# SEI IPA/SEC Research Status

William R. Nichols

Software Engineering Institute
Carnegie Mellon University
Pittsburgh, PA 15213




DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution

Copyright 2018 Carnegie Mellon University. All Rights Reserved.

This material is based upon work funded and supported by the Department of Defense under Contract No. FA8702-15-D-0002 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center.

The view, opinions, and/or findings contained in this material are those of the author(s) and should not be construed as an official Government position, policy, or decision, unless designated by other documentation.

NO WARRANTY. THIS CARNEGIE MELLON UNIVERSITY AND SOFTWARE ENGINEERING INSTITUTE MATERIAL IS FURNISHED ON AN "AS-IS" BASIS. CARNEGIE MELLON UNIVERSITY MAKES NO WARRANTIES OF ANY KIND, EITHER EXPRESSED OR IMPLIED, AS TO ANY MATTER INCLUDING, BUT NOT LIMITED TO, WARRANTY OF FITNESS FOR PURPOSE OR MERCHANTABILITY, EXCLUSIVITY, OR RESULTS OBTAINED FROM USE OF THE MATERIAL. CARNEGIE MELLON UNIVERSITY DOES NOT MAKE ANY WARRANTY OF ANY KIND WITH RESPECT TO FREEDOM FROM PATENT, TRADEMARK, OR COPYRIGHT INFRINGEMENT.

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution. Please see Copyright notice for non-US Government use and distribution.

This material may be reproduced in its entirety, without modification, and freely distributed in written or electronic form without requesting formal permission. Permission is required for any other use. Requests for permission should be directed to the Software Engineering Institute at permission@sei.cmu.edu.

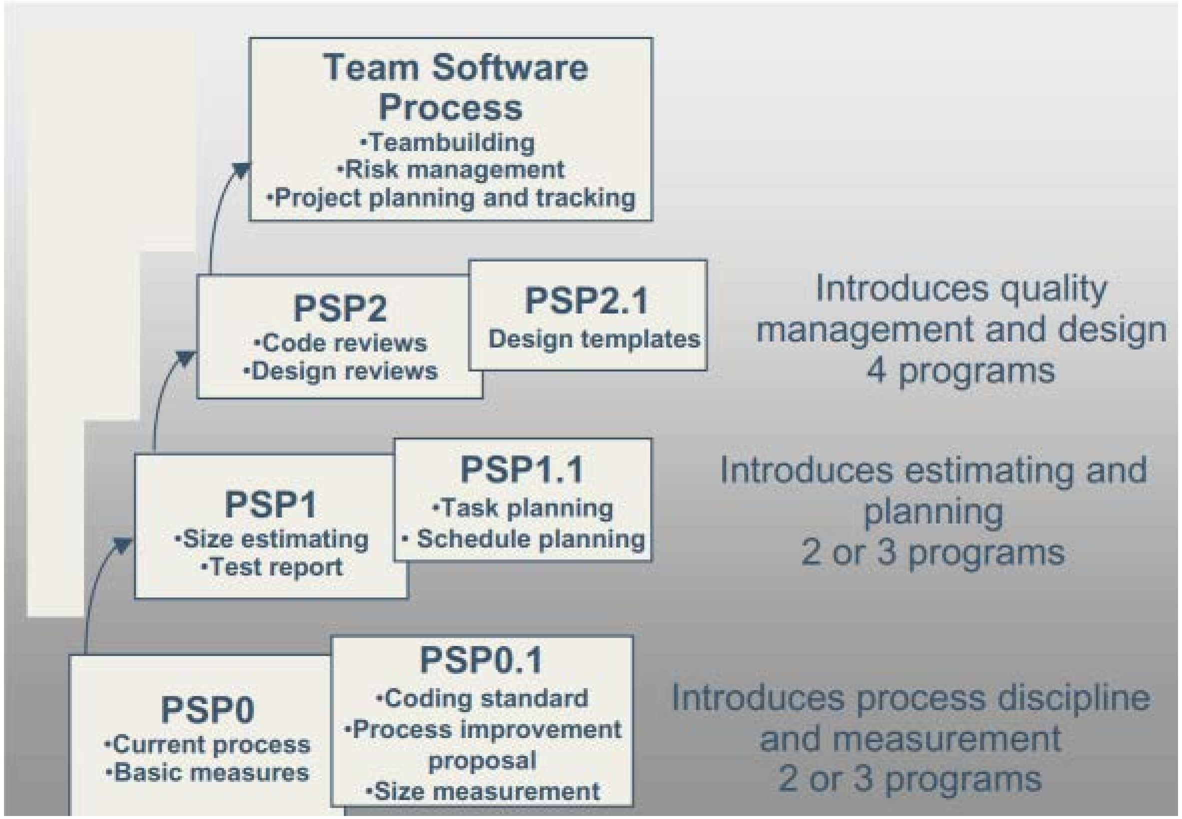
PSP[SM] and TSP[SM] are service marks of Carnegie Mellon University.

DM18-0185


[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

# PSP database

Suitable  for Quasi-expermiental methods.





[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

# PSP Data Overview

When using the PSP, developers gather and use data.

- Time data
  - The time in minutes spent by development task
  - Interruption time is not counted.
- Size data
  - Product size in db elements, pages, LOC, etc.
  - Categories: base, added, deleted, modified, reused
- Defect data
  - All defects removed in compile, test, review, Type, phases injected & removed, fix time, description


[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

# PSP Data by Language

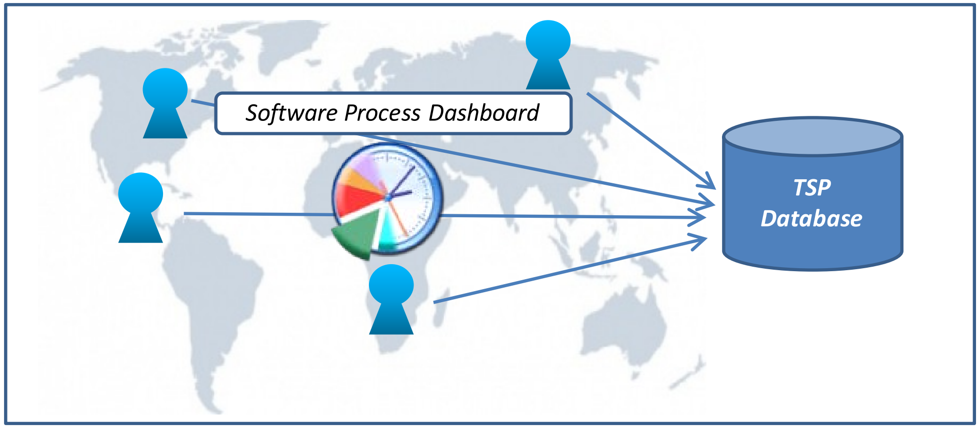
| Language | Programs | LOC | Hours | Defects |
|---|---|---|---|---|
| C | 4,984 | 532,529 | 21,460.80 | 36,426 |
| C++ | 3,255 | 448,517 | 14,913.40 | 30,785 |
| C# | 1,213 | 163,233 | 3,696.60 | 6,661 |
| VB | 1,353 | 144,621 | 5,108.50 | 7,405 |
| Java | 1,383 | 199,493 | 6,311.00 | 11,131 |
| Ada | 286 | 33,060 | 1,869.00 | 3,477 |
| Total | **12,474** | **1,521,453** | **53,359.20** | **95,885** |


[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

# The TSP SEMPR Database

## Stores project data

- >900 projects launched after 2009
- used the Software Process Dashboard
- In a relational database (MySQL 5.6) via SQL



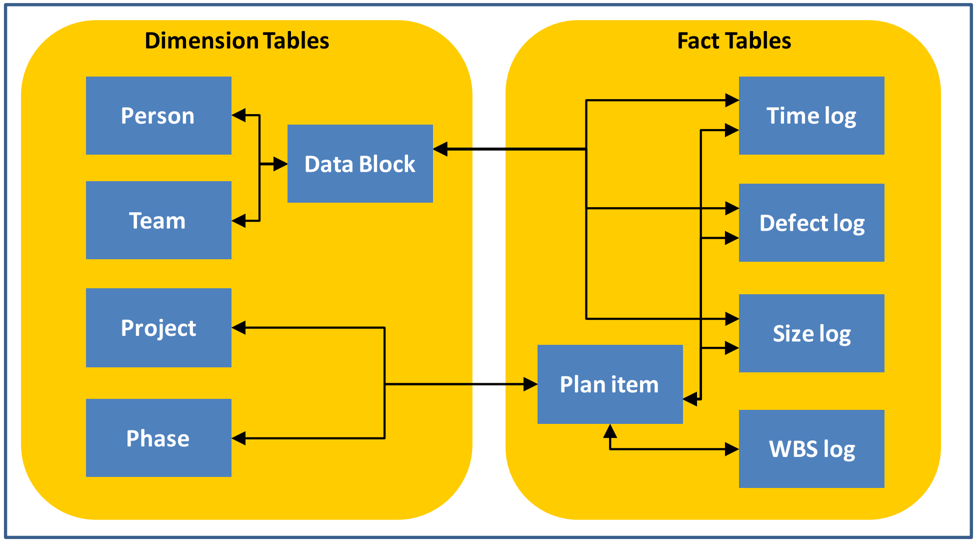
# The TSP Database

*Follows the schema of the Team Process Data Warehouse.*

*Includes Fact Tables and Dimension Tables.*

*Connects Dimension Tables to Fact Tables for data analysis from many perspectives.*

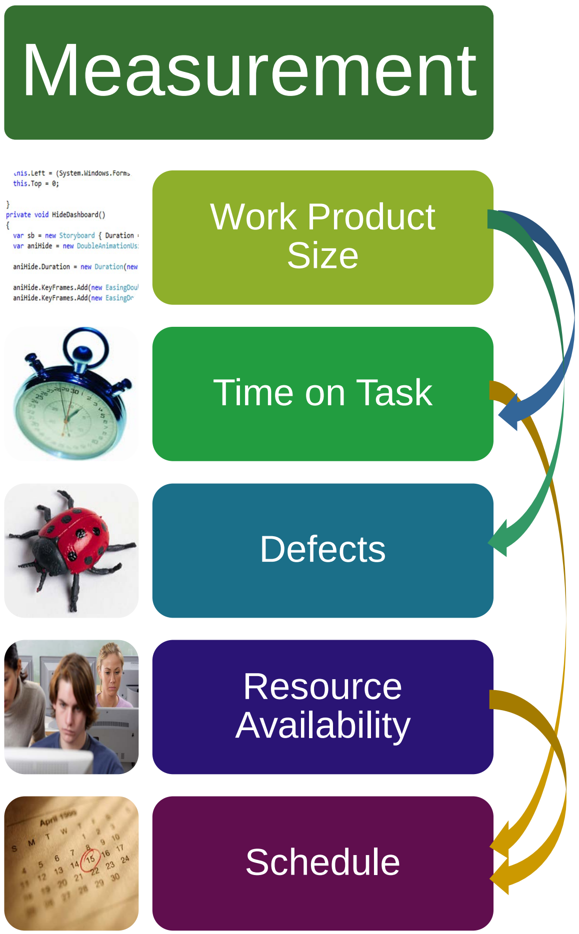
**http://www.processdash.com/**






[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

# Team Software Process (TSP$_{SM}$) measurement framework




Five direct measures

Team and team member data

Estimated during planning

Measured while working

Evaluated weekly or when a

- task is completed
- process phase is completed
- component is completed
- cycle is completed
- project is completed



[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

# Types of Data, Individual

Planned total effort per period

Actual total effort per period

Planned task effort for each work item

Actual task effort actual task effort for each work item,

Defects found

Defects removed


[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

# Types of Data, Component

Planned component size

Actual component size,

Planned effort in each development phase

Actual effort in each development phase

Planned completion date for each task

Actual completion date for each task

Defects injected in each development phase

Defects removed in each development phase

Number of individuals who worked on each component


[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

# Types of data, Project Context

Project Characteristics and Site Charactistics

See PACE Application Forms

Data includes (not limited to)

- Programming Language
- Organization size
- Goals priority
- Business category
- Application category
- Project Lifecycle Stage
- Tools used


[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

# Types of data, Project outcomes

Planned effort

Actual effort

Number of developers

Planned delivery date

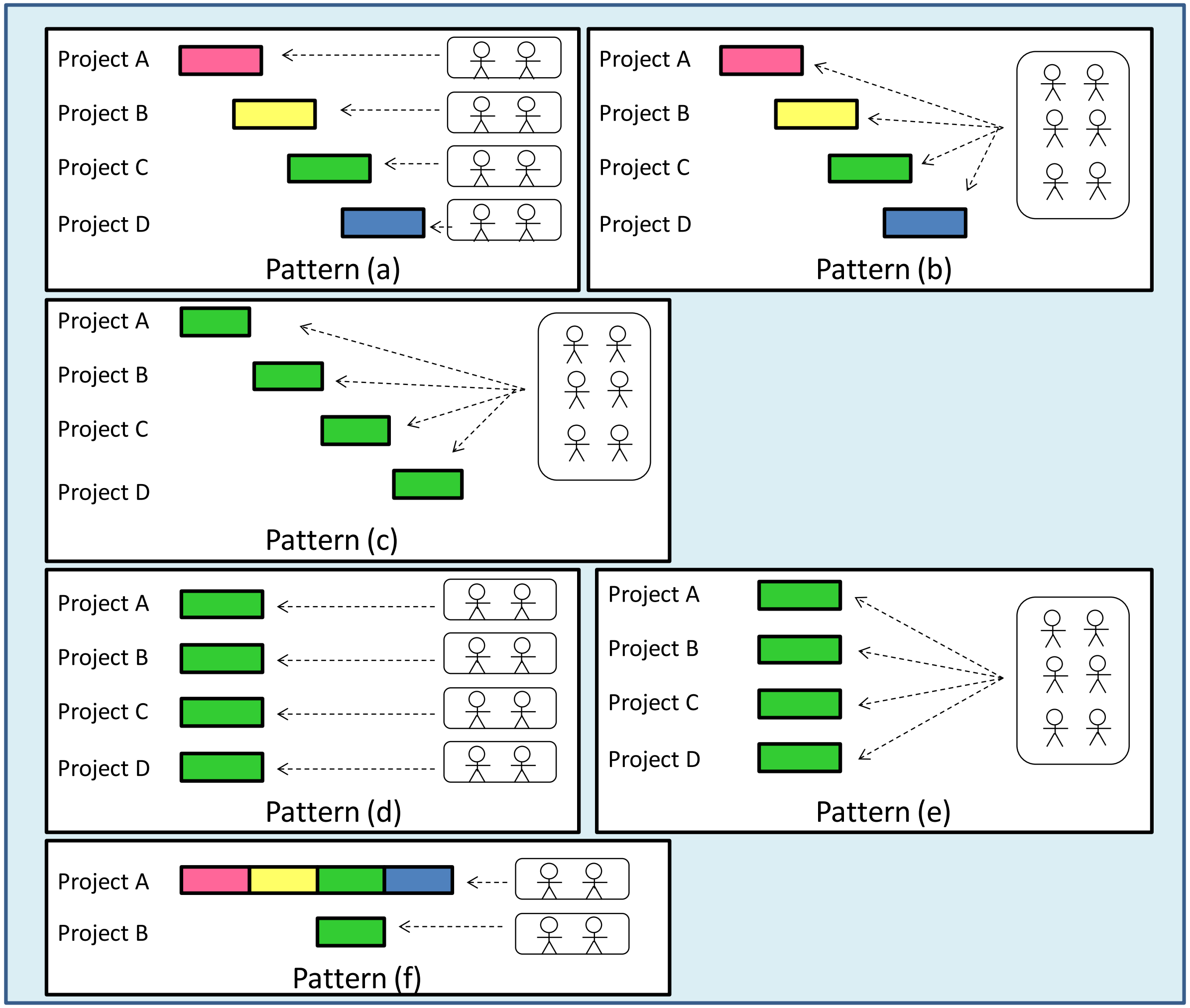
Actual delivery date

Customer Satisfaction results (for PACE projects)

Planned and actual effort in each development phase


[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

# Patterns of Project Organization






[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

# TSP IPA/SEC Research Questions

Are the correlations identified in IPA/SEC data also observed in other data sets?

Which development measures have the highest correlation with external, fielded measures of quality?

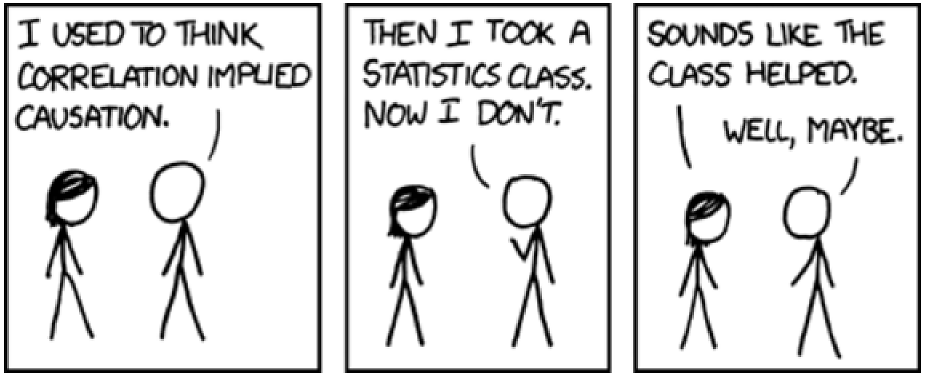
Which measures of quality are candidates for causes of product quality-in-use?

Which internal, development measures have the highest correlation with external, schedule performance?

Does performance differ with  project organizational structure?


[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

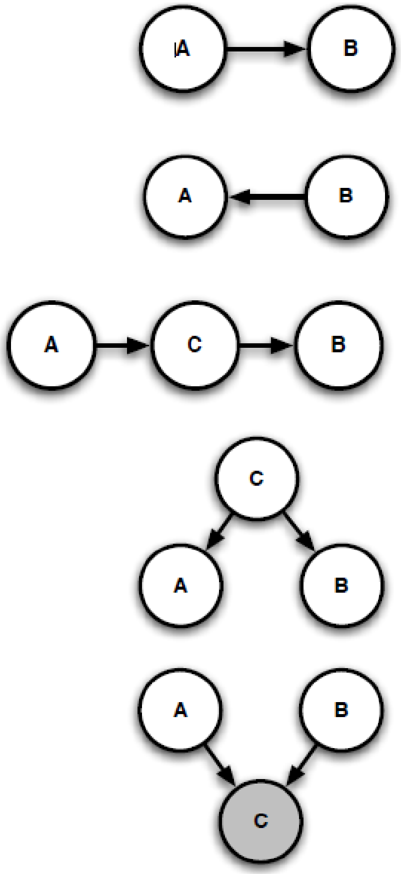
# Investigating Project Success Factors Correlation and Inferring Causation





[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

# Types of association




**Direct causation** — A causes B in the expected direction.

**Reverse causation** — A causes B in the reverse direction.

**Causal chain** —A indirectly causes B through C.

**Common cause** —The variable C causes both A and B, thus inducing a dependence between A and B.

**Conditioning on common effect** — A and B share a common effect C,and conditioning on this variable can Induce a dependence between A and B.


[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

# Analysis approach

Begin with correlational studies (R, Minitab)

Include Quasi-expermental analysis

Apply tools to infer causation

- Tetrad.
  - ❑ Useful for large data samples.
  - ❑ Applied to observational data.
  - ❑ Useful where unobserved confounders are present.
  - ❑ Uses Categorical Data

- Strata and BayesiaLab
  - ❑ Evaluate a causal model
  - ❑ Quantify degree of cause and effect between factors


[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

# Initial Results

Data preparation

Cuts on data

Include only

- New development
- LOC measured (Added and Modified)
- Blank lines and comments not included
- Defects (5269 (released) OR 5253, 5254 (int and sys test) )
- Phased efforts reported

Remove all production rates > 30 LOC/Hr (data went to 600LOC/Hr!)

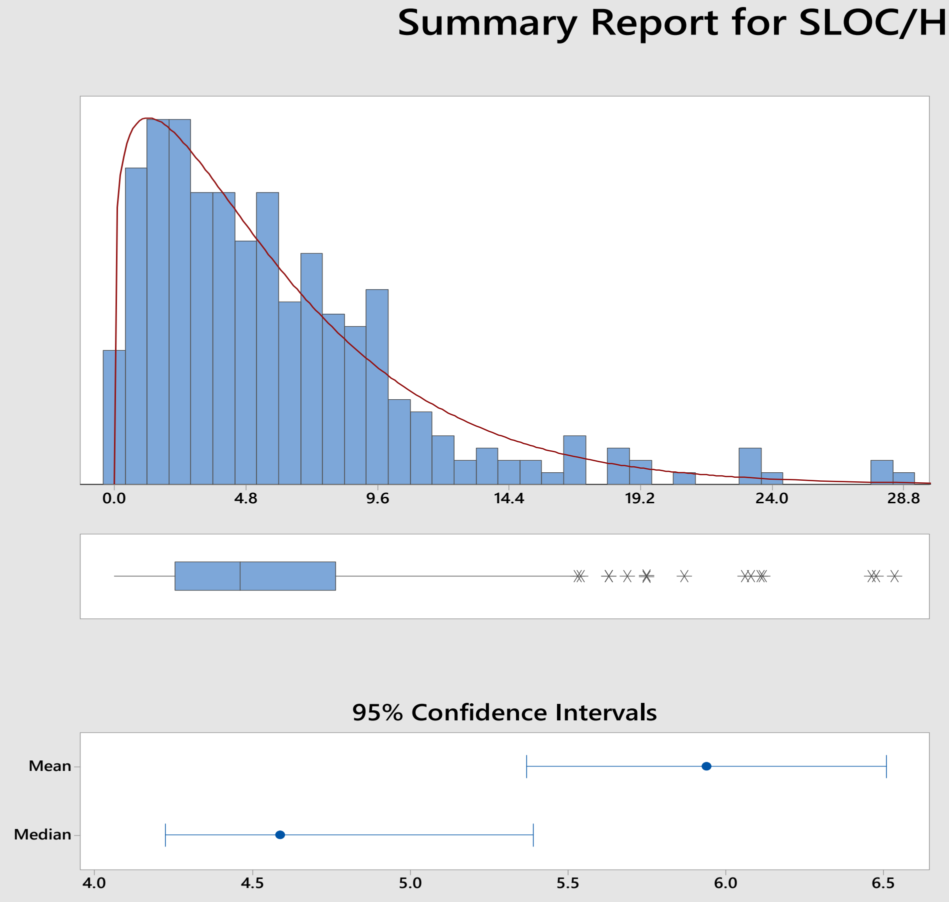
Normalize data by product size


[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

# Code Production Rates

50% of the data is between 4.5 and 8 LOC/Hr

But a substantial number are well outside of this range.

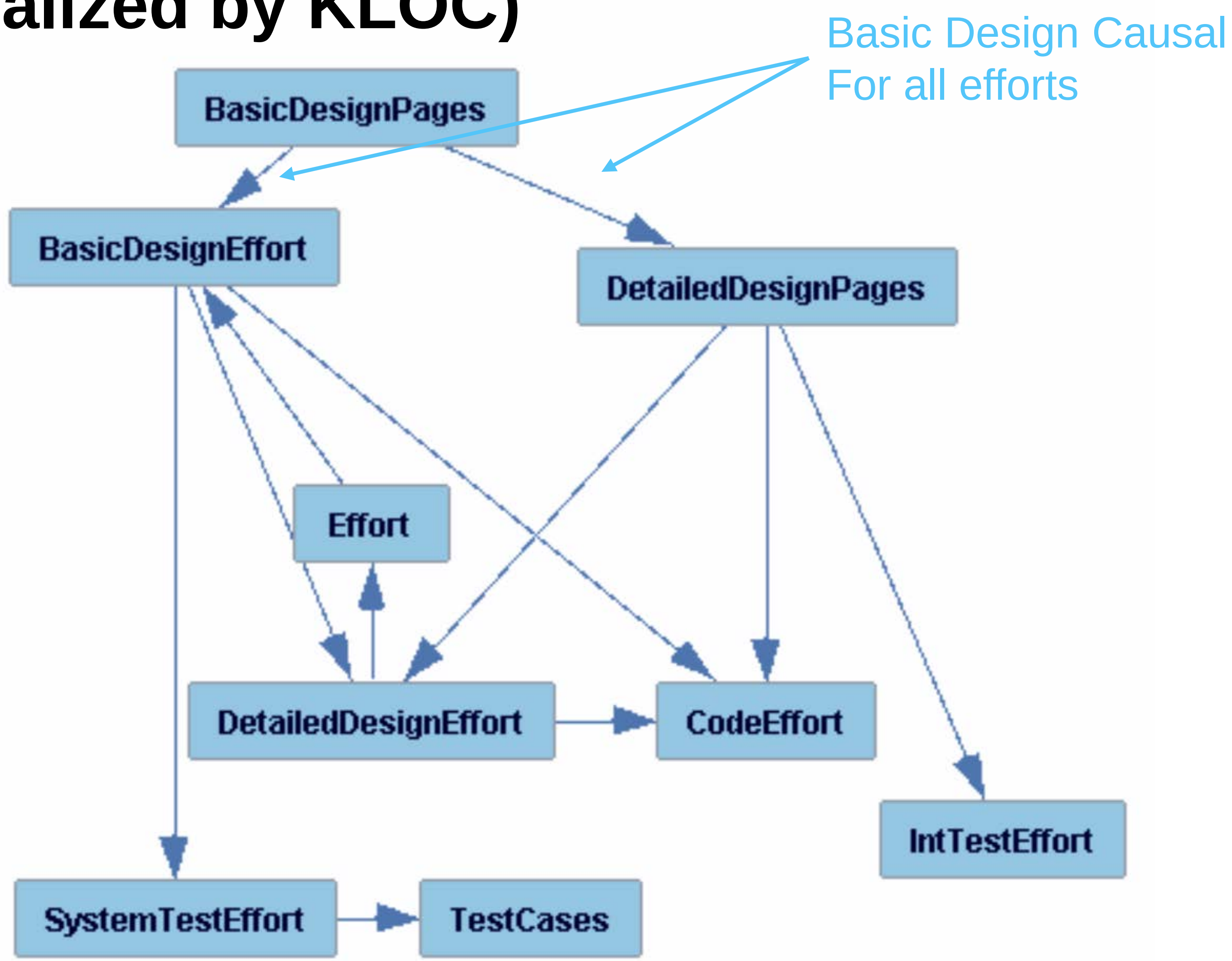



| | |
|---|---|
| **Anderson-Darling Normality Test** | |
| A-Squared | 10.55 |
| P-Value | <0.005 |
| Mean | 5.9390 |
| StDev | 5.0999 |
| Variance | 26.0092 |
| Skewness | 1.74541 |
| Kurtosis | 3.93287 |
| N | 310 |
| Minimum | 0.0065 |
| 1st Quartile | 2.2241 |
| Median | 4.5867 |
| 3rd Quartile | 8.0819 |
| Maximum | 28.4598 |
| **95% Confidence Interval for Mean** | |
| 5.3691 | 6.5090 |
| **95% Confidence Interval for Median** | |
| 4.2240 | 5.3894 |
| **95% Confidence Interval for StDev** | |
| 4.7276 | 5.5364 |



[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

# Results of a causal search (normalized by KLOC)





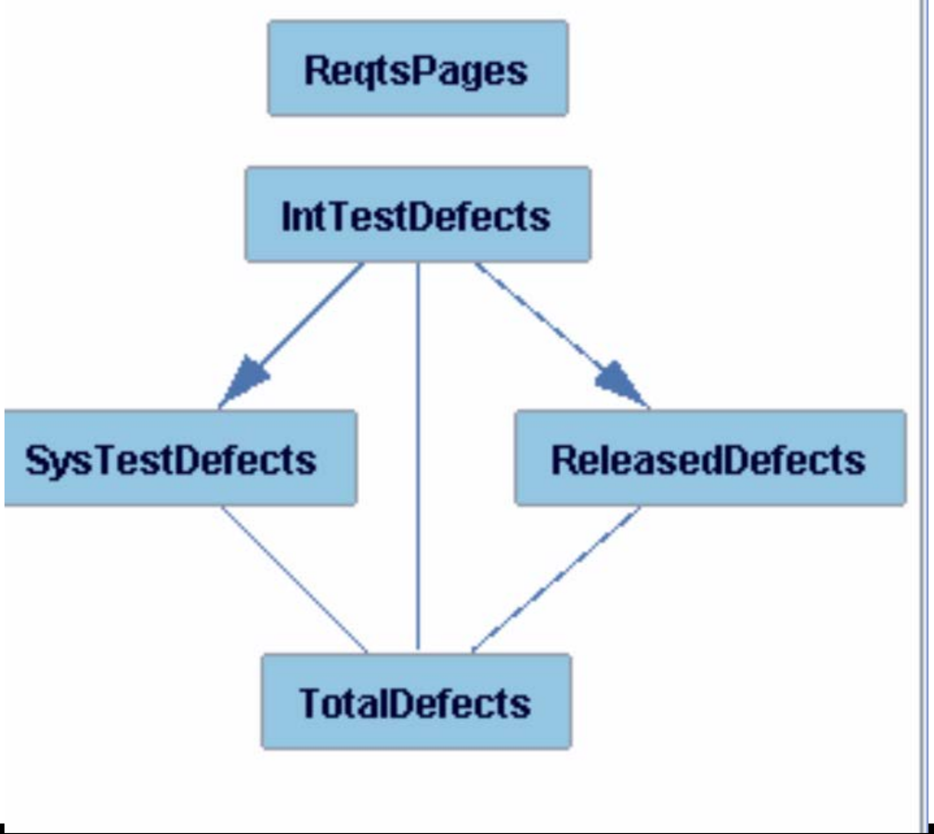
[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

# Surprisingly, other factors were isolated

In what dense to Integration Test defects “cause” later defects?

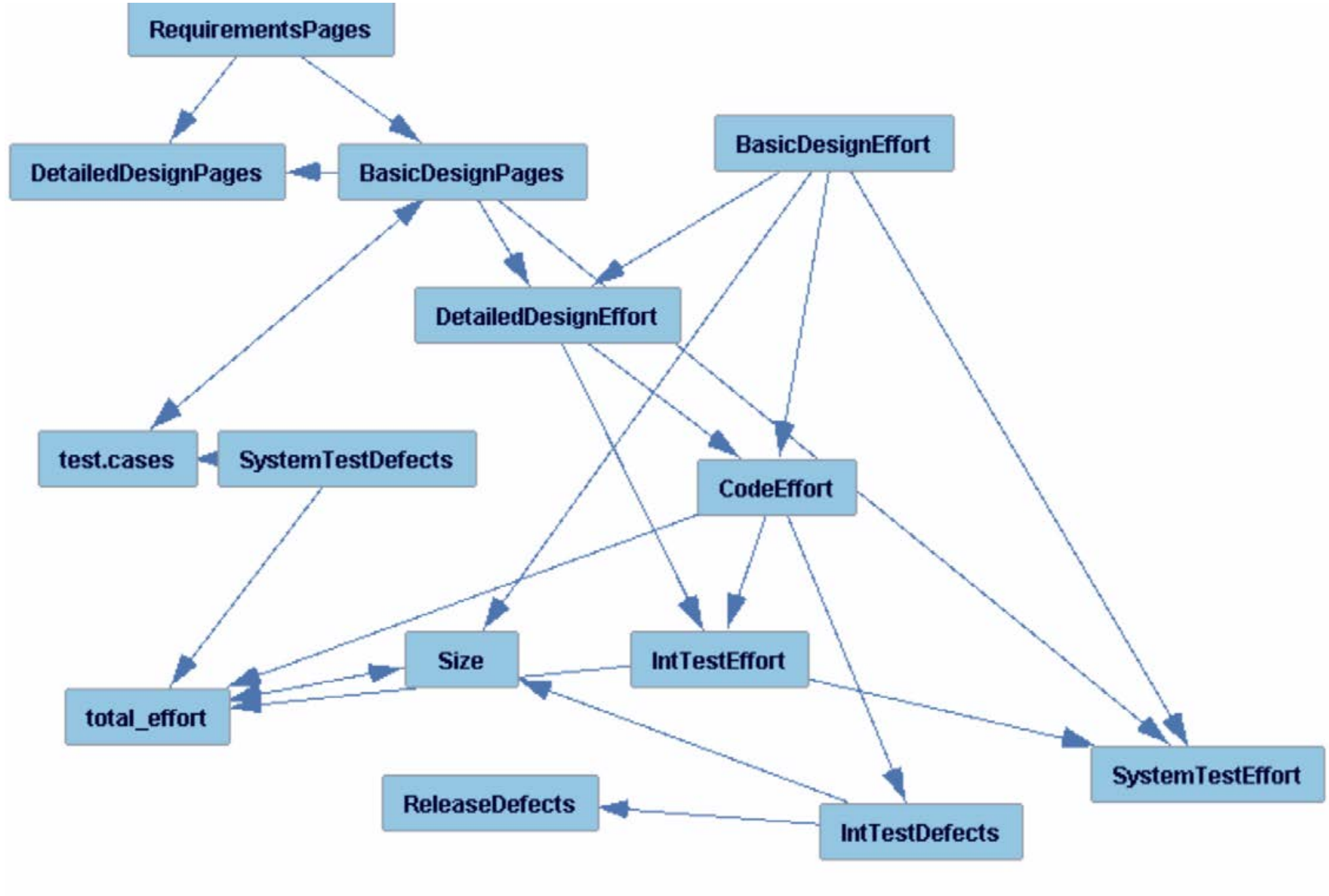
Why is requirements documentation isolated?

Defects are caused by document pages using FGRES algorighm



# Causal search (unnormalized factors)

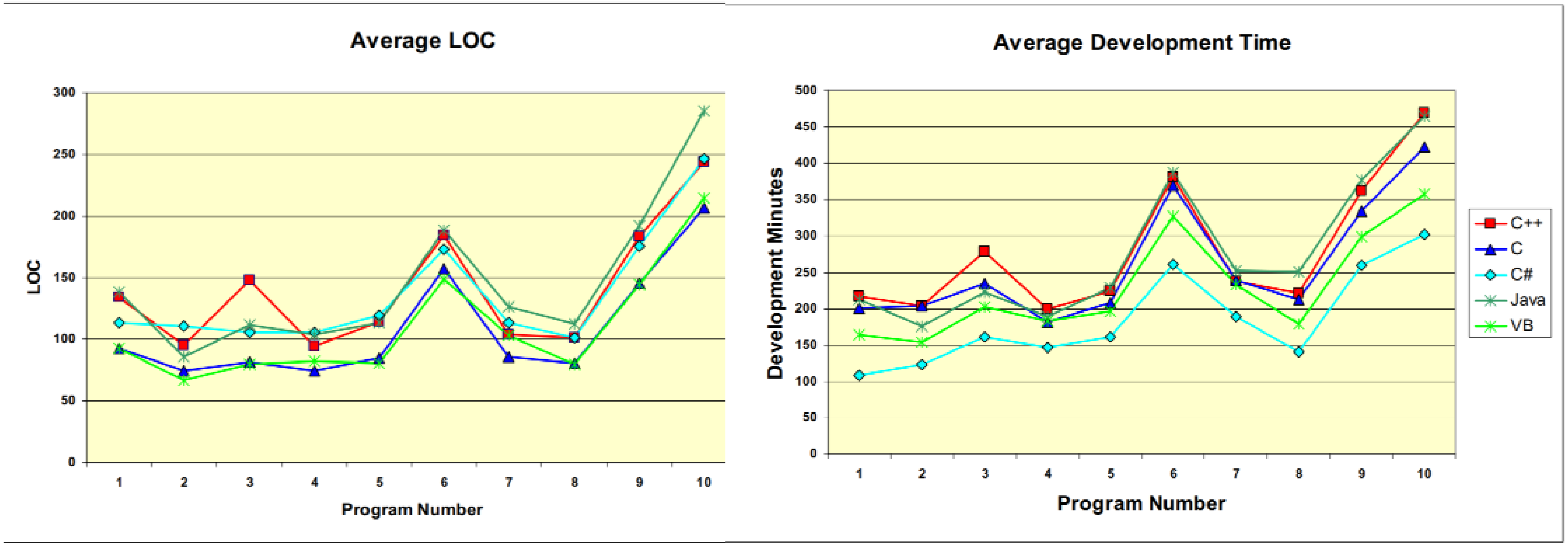
Spearman correlation with Effort and Size (KSLOC) is 0.83



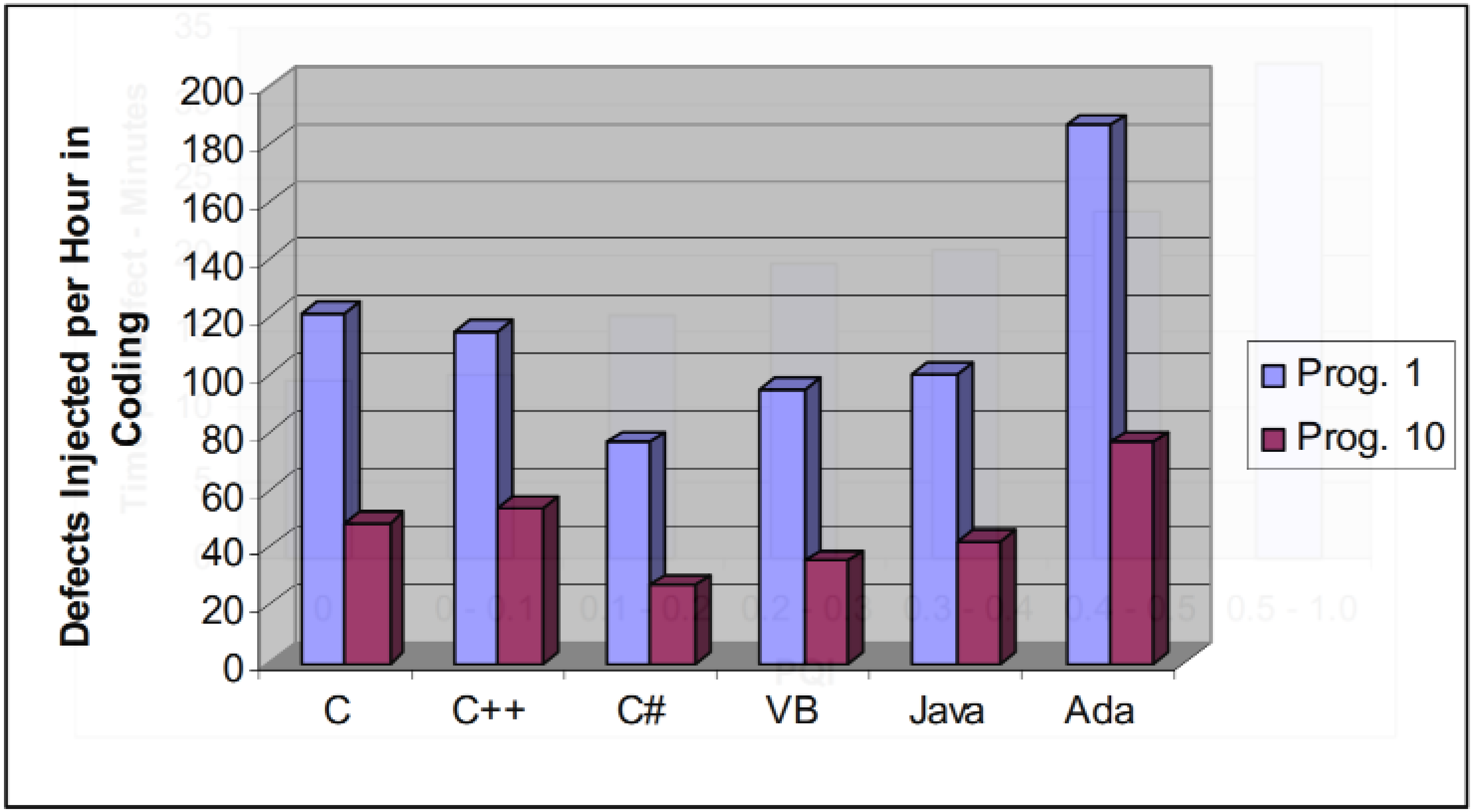
# Effect of programming language:

PSP data can help adjust for language factors for size/effort.

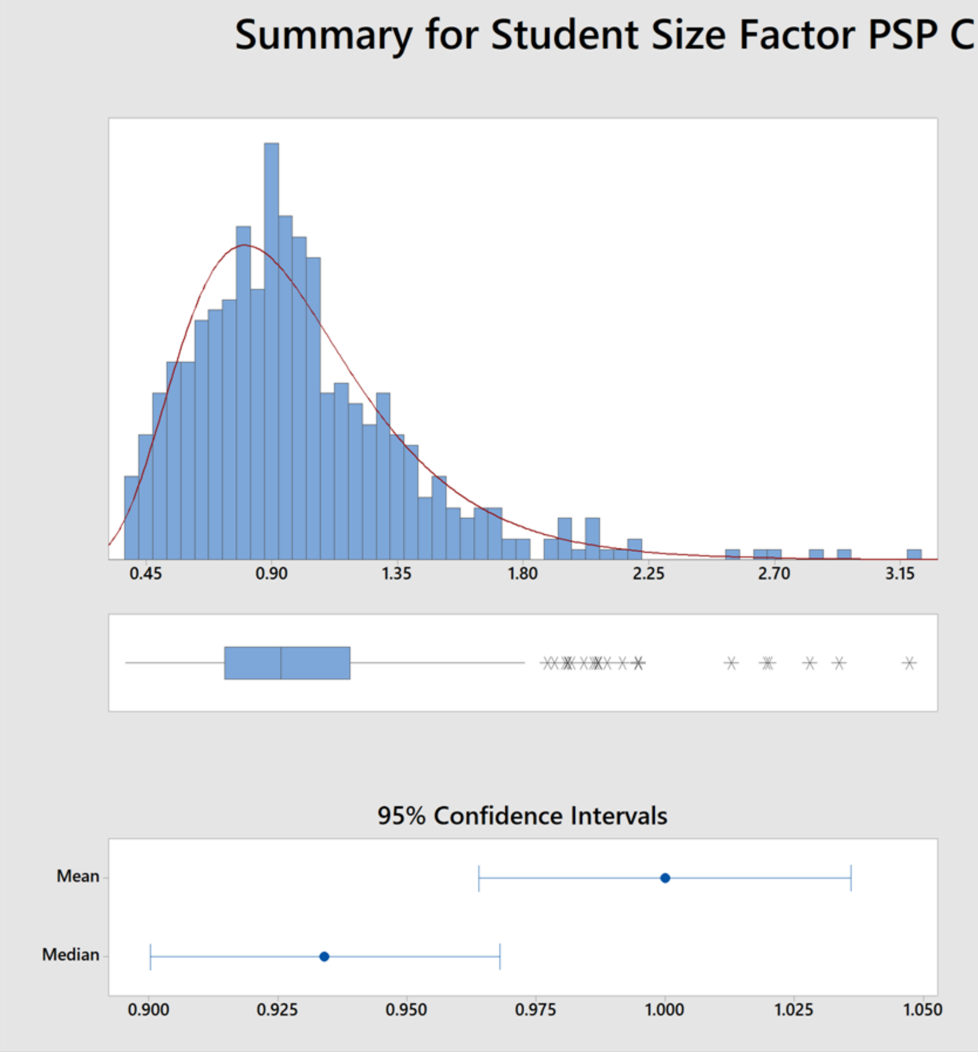
The size factor can supplement benchmarks for KLOC/FP



# Languages have different defect proneness



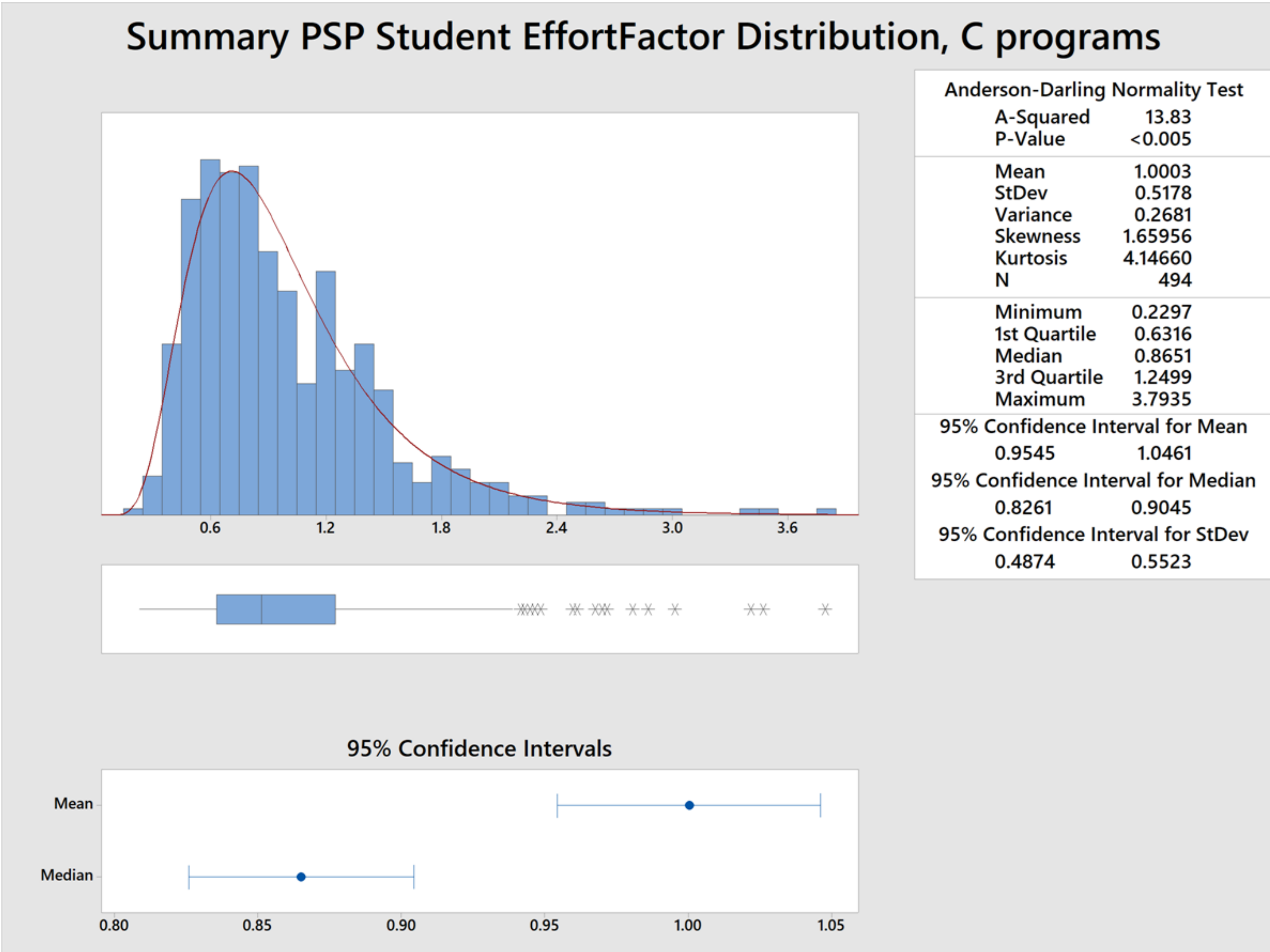
# Size of the programs varied by students!

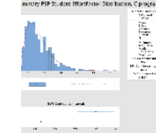



| Anderson-Darling Normality Test | |
|---|---|
| A-Squared | 10.28 |
| P-Value | <0.005 |
| Mean | 1.0000 |
| StDev | 0.4081 |
| Variance | 0.1665 |
| Skewness | 1.59384 |
| Kurtosis | 4.27344 |
| N | 495 |
| Minimum | 0.3769 |
| 1st Quartile | 0.7327 |
| Median | 0.9341 |
| 3rd Quartile | 1.1814 |
| Maximum | 3.1819 |
| 95% Confidence Interval for Mean | |
| 0.9640 | 1.0360 |
| 95% Confidence Interval for Median | |
| 0.9004 | 0.9681 |
| 95% Confidence Interval for StDev | |
| 0.3841 | 0.4352 |


[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

# Student total effort also varies

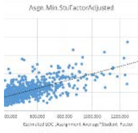


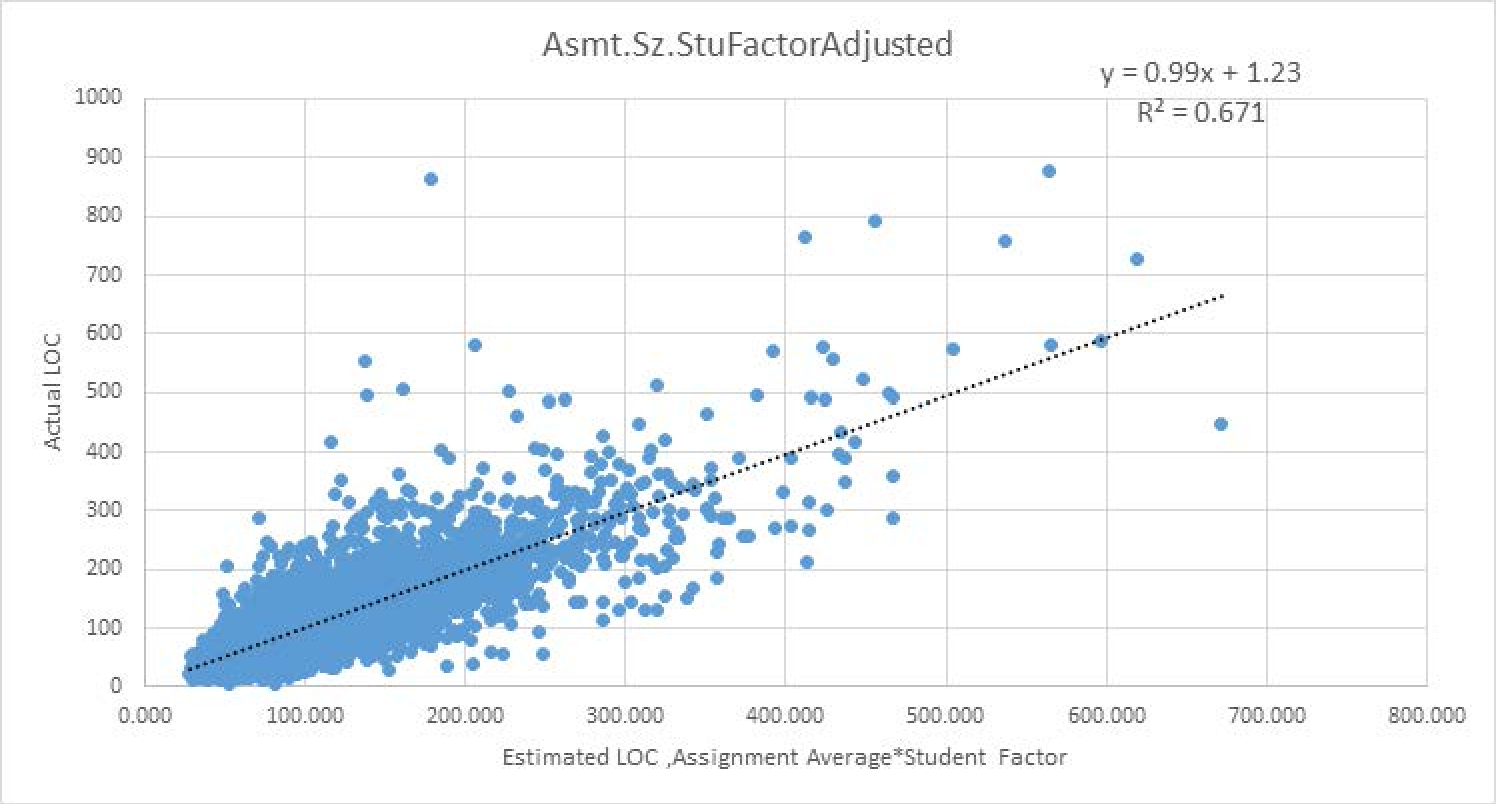





[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

# Adding individual size factors accounts for 67% of the variation

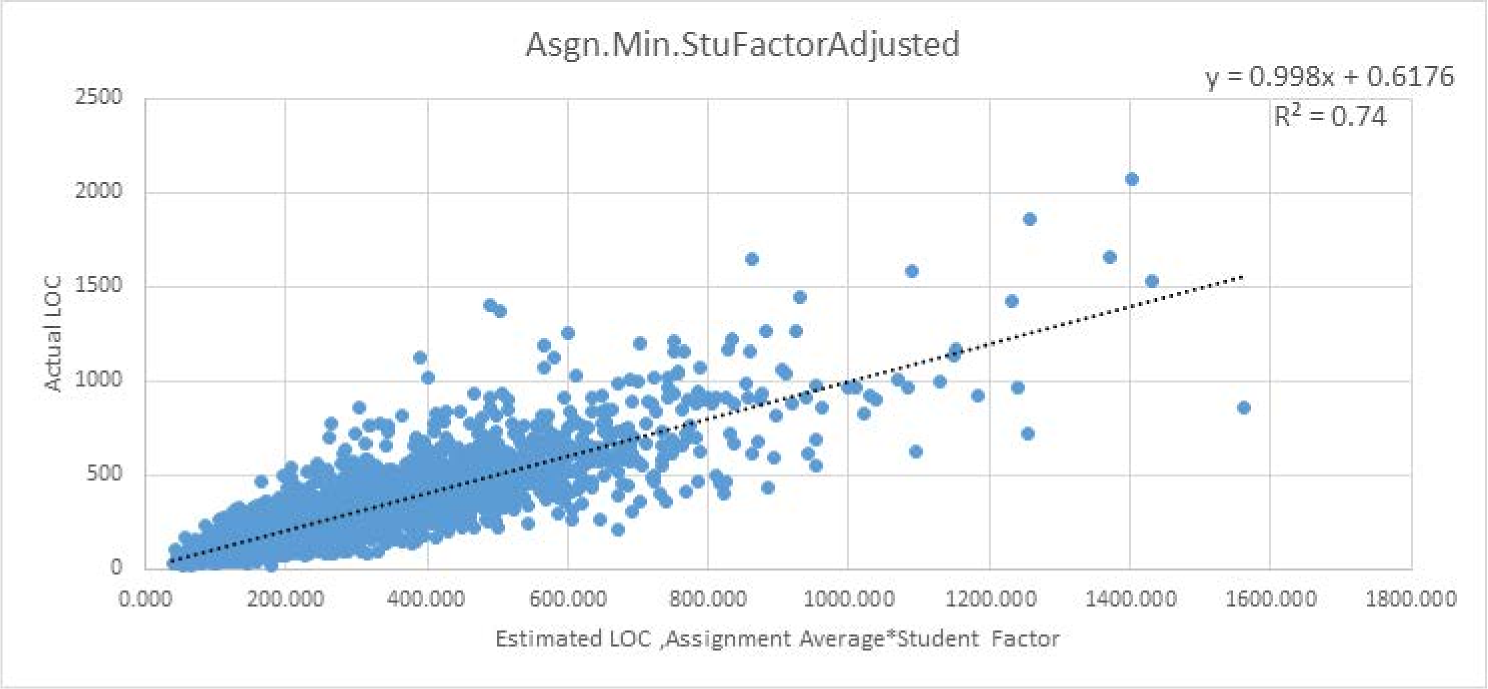







[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

# Student effort factor predicts 74% of variation in program effort

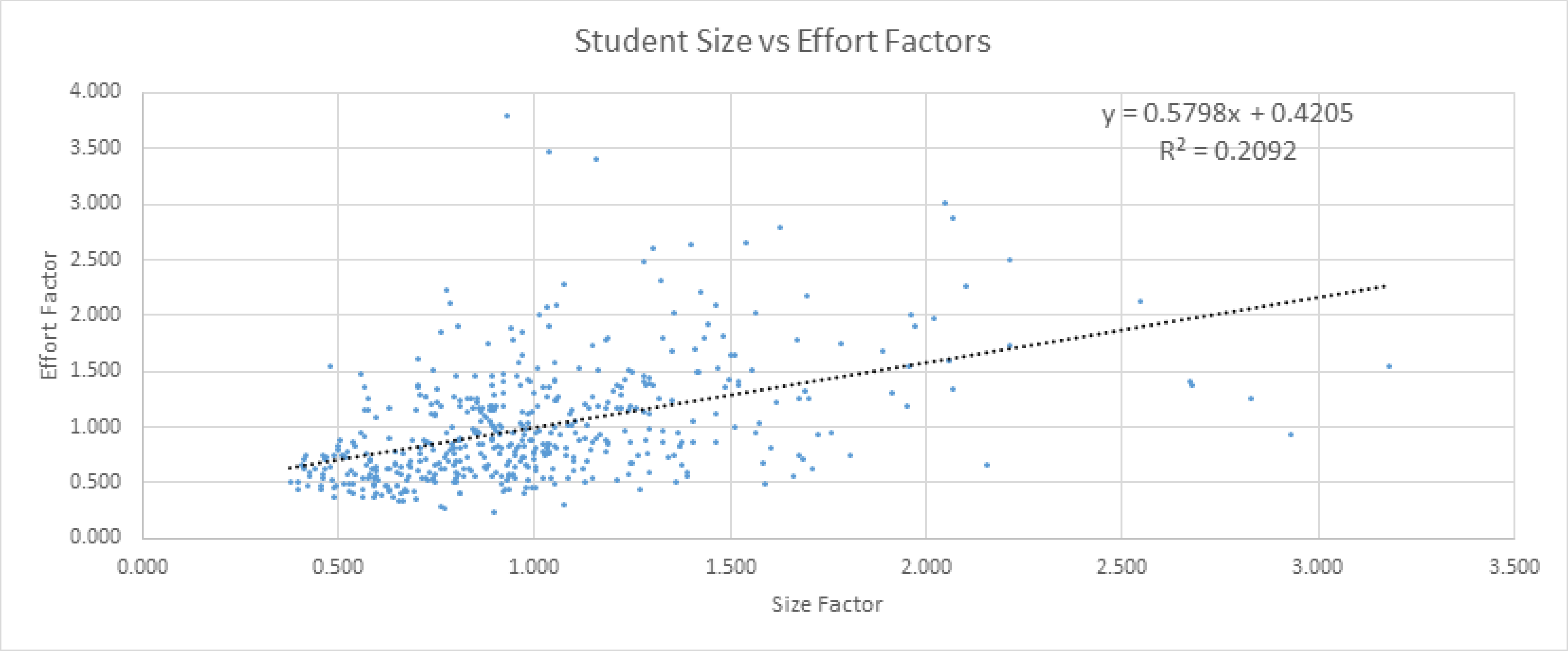




[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

# Effort and size factors do NOT correlate strongly for the overall group.

Size is highly predictive for individual students, but the individual rates vary widely.  Factors are highly local and do not generalize.





[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

# Summary

We have some evidence of causal influence from Design documents and effort

While size appears to be the big factor associated with effort, other factors contribute. (Design, review, programming language, individual developers)


[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

# Next Steps

Analyze other search algorithms and characterize the strength and direction of effects.

Replicate IPA results with separate dataset (TSP)

Combine multiple sets of data in a more complete model


[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.