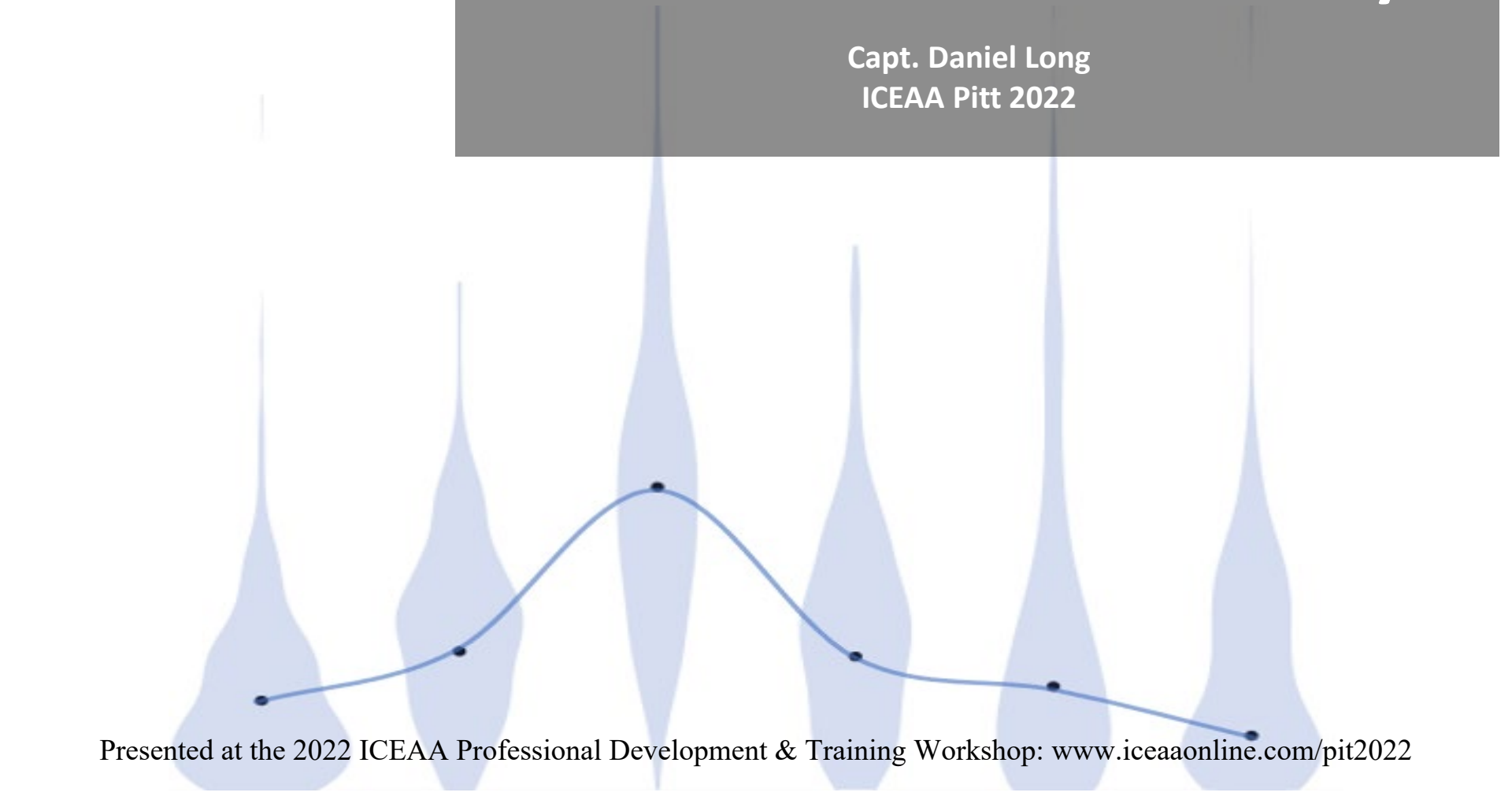
Software Development Phasing & Schedule Growth Analysis

Capt. Daniel Long

ICEAA Pitt 2022



Presented at the 2022 ICEAA Professional Development & Training Workshop: www.iceaaonline.com/pit2022

SW Development in Government can improve

	Successful	Challenged	Failed
Retail	35%	49%	16%
Banking	30%	55%	15%
Financial	29%	56%	15%
Healthcare	29%	53%	18%
Services	29%	52%	19%
Other	29%	48%	23%
Manufacturing	28%	53%	19%
Telecom	24%	53%	23%
Government	21%	55%	24%

Outcome of Software Projects by Industry - Standish Group (2015)

Purpose

Improve software development planning and elucidate phasing interactions & effects within DoD SRDR database

Focus Areas:

- Phasing Distributions (Rules of Thumb)
- Effort vs. Schedule

Gaps in Existing Research:

- No phase analysis related to DoD systems in past 20 years
- No Rules of Thumb verified for current DoD data
  - ACFAA Handbook recommends 40/20/40 split, but not empirically based

Research Questions

1. What phase characteristics do DoD software projects show?
2. How accurate are we with phase planning (do we get our estimates right)?
3. How does effort allocation differ between programs with low schedule growth and high schedule growth projects?
4. How do errors in effort in early phases relate to errors in subsequent phases?
5. How do the results (above) differ by program types and categorical parameters (Waterfall vs Agile, Service, Domain, etc.)?

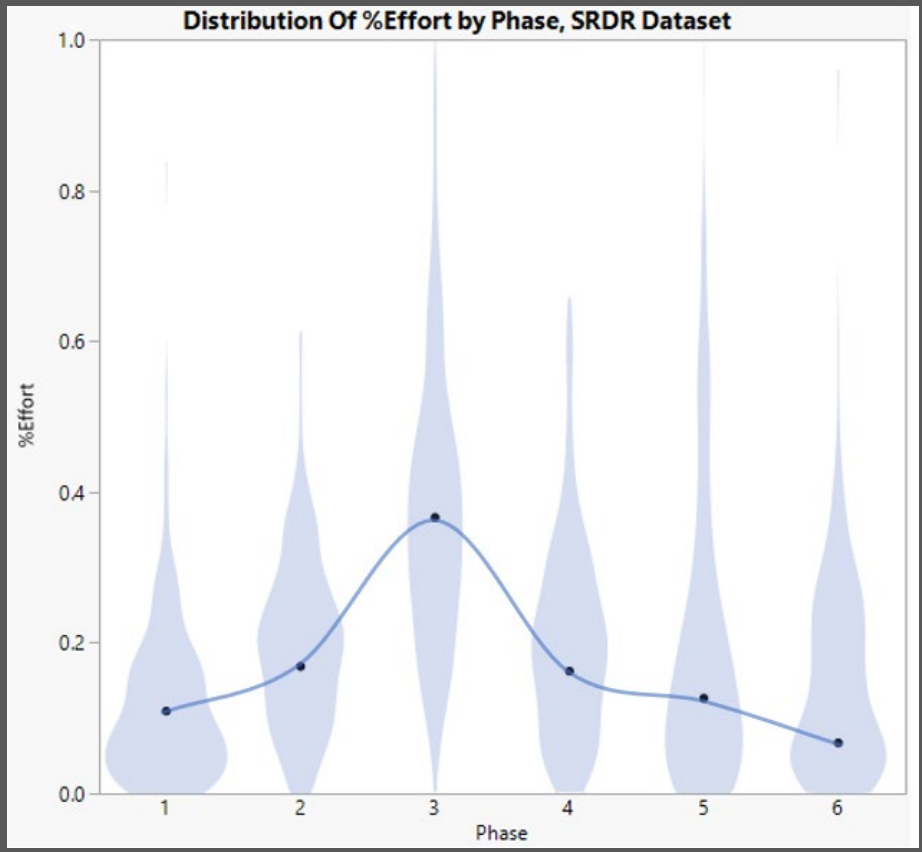
DoD Focus

Last decade of research is on variables that predict program costs, *not* on relationships between SW phases. Older SW research focused on calibrating commercial estimating SW for DoD use (e.g. PRICE, SEER-SEM).

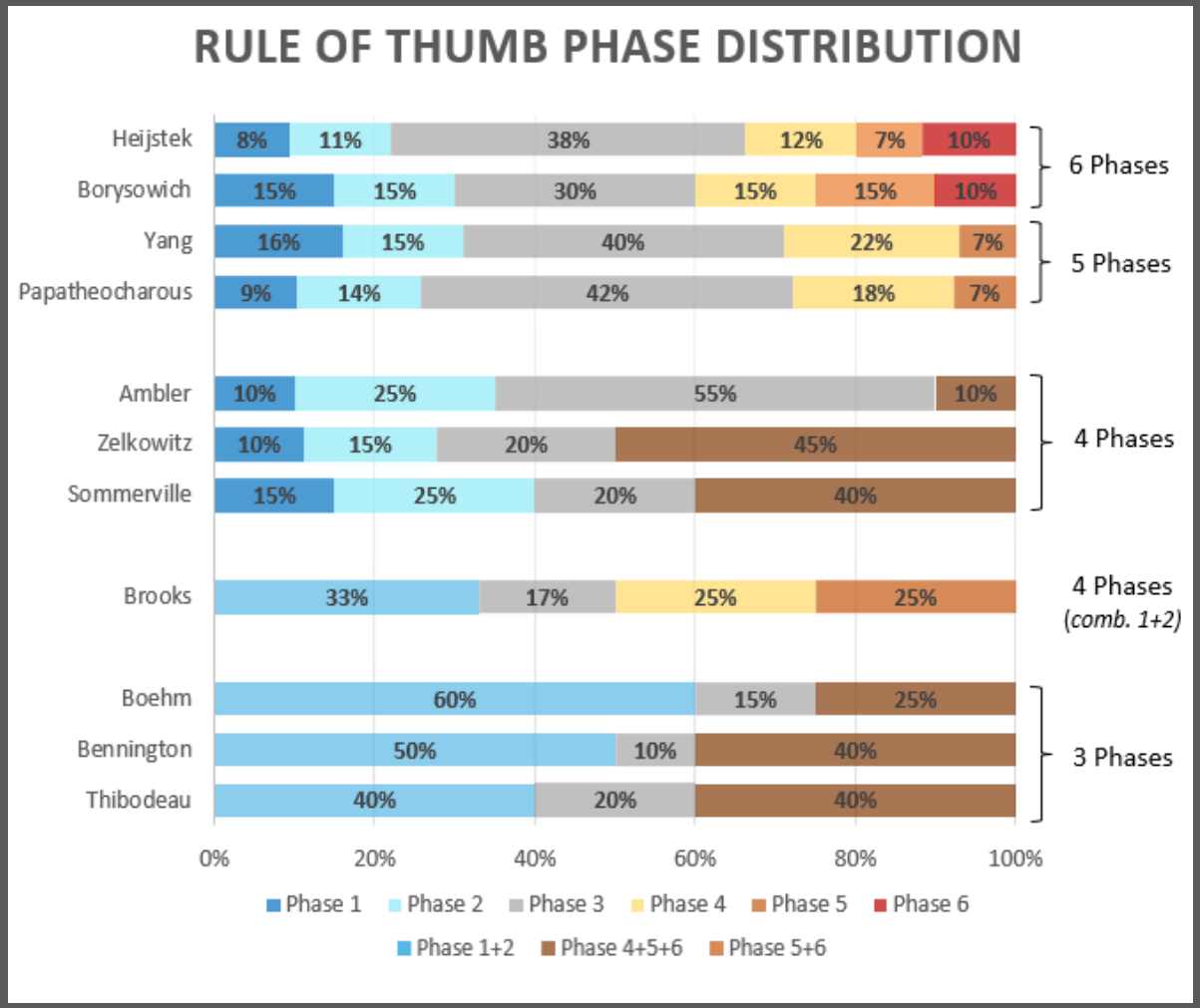
Focus	DoD Researchers
CER Estimating Software	Blalock (’88); Daly (‘90); Coggins (‘93)
Stage Analysis (general Acq, not SW-focused)	Unger (’01); T. Brown (‘02); G. Brown (‘15); D’Amico (‘17)
CER Variables and Methods	Amato (‘21); Goljan (‘21); Violette (‘21)

DoD Phasing Definitions

1. Requirements Analysis
2. Architecture & Design
3. Coding & Testing
4. System Integration
5. Qualification Testing
6. Development Test & Evaluation



Rules of Thumb



← AFCAA Cost Est. Handbook: 40/20/40 split (~Thibodeau)

Data Source

- DoD Software Resources Data Report (SRDR)
- 5074 CSCIs, 2000+ programs, 150+ metrics

After Data Exclusion process:

- 1128 reports for Rule of Thumb Analysis
- 308 initial/final pairs (616 reports) for Schedule Analysis

Methods

Rule of Thumb:

- Least Squared Means – OLS-derived, calculates magnitude of difference
- Hotelling's T^2 test – test vector equivalency
  - More rigorous version of MANOVA, evaluates vectors of means against ea. other

Schedule:

- Contingency analysis, t-tests – evaluate how %Phase X affects schedule growth
  - Two cohorts: higher than & lower than mean/median schedule growth
- Service, Super Domain, Application Domain, Op. Environment, Dev. Process, Size (ESLOC)

Categories Analyzed

Service	Air Force, Army, Navy
Operating Environment	Surface, Air, Sea, Space
Super Domain	AIS, ENG, MS, RT
Application Domain	C&C, CAD, COM, M&F, MP, PC, RTE, S&S, SP, SS, TMDE, DOOL, TRN, VC, VP
Size (ESLOC)	<5K, 6-20K, 21-50K, 51-100K, 101-250K, >250K

Results

		Final Percent Effort, Mean Values					
	N	Phase 1	Phase 2	Phase 3	Phase 4	Phase 5	Phase 6
Full Dataset	722	11%	18%	37%	17%	12%	6%
Service							
Air Force	230	8%	19%	40%	19%	8%	6%
Army	216	16%	18%	33%	14%	17%	3%
Navy	276	11%	18%	36%	17%	11%	7%
Super Domain							
AIS	48	9%	21%	44%	16%	6%	4%
ENG	114	9%	21%	35%	18%	11%	6%
MS	26	18%	22%	30%	15%	6%	9%
RT	534	12%	17%	37%	17%	12%	6%
Application Domain							
C&C	164	13%	16%	42%	12%	11%	6%
CAS	15	11%	23%	41%	16%	6%	3%
COM	68	7%	24%	32%	21%	11%	4%
M&F	1	7%	26%	33%	30%	0%	4%
MP	33	8%	20%	46%	15%	7%	4%
PC	1	7%	18%	60%	5%	4%	6%
RTE	119	13%	18%	35%	18%	11%	5%
S&S	33	10%	18%	32%	17%	15%	8%
SP	77	12%	18%	33%	22%	10%	5%
SS	60	9%	23%	36%	21%	5%	6%
TMDE	20	5%	18%	37%	14%	22%	4%
TOOL	16	12%	27%	27%	23%	5%	6%
TRN	10	26%	16%	34%	4%	7%	13%
VC	62	13%	14%	35%	14%	17%	7%
VP	43	10%	12%	35%	19%	20%	4%
OE - Summary							
Air Vehicle	268	9%	16%	39%	17%	13%	7%
Sea System	78	9%	18%	37%	17%	18%	1%
Space System	48	10%	14%	39%	14%	8%	14%
Surface	339	15%	21%	34%	17%	10%	5%
Process							
Agile	123	22%	13%	41%	10%	9%	5%
Traditional	595	9%	19%	36%	18%	12%	6%
Initial Size							
<5K	104	13%	14%	34%	20%	15%	4%
5-20K	188	11%	19%	37%	17%	12%	4%
20-50K	174	9%	20%	36%	18%	11%	5%
50-100K	113	10%	17%	39%	15%	10%	9%
100-250K	105	14%	19%	35%	14%	11%	7%
>250K	38	13%	20%	36%	15%	13%	6%

Final %Phase Allocation

Darker blue = higher % effort (horizontally)

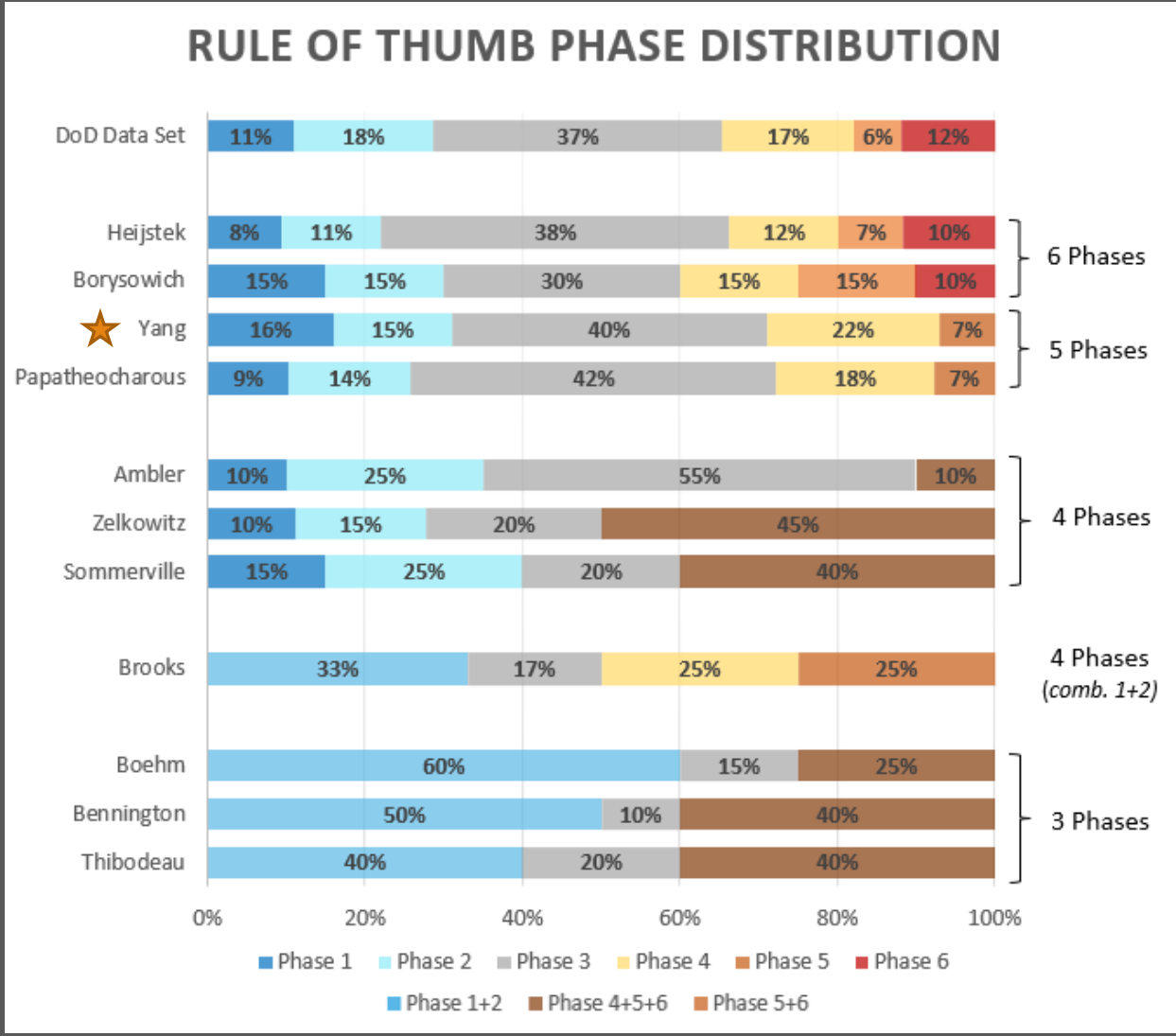
Similar phasing, regardless of category

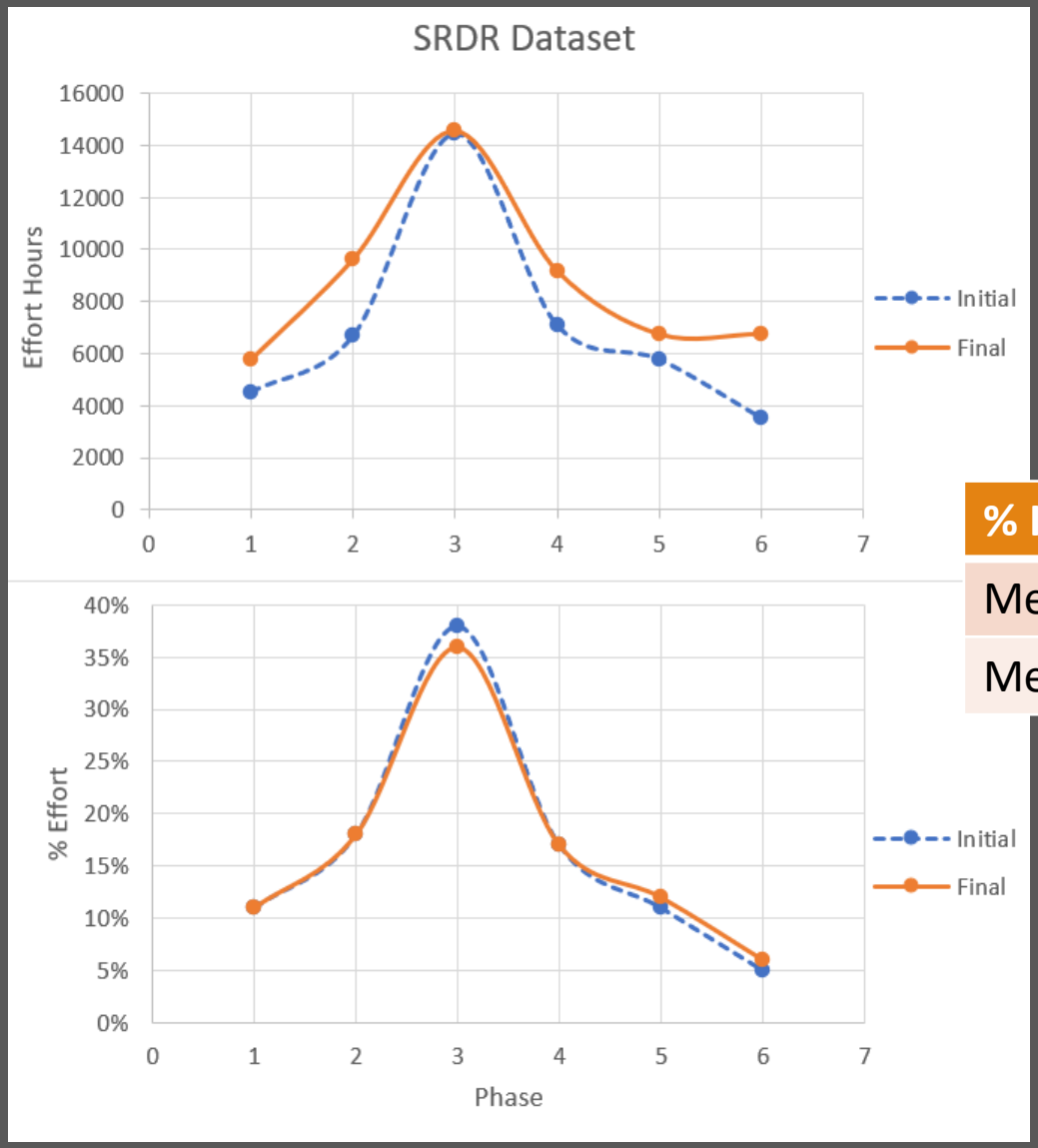
SRDR vs Rules of Thumb

LSM method showed Yang, Papatheocharous, & Borysowich the closest matches

Worst RoT were Ambler, Boehm

Hotelling Test inconclusive





Overall Effort

% Increase Effort	% Increase Size
Mean: 113%	Mean: 136%
Median: 33%	Median: 37%

Schedule v. Effort

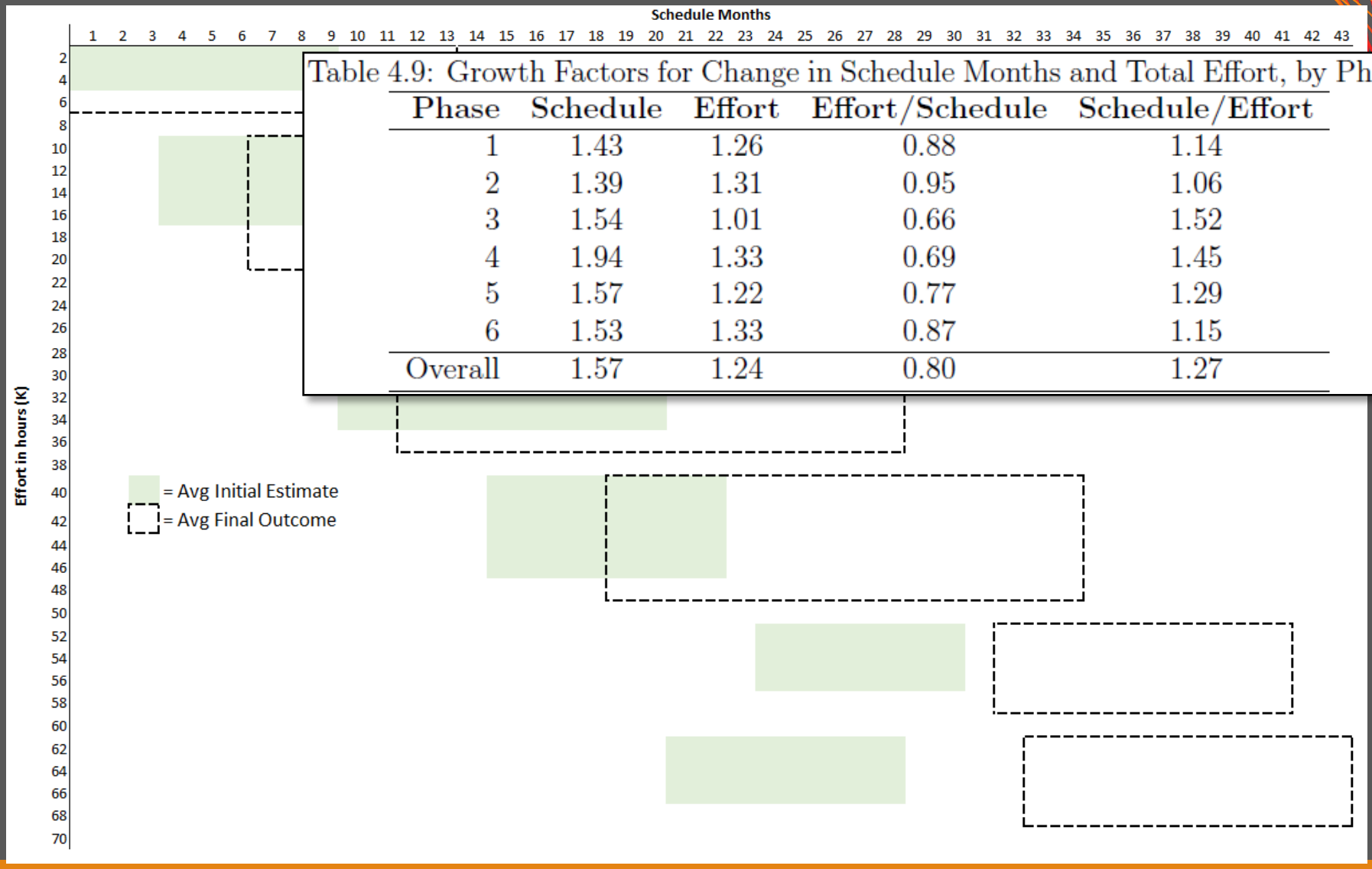
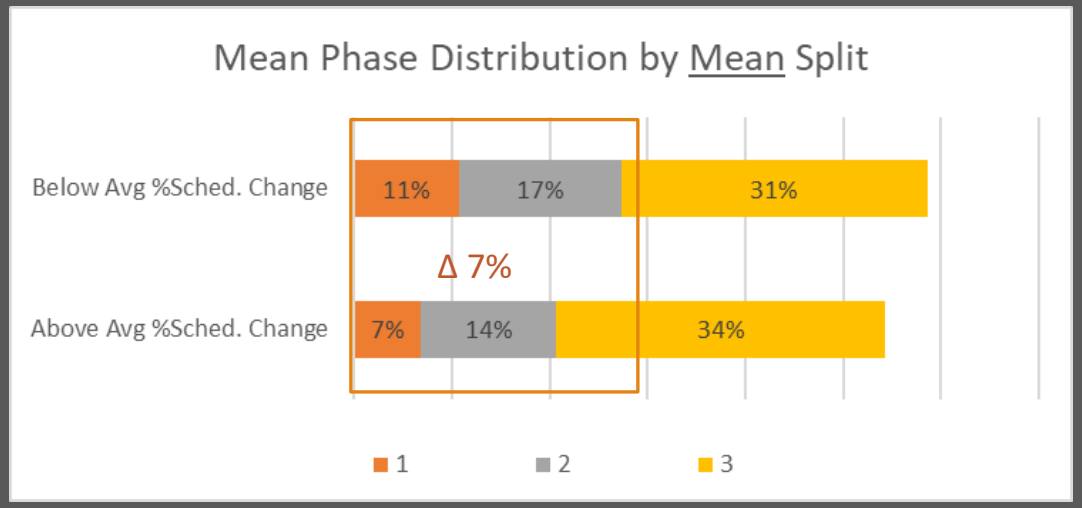
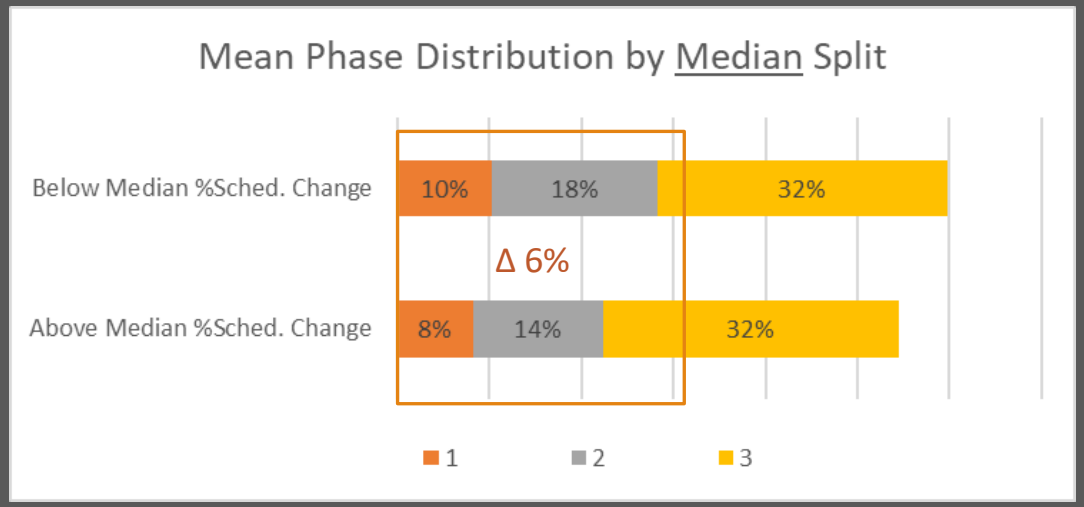


Table 4.9: Growth Factors for Change in Schedule Months and Total Effort, by Phase

Phase	Schedule	Effort	Effort/Schedule	Schedule/Effort
1	1.43	1.26	0.88	1.14
2	1.39	1.31	0.95	1.06
3	1.54	1.01	0.66	1.52
4	1.94	1.33	0.69	1.45
5	1.57	1.22	0.77	1.29
6	1.53	1.33	0.87	1.15
Overall	1.57	1.24	0.80	1.27





Schedule Growth: Two Cohorts

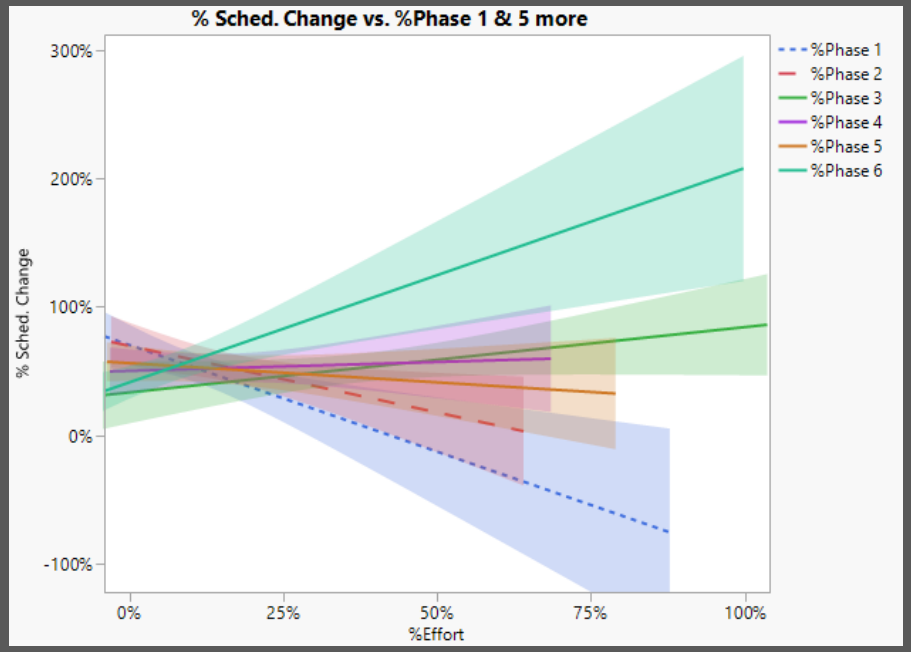
Mean growth = 52%

Median growth = 17%

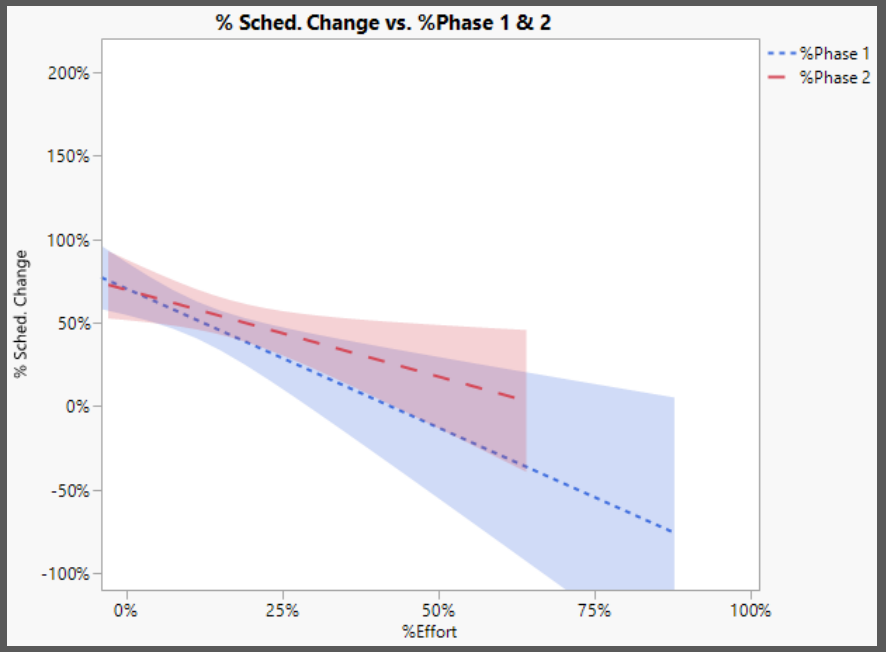
Comparison of Avg. Percent Effort in Phase 1+2 Between the Above-Mean Cohort and the Below-Mean Cohort

	N	Below Mean of 52%	Above Mean of 52%	Delta	t Ratio	p-value
Overall	308	30%	21%	9%	-3.966	0.0001
Service						
Air Force	76	28%	20%	8%	-2.129	0.0366
Army	96	29%	17%	12%	-1.775	0.0791
Navy	137	32%	23%	9%	-3.437	0.0008
Super Domain						
AIS	25	36%	21%	15%	-2.249	0.0344
ENG	39	29%	18%	11%	-2.155	0.0377
MS	12	36%	31%	5%	-0.435	0.6726
RT	233	29%	21%	8%	-3.027	0.0027
Application Domain						
C&C	66	26%	21%	5%	-1.119	0.2711
CAS	11	42%	19%	23%	-3.541	0.0489
COM	22	38%	31%	7%	-0.375	0.4664
M&F	1	.	33%			
MP	14	30%	23%	7%	-1.158	0.273
RTE	54	30%	20%	10%	-2.307	0.0250
S&S	12	28%	17%	11%	-1.197	0.2588
SP	35	28%	29%	-1%	0.219	0.828
SS	20	31%	12%	19%	-2.163	0.0442
TMDE	7	25%	26%	-1%	0.116	0.912
TOOL	6	35%	33%	2%	-0.086	0.9352
TRN	6	39%	31%	8%	-0.463	0.667
VC	33	29%	11%	18%	-2.451	0.0201
VP	22	24%	12%	12%	-1.698	0.105
Operating Env.						
Air	137	25%	22%	4%	-1.273	0.2052
Sea	32	30%	18%	12%	-1.401	0.1715
Space	15	19%	10%	9%	-1.620	0.1286
Surface	125	34%	25%	9%	-2.580	0.0110
Process						
Agile	206	30%	20%	10%	-3.914	0.0001
Traditional	103	29%	25%	4%	-1.225	0.2235
Initial Size						
<5K	52	26%	14%	12%	-1.678	0.0995
5-20K	97	31%	21%	10%	-2.952	0.004
20-50K	80	28%	21%	7%	-2.177	0.0325
50-100K	40	33%	25%	8%	-1.116	0.2716
100-250K	26	30%	20%	10%	-1.102	0.2837
>250K	12	29%	25%	4%	-0.296	0.7734

Trends for %Effort vs. %Schedule



All phases



Phase 1 & 2

Conclusions

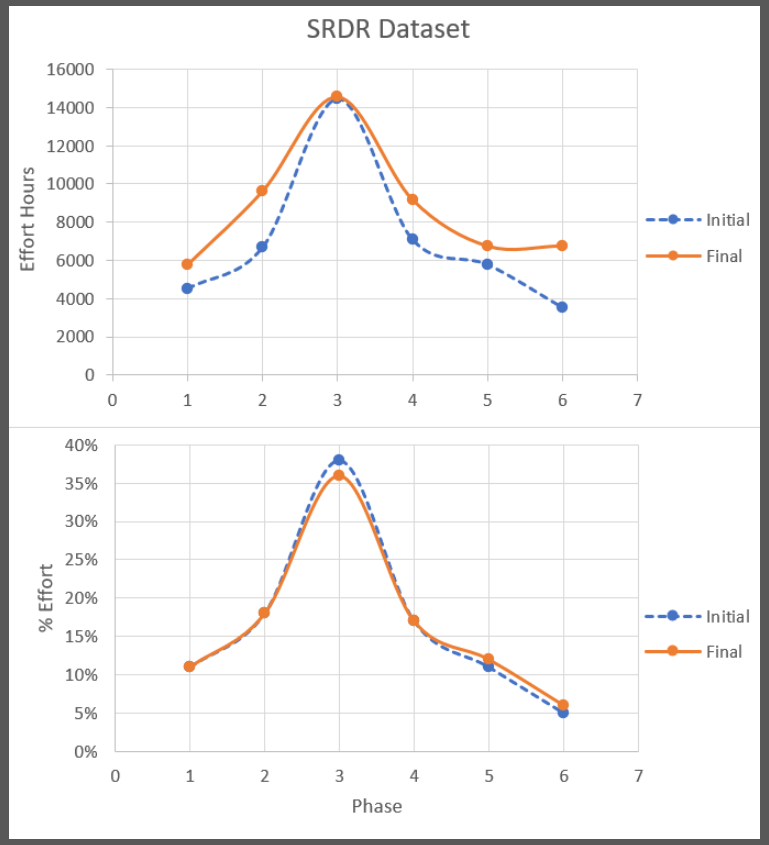
Overall Phasing Dist. Change

Minimal change in avg. distribution

Increase in size

Overall Percent Effort Change from Initial to Final

	1	2	3	4	5	6
Overall	0%	1%	1%	0%	-2%	-1%



Rules of Thumb

- Cannot statistically ID specific RoT (Hotelling Test)

- As general guide, can point to RoT that are more helpful, match SRDR means better

Table 5.1: Means of Dataset and Proposed Rules of Thumb

	Phase 1	Phase 2	Phase 3	Phase 4	Phase 5+6
SRDR Mean	11%	18%	37%	16%	18%
New Rule of Thumb	**15%**	**15%**	**40%**	**20%**	**10%**
Yang	16%	15%	40%	22%	7%
Papatheocharous	9%	14%	42%	18%	7%
Borysowich	15%	15%	30%	15%	25%

Simplified: 30/40/30

Schedule

Projects with larger-than-average schedule growth spent less %effort in early phases (1 and 2)

→ Increase early effort phase allocation, reduce Pareto effect

Phases	1	2	3	4	5	6
>Mean Schedule Growth	7%	14%	39%	18%	11%	11%
Zero/Neg Schedule Growth	13%	19%	35%	16%	10%	6%

Adjusted RoT: 33/35/32

~ even split among Design/Coding/Testing

1/3 Req's & Design
1/3 Coding
1/3 Testing & Integration

Recommendations

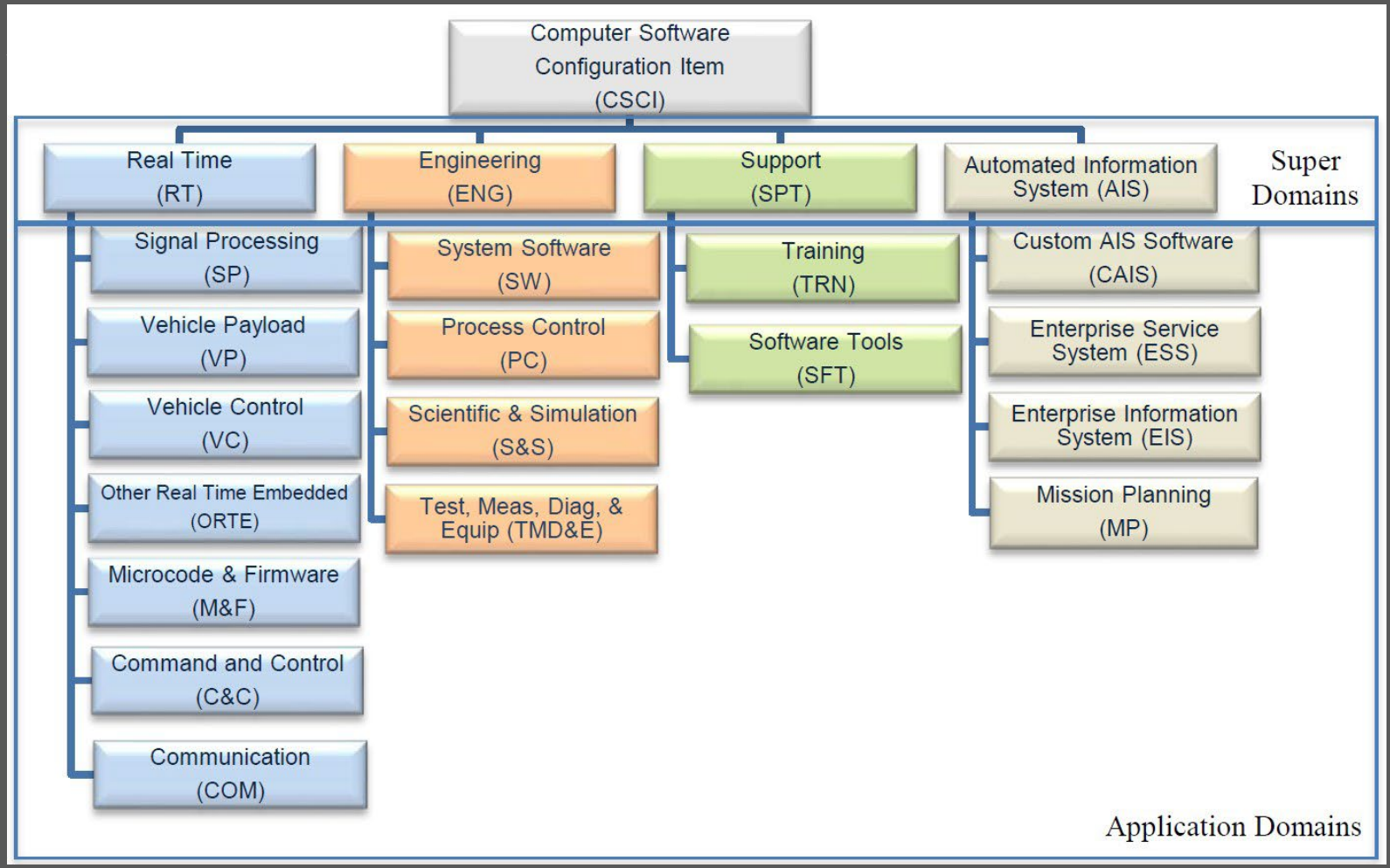
- Recommend 33/34/33 split when first planning new SW starts
- If the above RoT is not used, consider allocating >30% to phase 1 and 2
- If schedule slips and early phases are extended, don't try to start subsequent phases on time or rush early work (encourages Pareto rule)

Limitations

1. High variability in data – limits applicability of any prescribed RoT
2. Phase definitions between SRDR and RoT
3. Inconsistent data entry by Contractors in SRDR
4. Small sample compared to population (~300 v. 5000, just 6%); rejected reports could have different phase distribution

Questions?

Domains


Computer Software Configuration Item (CSCI)
Real Time (RT)
Engineering (ENG)
Support (SPT)
Automated Information System (AIS)
Super Domains
Signal Processing (SP)
Vehicle Payload (VP)
Vehicle Control (VC)
Other Real Time Embedded (ORTE)
Microcode & Firmware (M&F)
Command and Control (C&C)
Communication (COM)
System Software (SW)
Process Control (PC)
Scientific & Simulation (S&S)
Test, Meas, Diag, & Equip (TMD&E)
Training (TRN)
Software Tools (SFT)
Custom AIS Software (CAIS)
Enterprise Service System (ESS)
Enterprise Information System (EIS)
Mission Planning (MP)
Application Domains

Operating Environments

Operating Environment (OE)		Examples
Surface Fixed (SF)	Fixed (SF)	software is at a fixed site
Surface Mobile (SM)	Manned (SMM)	software is moved somewhere and set up
	Unmanned (SMU)	
Surface Portable (SP)	Manned (SPM)	software is in a handheld device
Surface Vehicle (SV)	Unmanned (SVM)	software is embedded in as part of a moving ground vehicle
	Unmanned (SVU)	
Air Vehicle (AV)	Manned (AVM)	software is embedded as part of an aircraft
	Unmanned (AVU)	
Sea System (SS)	Manned (SSM)	software is embedded as part of a surface or underwater boat/ship or boat
	Unmanned (SSU)	
Missile System (MS)	Unmanned (MSU)	Software is embedded as part of a missile system
Ordnance System (OS)	Unmanned (OSU)	software is embedded as part of an ordnance system
Space System (SPS)	Manned (SPSM)	software is embedded as part of a spacecraft
	Unmanned (SPSU)	