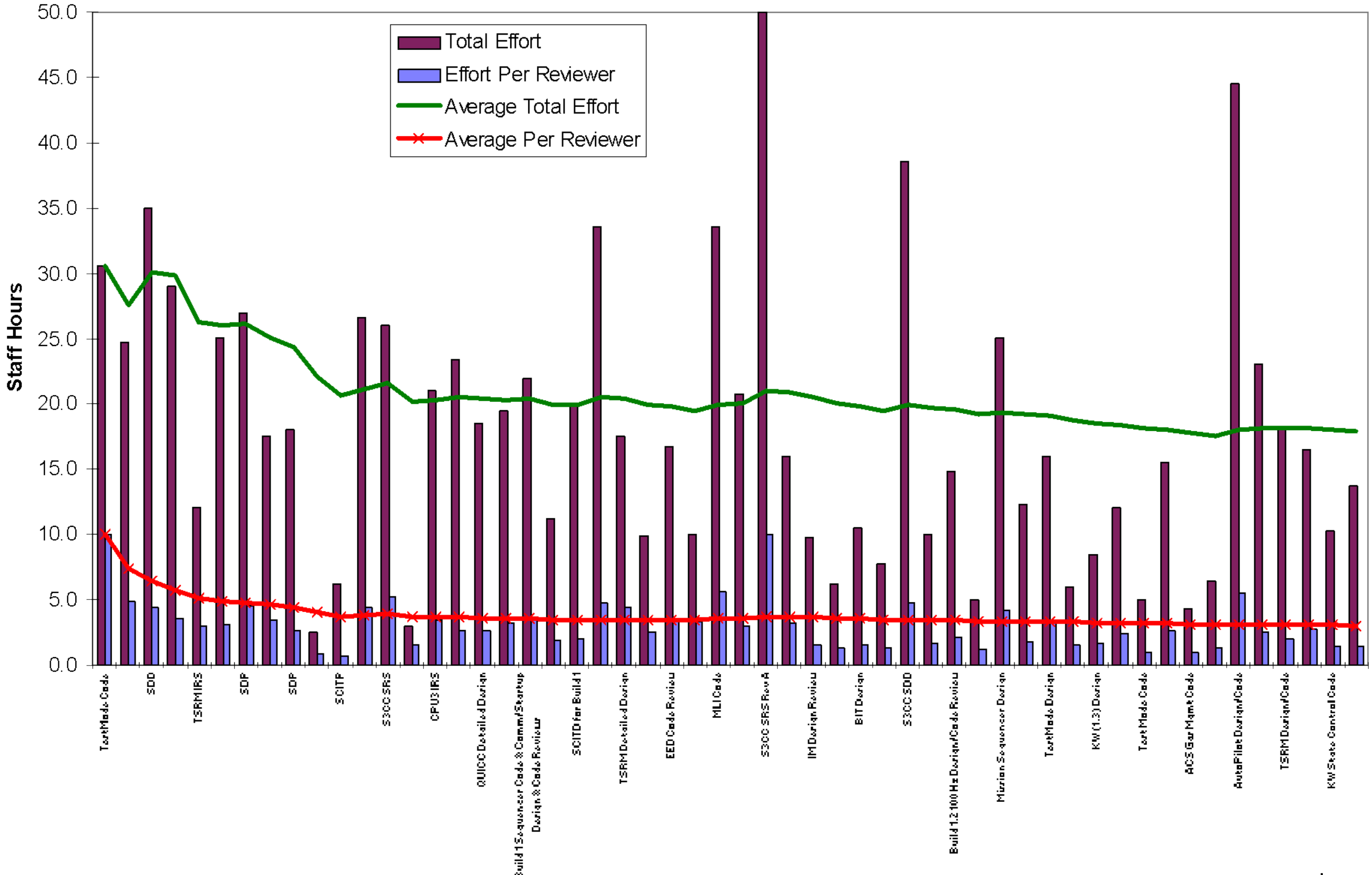
Number of people was eye opening - didn't make too much difference
  Since more people cost more money, keep it at 2

Criteria was a shock

Need to do follow-up to see why this was counter-intuitive

Follow up revealed criteria limited the scope of the review, reviewed only what was there, did not use as intended

Other Measures Were Monitored, But Not Part of DOE



Tool Usage

Thought Process Map - big benefit when reviewing project with others. Helps avoid heading down a dead end path before you even start. Makes you ask questions and identify barriers well before you actually get there. Much mor … itical than first thought. Is a pain to keep up to date

Process Map - … tually operating, not how it is designed to operate. Important to get a varying … iews. Key to almost everything, need to know what x's are involved i …

FMEA - Id … that should be taken. Importan …

S/W Worksheet (Product Scorecard) - Collec … cycle. Facilitates to baseline data and track improvement. … you need to show improvement

Pareto Chart - Bar chart ordered from largest to smallest. Helped immediately determine which … to focus improvements upon for the largest payback and where to ask initial questions.

Control Charts - Distinguishes special and common variation in the process. Helps to develop appropriate action for the type of variation. Showed how process would perform if nothing done to improve it.

DOE - Determined what factors were influential and were to set those factors. Showed that the obvious accepted conclusions are not always an improvement.

Bottom line -
Until the DOE all the tools were useful and we could see their value. They confirmed what was already known. We were able to move forward confident that we were working the right issues and had the metrics to back up my assumptions.
After the DOE a valuable lesson was learned. We do not know everything. If we hadn't used the data to run the DOE, we would have made things worse!!

Results

Raytheon

	Planning	Customer	Rqmts. Analysis	Design	Implementation	Test	Formal Test	Customer Before	TOTAL
Planning	29	0	0	0	0	0	0	0	29
Customer	0	12	2	0	2	0	5	0	21
Rqmts. Analysis	0	0	61	14	29	26	71	1	202
Design	0	0	1	323	82	29	38	2	475
Implementation	0	0	1	5	220	43	44	10	323
Test	0	0	0	2	1	249	30	0	282
Formal Test	0	0	0	0	0	13	597	0	610
Customer Before	0	0	0	0	0	0	1	4	5
TOTAL	29	12	65	344	334	360	786	17	1947

Number of defects identified by phase introduced/phase detected (from Product Scorecard)

AFTER IMPROVEMENTS

PHASE DETECTED

PHASE INTRODUCED	Planning	Customer	Rqmts. Analysis	Design	Implementation	Test	Formal Test	Customer Before	TOTAL
Planning	235	2	0	3	0	0	0	0	240
Customer	0	22	22	7	4	0	0	0	55
Rqmts. Analysis	0	1	161	13	7	2	0	0	184
Design	0	0	3	173	111	2	0	0	289
Implementation	0	0	0	2	342	20	0	1	365
Test	0	0	0	2	2	22	1	0	27
Formal Test	0	0	0	0	0	0	2	1	3
Customer Before	0	0	0	0	0	0	0	3	3
TOTAL	235	25	186	200	466	46	3	5	1166

Raytheon

Percentage of defects identified by phase introduced/phase detected (from Product Scorecard)

	Planning	Customer	Rqmts. Analysis	Design	Implementation	Test	Formal Test	Customer Before	TOTAL
Planning	100%	0%	0%	0%	0%	0%	0%	0%	1%
Customer	0%	57%	10%	0%	10%	0%	24%	0%	1%
Rqmts. Analysis	0%	0%	30%	7%	14%	13%	35%	0%	10%
Design	0%	0%	0%	68%	17%	6%	8%	0%	24%
Implementation	0%	0%	0%	2%	68%	13%	14%	3%	17%
Test	0%	0%	0%	1%	0%	88%	11%	0%	14%
Formal Test	0%	0%	0%	0%	0%	2%	98%	0%	31%
Customer Before	0%	0%	0%	0%	0%	0%	20%	80%	0%
TOTAL	1%	1%	3%	18%	17%	18%	40%	1%	100%

AFTER IMPROVEMENTS

PHASE DETECTED

PHASE INTRODUCED		Planning	Customer	Rqmts. Analysis	Design	Implementation	Test	Formal Test	Customer Before	TOTAL
I	Planning	98%	1%	0%	1%	0%	0%	0%	0%	21%
N	Customer	0%	40%	40%	13%	7%	0%	0%	0%	5%
T	Rqmts. Analysis	0%	1%	88%	7%	4%	1%	0%	0%	16%
R	Design	0%	0%	1%	60%	38%	1%	0%	0%	25%
O	Implementation	0%	0%	0%	1%	94%	5%	0%	0%	31%
D	Test	0%	0%	0%	7%	7%	81%	4%	0%	2%
U	Formal Test	0%	0%	0%	0%	0%	0%	67%	33%	0%
C	Customer Before	0%	0%	0%	0%	0%	0%	0%	100%	0%
D	TOTAL	20%	2%	16%	17%	40%	4%	0%	0%	100%

Raytheon

Cost of Rework due to Defects in Days (from Product Scorecard)

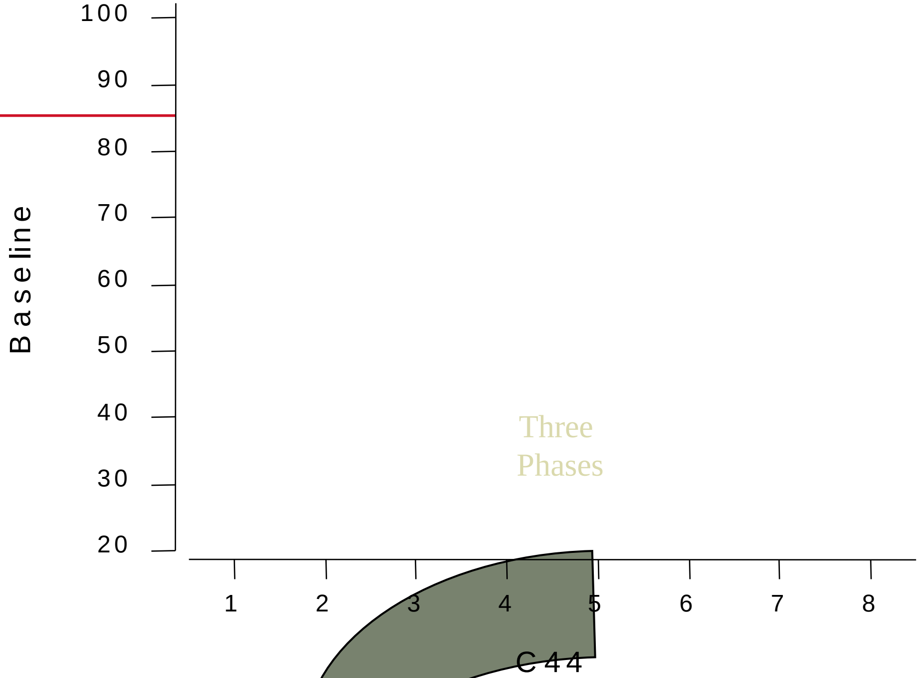
	Planning	Customer	Rqmts. Analysis	Design	Implementation	Test	Formal Test	Customer Before	TOTAL	Leaked
Planning	2.03	0	0	0	0	0	0	0	2.03	0
Customer	0	1.8	1.4	0	4.06	0	16.15	0	23.41	21.61
Rqmts. Analysis	0	0	7.32	12.04	32.77	41.6	56.09	0.79	150.61	143.29
Design	0	0	0.13	41.99	8.2	23.2	118.94	5.28	197.74	155.75
Implementation	0	0	0.17	0.5	154	90.3	88.88	23.3	357.15	203.15
Test	0	0	0	0.16	0.03	19.92	4.5	0	24.61	4.69
Formal Test	0	0	0	0	0	2.34	149.25	0	151.59	2.34
Customer Before	0	0	0	0	0	0	2.7	13.6	16.3	13.6
TOTAL	2.03	1.8	9.02	54.69	199.06	177.36	436.51	42.97	923.44	544.43

AFTER IMPROVEMENTS

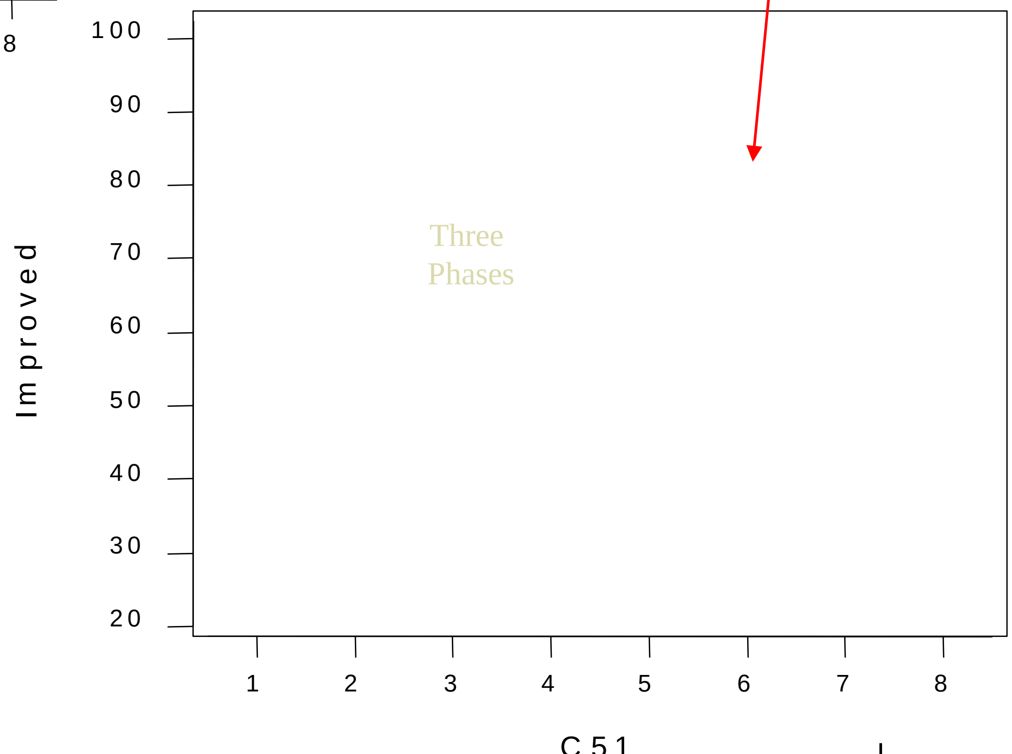
PHASE DETECTED

PHASE		Planning	Customer	Rqmts. Analysis	Design	Implementation	Test	Formal Test	Customer Before	TOTAL	Leaked
I	Planning	16.45	2	0	0.51	0	0	0	0	18.96	2.51
N	Customer	0	3.3	15.4	7	8.12	0	0	0	33.82	30.52
T	Rqmts. Analysis	0	1	19.32	11.18	7.91	3.2	0	0	42.61	23.29
R	Design	0	0	0.39	22.49	11.1	1.6	0	0	35.58	13.09
O	Implementation	0	0	0	0.2	239.4	42	0	2.33	283.93	44.53
D	Test	0	0	0	0.16	0.06	1.76	0.15	0	2.13	0.37
U	Formal Test	0	0	0	0	0	0	0.5	0.58	1.08	0.58
C	Customer Before	0	0	0	0	0	0	0	10.2	10.2	10.2
D	TOTAL	16.45	6.3	35.11	41.54	266.59	48.56	0.65	13.11	428.31	125.09

95% Confident That Two of The Three Phase Will be World-Class. The Third Had Drastic Improvement

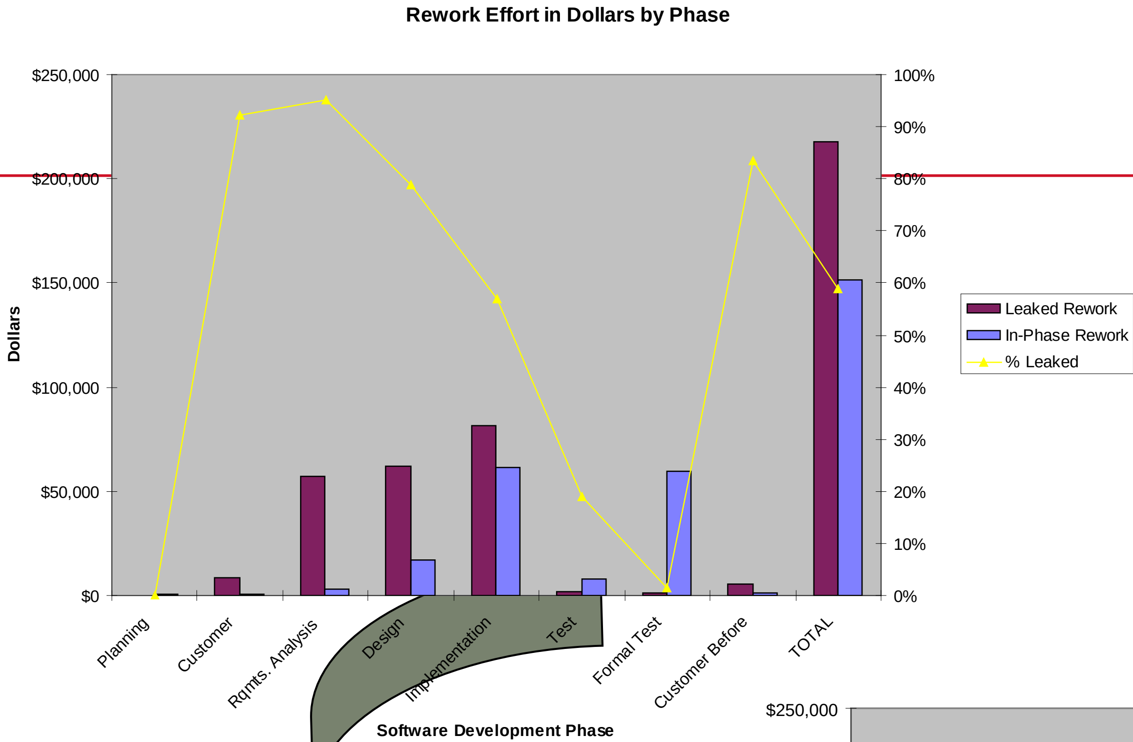


What happened here? Remember we are measuring % not #. We went from 282 to 27, a huge improvement in rework $ and effort, but % wise it was a decline

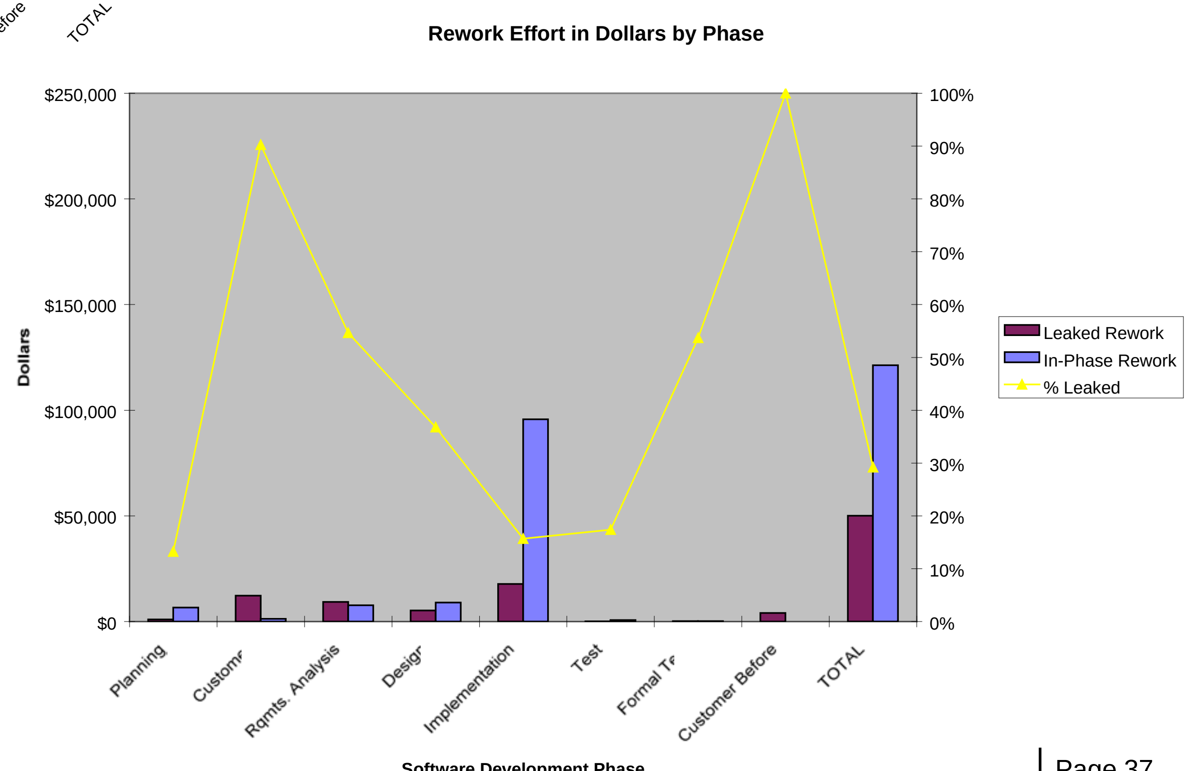


AFTER IMPROVEMENTS

Cost of Rework due to Defects

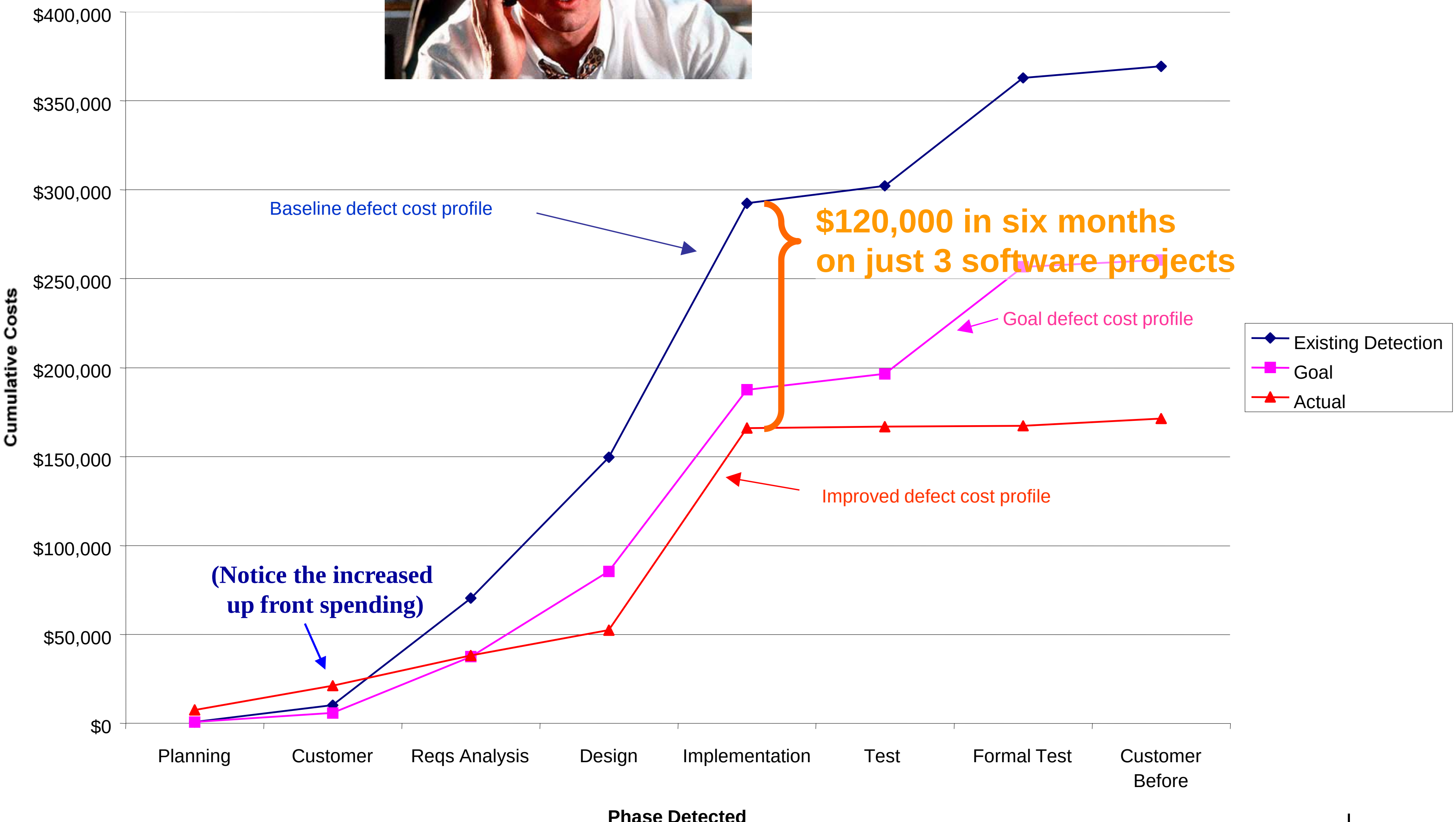


AFTER IMPROVEMENTS

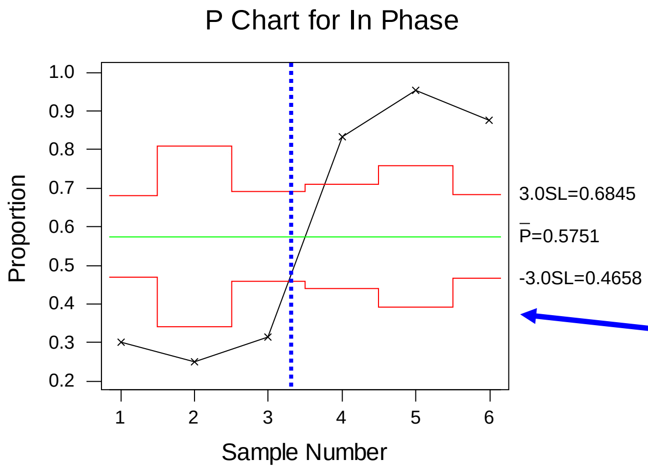


Bottom Line Savings vs Goal - Show Me the Money

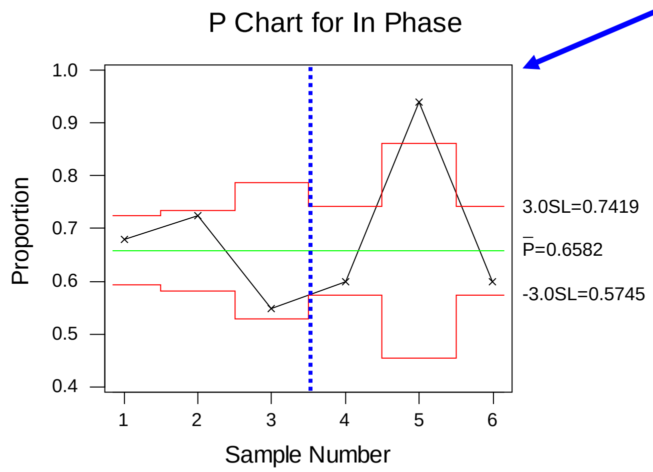
Raytheon

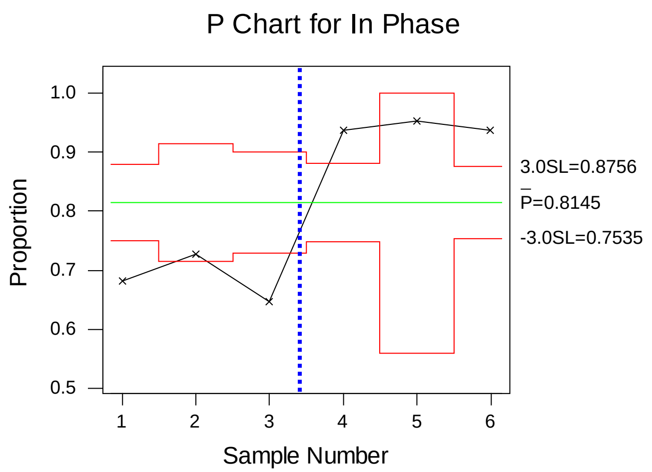


Three Phase Where Project Concentrated (Three Phases With > 92% of Rework)

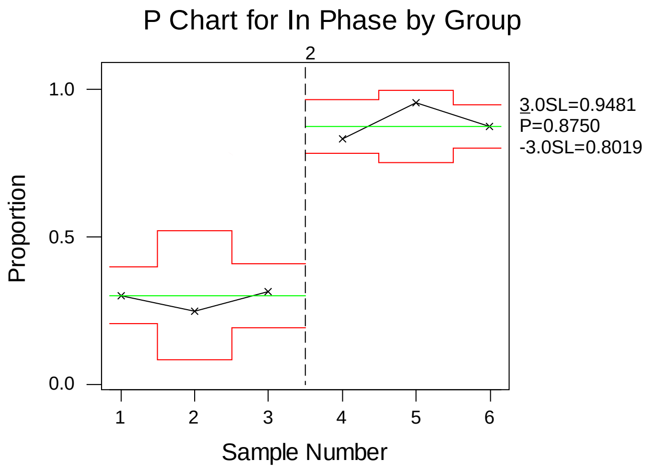


AFTER IMPROVEMENTS



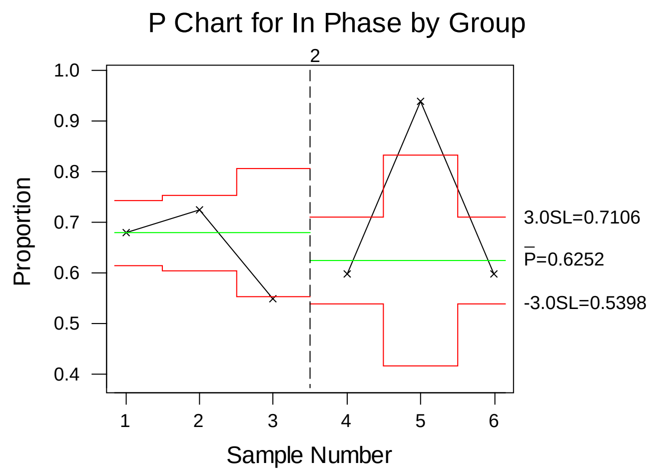


Recalculating the control limits to see if there is significance….



Control Limits for Requirements Analysis Phase has no Overlap Whatsoever

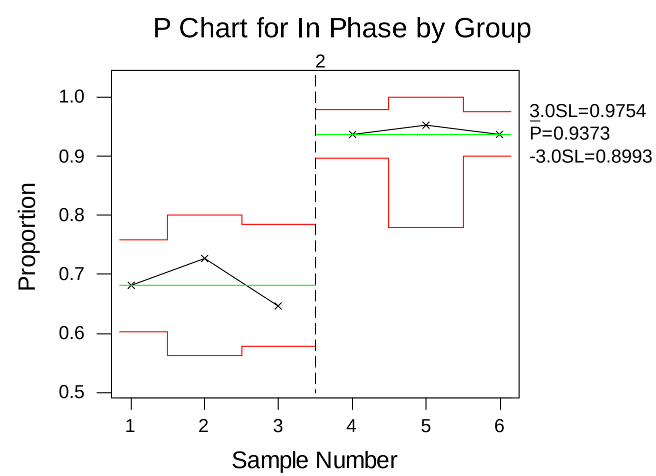
Design Phase is More Interesting



Range is much larger which widens the control limits.

However, looking at the scorecard,
- only 1% of defects made it to test, whereas before it was 14%.
- Rework due to leakage was a mere 13 days compared to 156 days!

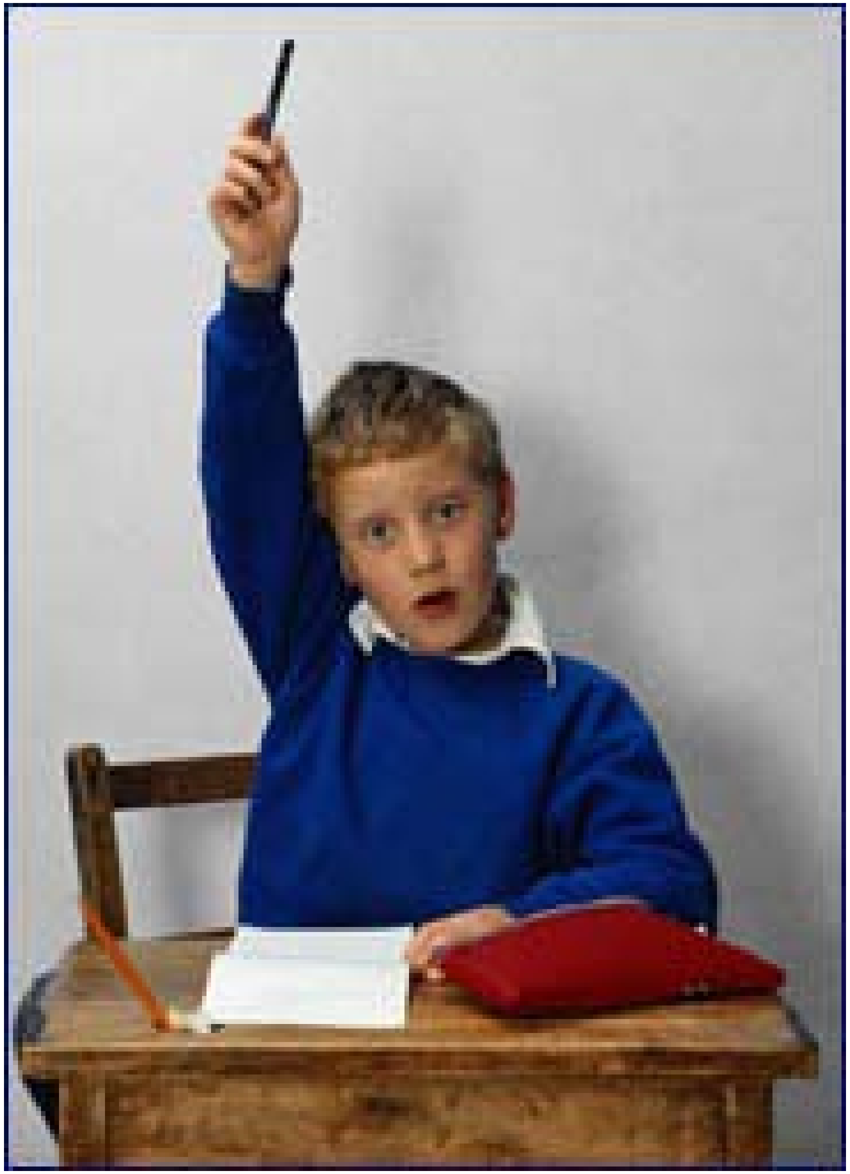
Control Limits for Implementation Phase Have Significantly Changed With Little Overlap



Conclusion:

- Make decisions based on data.
  - Experience is only one input parameter and can steer you wrong
- The Six Sigma tools can and do apply to software
- A few simple process changes resulted had BIG impact to bottom line!
- The three phases that were concentrated upon improved greatly
  - Two of the three are above the threshold
- Quick look at charts show that the process is now in control
- Just measuring the percentage of defect leakage is not the whole story
  - Test injected defects were significantly reduced which resulted in a major cost savings

Questions



That's All Folks

Tom Lienhard

Thomas_G_Lienhard@Raytheon.com

(520) 794-2989