



**Carnegie Mellon
Software Engineering Institute**

Pittsburgh, PA 15213-3890

Measurement and Analysis: What Can and Does Go Wrong?

Maureen Brown

University of North Carolina Institute

Dennis R. Goldenson

Software Engineering Institute

10th International Symposium on Software Metrics

14 September 2004

Sponsored by the U.S. Department of Defense
© 2004 by Carnegie Mellon University



Acknowledgements

The authors thank Bob Ferguson, Shannon Schelin, Kenny Smith, Mike Zuccher, and Dave Zubrow for their respective contributions. Capability Maturity Model, Capability Maturity Modeling, Carnegie Mellon, CMM, and CMMI are registered in the U.S. Patent and Trademark Office by Carnegie Mellon University. CMM Integration, SCAMPI, SCAMPI Lead Appraiser, and SEI are service marks of Carnegie Mellon University.



Why Care?

Measurement very often is done well & adds value

- Can inform both management & technical decisions
- For both software & software intensive systems

We know this both experientially & evidentially

Yet, too often, measurement...

- is poorly integrated into education & practice
- remains challenging for all too many organizations



This Study

Analysis of findings from Software CMM[®] appraisals

- 1350 findings
- 663 appraisals
- Conducted between 1987 and 2002 inclusive.

Appraisal results augmented by survey of CIO's

- State & local governments
- Private sector

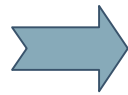
Analyses suggest several areas for better guidance about the use of measurement & analysis, for:

- Managers
- Engineers
- Appraisers



Carnegie Mellon
Software Engineering Institute

Today's Talk



Refresher on CMM models and appraisal methods
Measurement in CMM and CMMI
Appraisal findings
The CIO survey
Summary, conclusions & future research



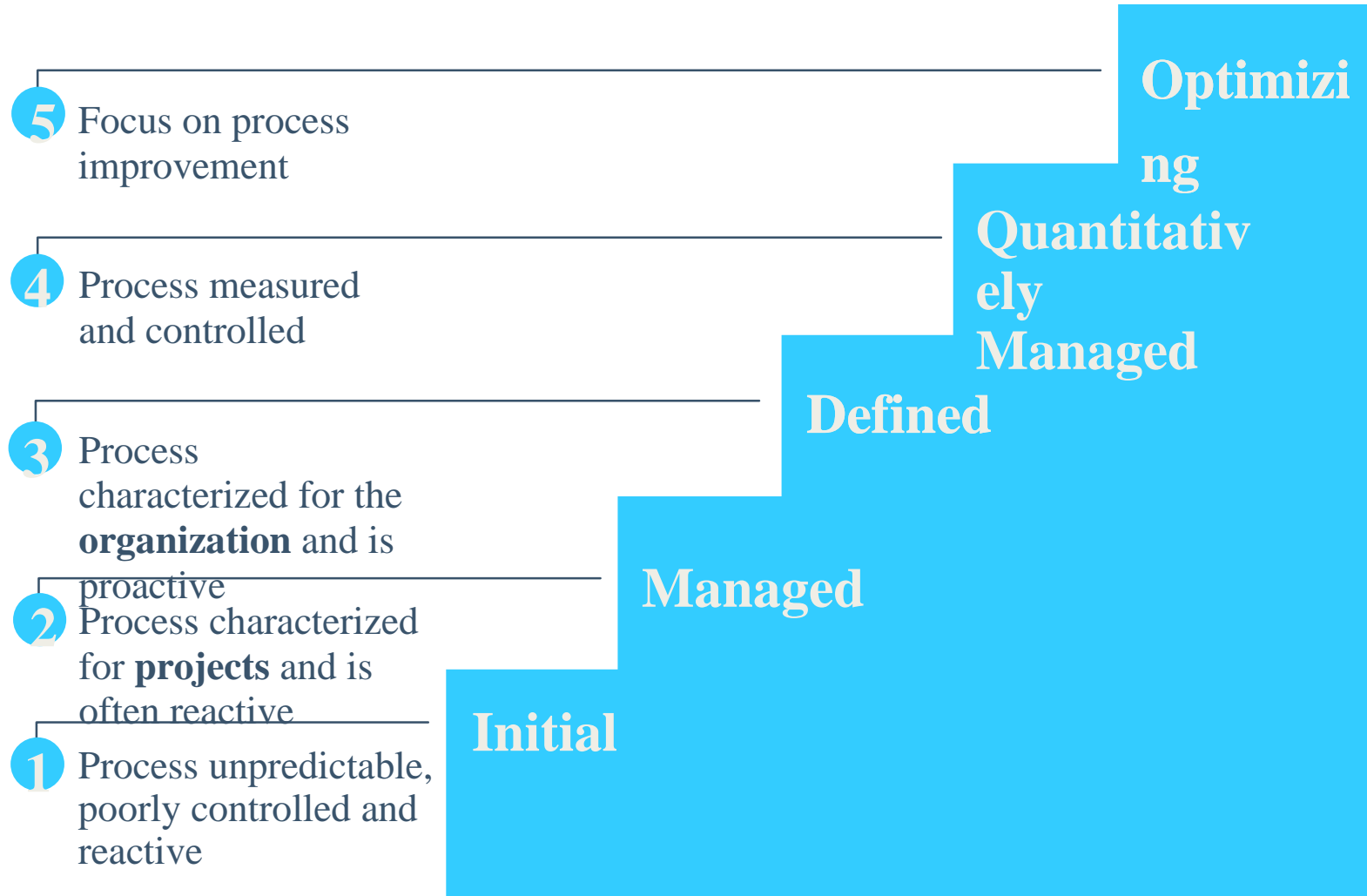


The SW-CMM Key Process Areas

Level	Focus	Key Process Areas
Level 5 Optimizing	Continuous process improvement	<ul style="list-style-type: none"> - Defect Prevention - Technology Change Management
Level 4 Managed	Product & process quality	<ul style="list-style-type: none"> - Process Change Management - Quantitative Process Management - Software Quality Management
Level 3 Defined	Engineering processes & organizational support	<ul style="list-style-type: none"> - Organization Process Focus - Organization Process Definition - Training Program - Integrated Software Management - Software Product Engineering - Intergroup Coordination
Level 2 Repeatable	Project management processes	<ul style="list-style-type: none"> - Requirements Management - Peer Review - Software Project Planning - Software Project Tracking & Oversight - Software Subcontract Management
Level 1 Initial	Competent people (and heroes)	<ul style="list-style-type: none"> - Software Quality Assurance - Software Configuration Management



The CMMI Maturity Levels





Level	Focus	Process Areas	
5 Optimizing	<i>Continuous Process Improvement</i>	Organizational Innovation and Deployment Causal Analysis and Resolution	<p>Quality Productivity</p> <p>Risk Rework</p>
4 Quantitatively Managed	<i>Quantitative Management</i>	Organizational Process Performance Quantitative Project Management	
3 Defined	<i>Process Standardization</i>	Requirements Development Technical Solution Product Integration Verification Validation Organizational Process Focus Organizational Process Definition Organizational Training Integrated Project Management	
2 Managed	<i>Basic Project Management</i>	Risk Management Requirements Management Decision Analysis and Resolution Project Planning Project Monitoring and Control Supplier Agreement Management Measurement and Analysis Process and Product Quality Assurance	
1 Initial		Configuration Management	



CMM Based Process Appraisals

Most widely known for quantitative benchmarks of

- Maturity levels
- KPA & goal satisfaction profiles

Also usually have textual findings

- Meant to provide additional qualitative context & clarification
- Presented verbally in formal presentations to sponsors & other appraisal participants

Verbatim findings typically short enough to fit on overhead slides

- Further clarification commonly provided verbally



Today's Talk

Refresher on CMM models and appraisal methods

Measurement in CMM and CMMI



Appraisal findings

The CIO survey

Summary, conclusions & future research



Measurement in SW-CMM

Explicit and focused guidance?

- Not a strong point
 - Well, it *was* always there ... but in the fine print
 - Well-integrated focus on measurement is noticeably lacking

An early focus?

- Again, not a strong point
- Especially important in a field where measurement isn't widely or well understood
- Do it right in the first place ... or expect rework



Measurement and Analysis in CMMI

Early emphasis introduced at Maturity Level 2

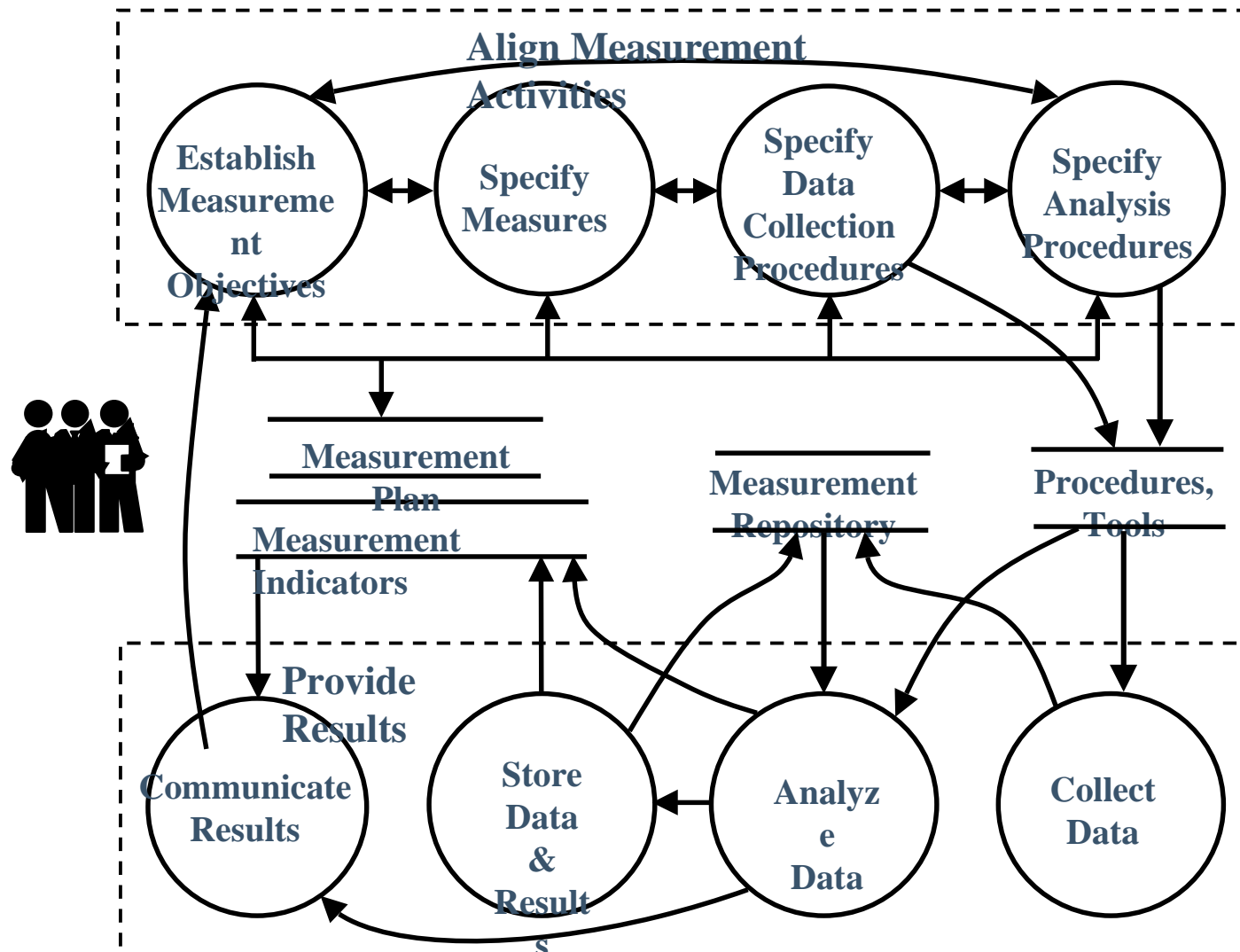
- Measurement and Analysis describes good measurement practice

But it's *not* just the new Process Area

- Maturing measurement processes at higher levels of organizational maturity
- Maturing measurement capability wherever it's applied



The Level 2 Process Area



Measurement Related Generic Practices

Practice	Focus
2.8 Monitor and control the process	Monitor and control the process against the plan for performing the process and take appropriate corrective action
3.2 Collect improvement information	Collect work products, measures, measurement results, and improvement information derived from planning and performing the process to support the future use and improvement of the organization's processes and process assets
4.1 Establish quality objectives	Establish and maintain quantitative objectives for the process about quality and process performance based on customer needs and business objectives
4.2 Stabilize sub-process performance	Stabilize the performance of one or more subprocesses to determine the ability of the process to achieve the established quantitative quality and process performance objectives
5.1 Ensure continuous process improvement	Ensure continuous improvement of the process in fulfilling the relevant business objectives of the organization
5.2 Correct common cause of problems	Identify and correct the root causes of defects and other problems in the process



Other Process Areas with Heavy Measurement Content

Organizational Process Definition

Decision Analysis & Resolution

Organizational Process Performance

Quantitative Process Management

Causal Analysis & Resolution

Organizational Innovation Deployment

Any process area that references Measurement and Analysis



What Typically Gets Measured?

Heavily influenced by SW-CMM

CMM models focus first on project planning & management

- Estimation (not always so well done)
- Monitoring & controlling schedule & budget

Followed by engineering

- Of course, some do focus on defects early ...



Measurement in High Maturity Organizations

By definition...

- Attention to organizational issues
- Bringing processes under management control
- Attention to process models
- Causal analysis & proactive piloting

At ML 3

- Focus on organizational definitions & a common repository

At ML 4

- Improve process adherence

(Especially at) ML 5

- Enhance & improve the processes themselves



How Well Do They Do It?

Well, it depends

Classes (if not nuances) of problems persist

- Even as organizational maturity increases

E.g., what about enterprise measures?

- How do you roll up measures from projects to enterprise relevance?
 - Asked by sponsor at a (deservedly) ML 5 organization
- Remains a pertinent, and difficult, issue for us as measurement experts today



Today's Talk

Refresher on CMM models and appraisal methods

Measurement in CMM and CMMI

Appraisal findings



The CIO survey

Summary, conclusions & future research



Findings Analyzed₁

Analysis limited to SW-CMM appraisal findings

- Many more SW-CMM than CMMI appraisals reported at time of study
- Treatment of measurement more explicit in CMMI

Exclusion of early CMMI appraisals

- Avoids confounding current results
- Allows better subsequent evaluation of effects of changes to models & appraisals



Findings Analyzed₂

Findings from

- CMM-Based Appraisals for Internal Process Improvement (CBA IPI)
- Software Process Assessments (SPA) replaced by CBA IPI in 1996

Data drawn from Process Appraisal Information System (PAIS)

- Contains all appraisal results submitted in confidence to SEI
- Part of authorized lead appraiser program

Findings from 2910 CBA IPI and SPA appraisals of SW-CMM

- Conducted from 19 February 1987 through 28 June 2003
- Total of 36,316 findings recorded as “weaknesses” or “opportunities for improvement”

663 appraisals in the same time period

- With 1350 weaknesses & opportunities for improvement
- That include the root word “measure.”



Typical Measurement Related Findings

Lack of a consistent approach for capturing quality and productivity measurement data and comparing actuals with forecasts

There is no common understanding, definition and measurement of Quality Assurance

Test coverage data is inconsistently measured and recorded

Measurements of the effectiveness and efficiency of project management activities are seldom made



Classifying the Findings

Initially had 48 categories of measurement related findings, based on:

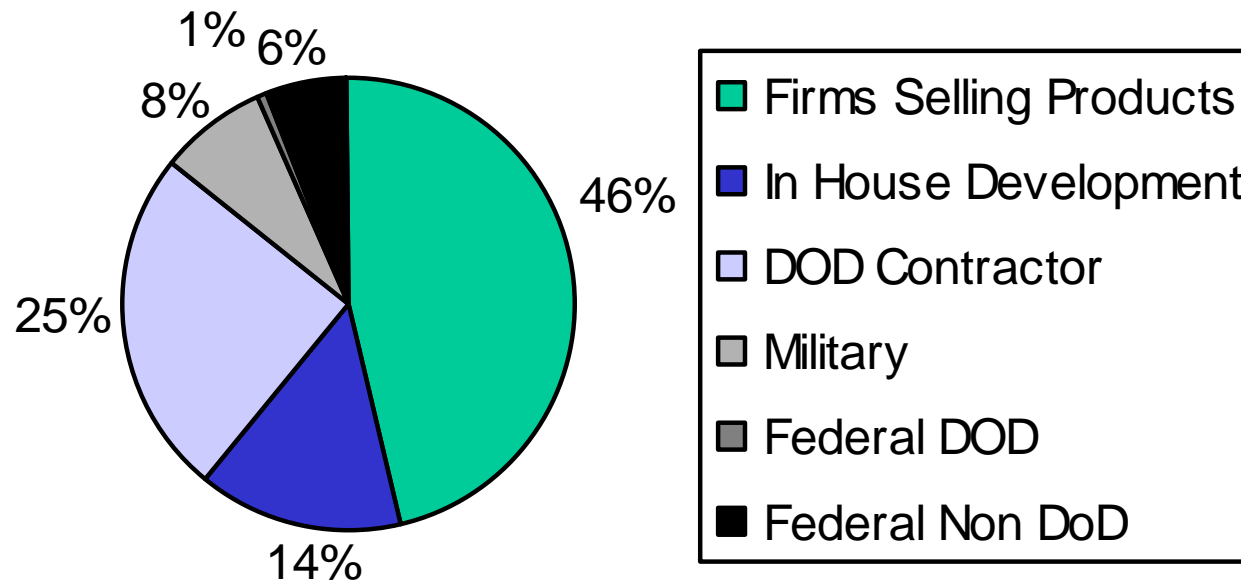
- Measurement categories from Practical Software and Systems Measurement (PSM) performance model
- & a few additional categories to accommodate findings related more directly to structure of SW-CMM

Some findings are classified into more than one of the 48 categories

- For a total of 1,549 coded findings

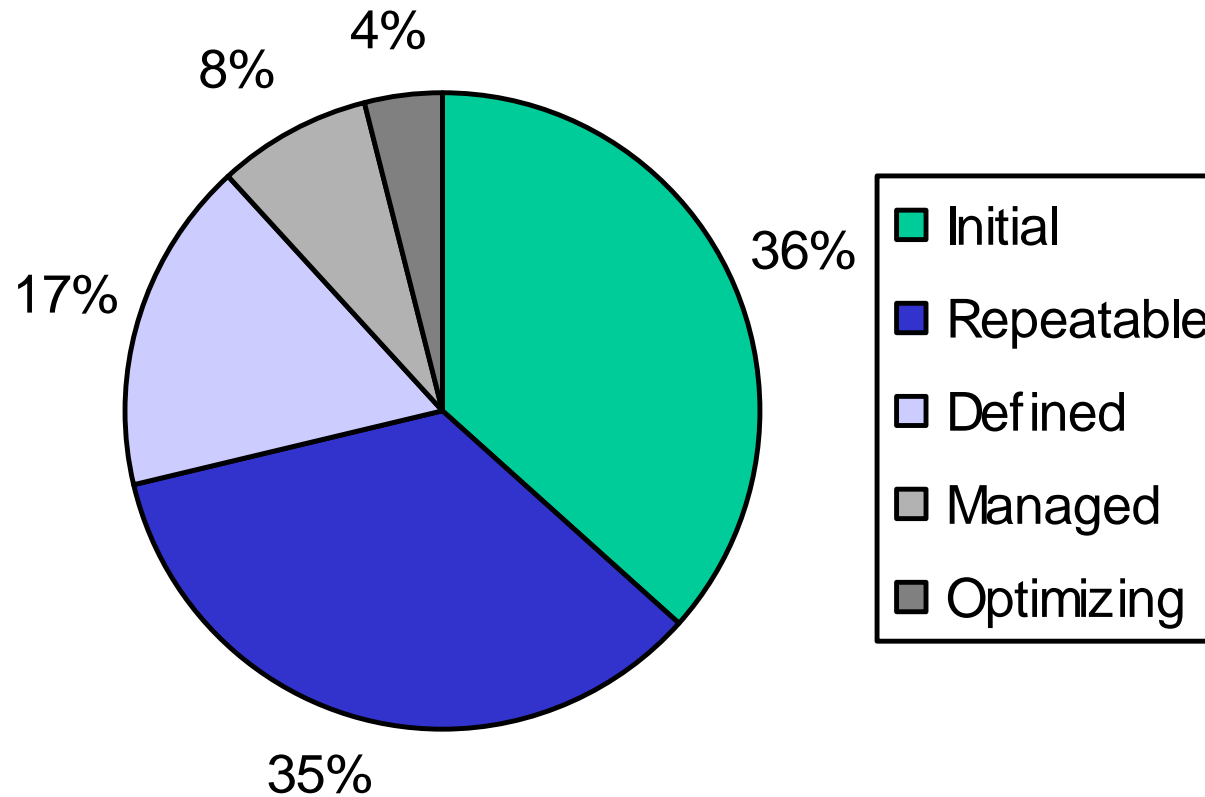


Sector of Organizations Appraised



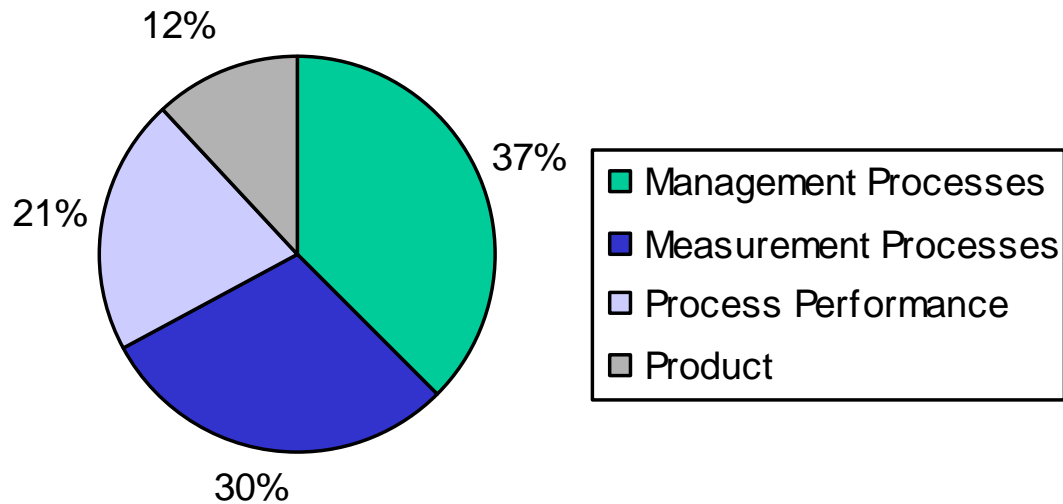


Maturity Levels





Grouped Measurement Findings



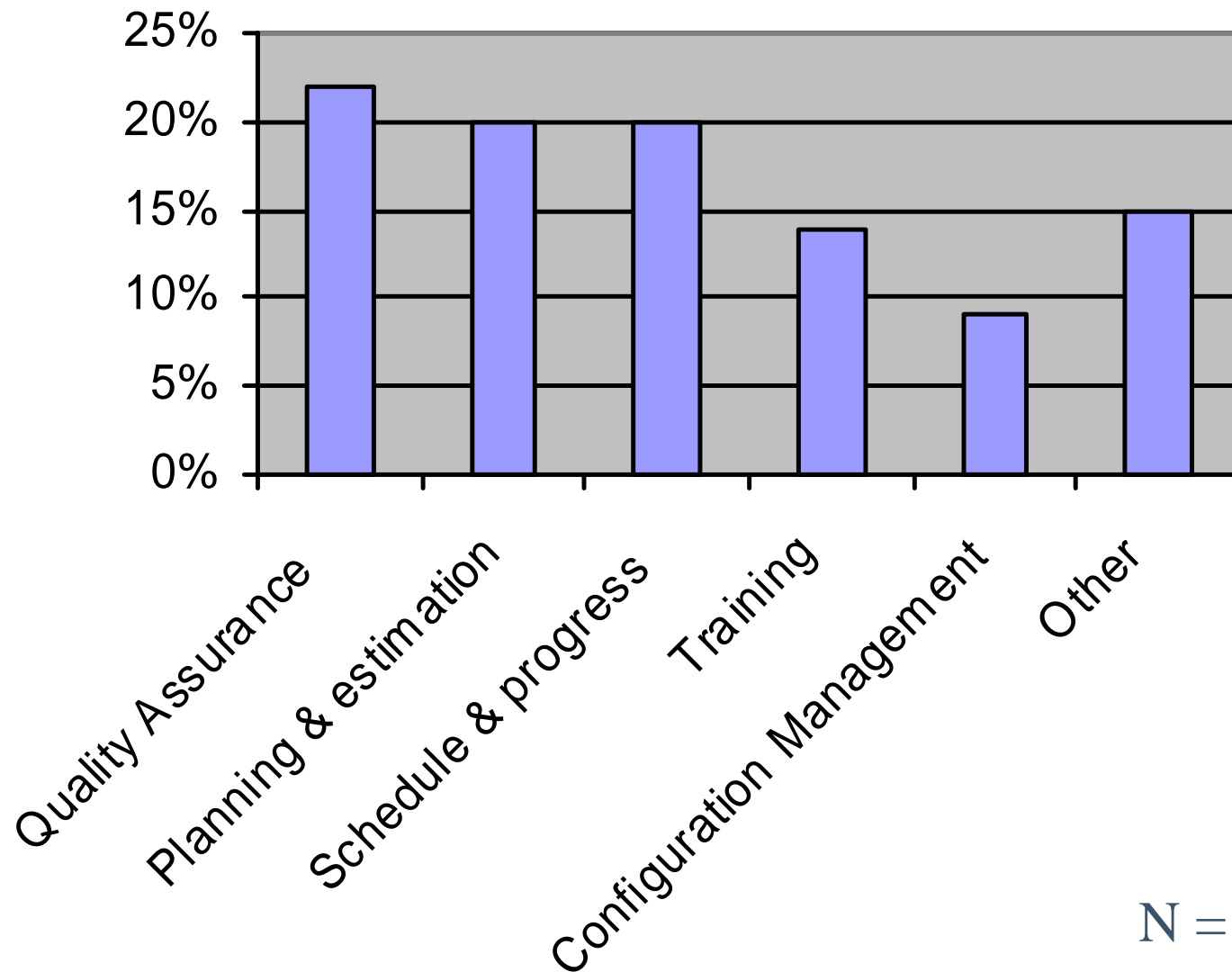
Appraisal findings typically arranged by KPA or other CMM model content

Not surprisingly: Largest of four groups addresses management

- Difficulties with, or lack of use, of measurement for management purposes



Measurement of Management Processes



N = 582 coded findings



Detail: Management

“Other” includes:

- Project management without further elaboration (38 instances)
- Resources & cost (28)
- Policies (14)
- Risk (7)
- ROI concerns (2)

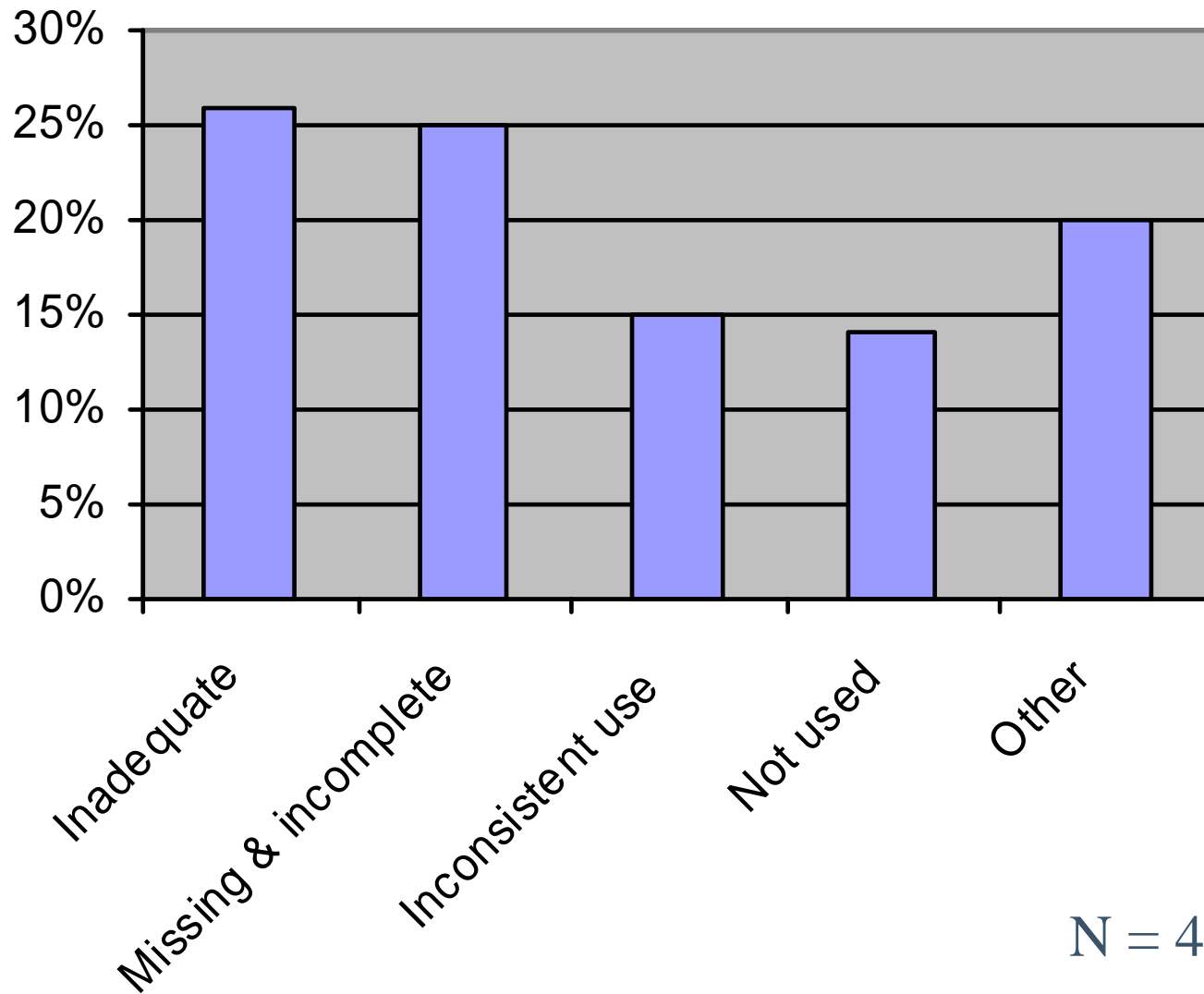
All six categories closely coupled to structure & content of SW-CMM

- First five categories map directly to model KPA structure

With possible exception of two references to measuring ROI, findings in “other” category map KPA’s or institutionalization common features



Measurement Processes Themselves



N = 461 coded findings



Detail: Measurement

Measurement findings particularly noteworthy

- Appraisers tend to focus on model structure & content
- Measurement related content in SW-CMM considerably less explicit & complete than CMMI®

26%: Existing measures inadequate for intended purposes

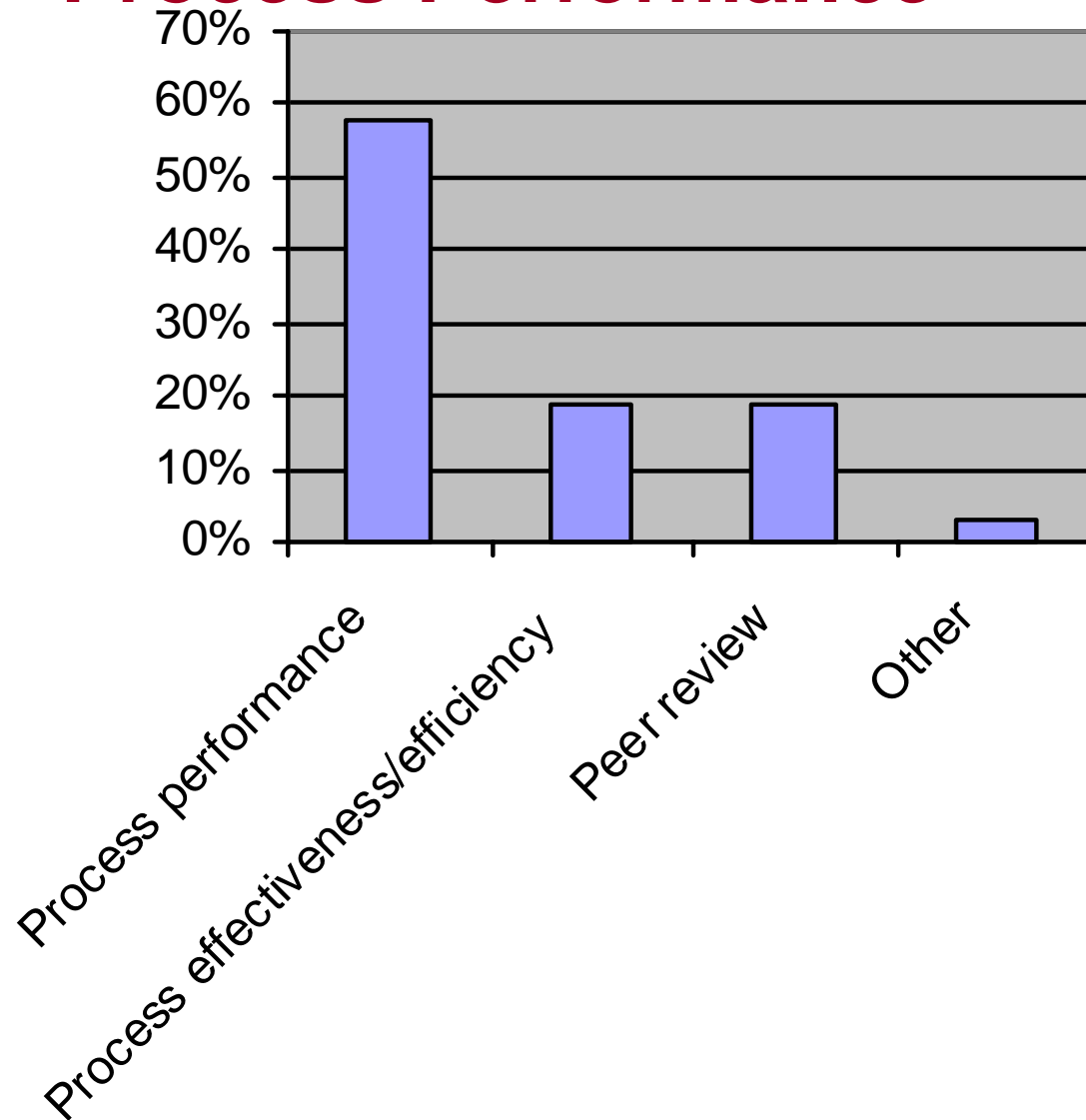
- Findings are terse, but...
- Many or most seem to say measurement is poorly aligned with business & technical needs

“Other” category includes:

- Improvement of measurement processes (43 instances)
- Inter group activities related to measurement (34)
- Measurements misunderstood / not understood (12)
- Leadership in the organization (3)



Process Performance



N = 319 coded findings



Detail: Process Performance₁

Findings describe problems with using measurement to understand & improve existing processes

Well over half of mention difficulties with measuring process performance in explicit terms

- Particularly noteworthy since measurement of process performance is often associated only with high maturity practices

19% refer to problems with measurement & analysis of process effectiveness or efficiency

19% refer to peer reviews

- A ML 3 KPA, but
- Similar issues raised in lower maturity organizations
 - Often re Software Project Tracking and Oversight
 - Also institutionalization common features, particularly Measurement and Analysis.



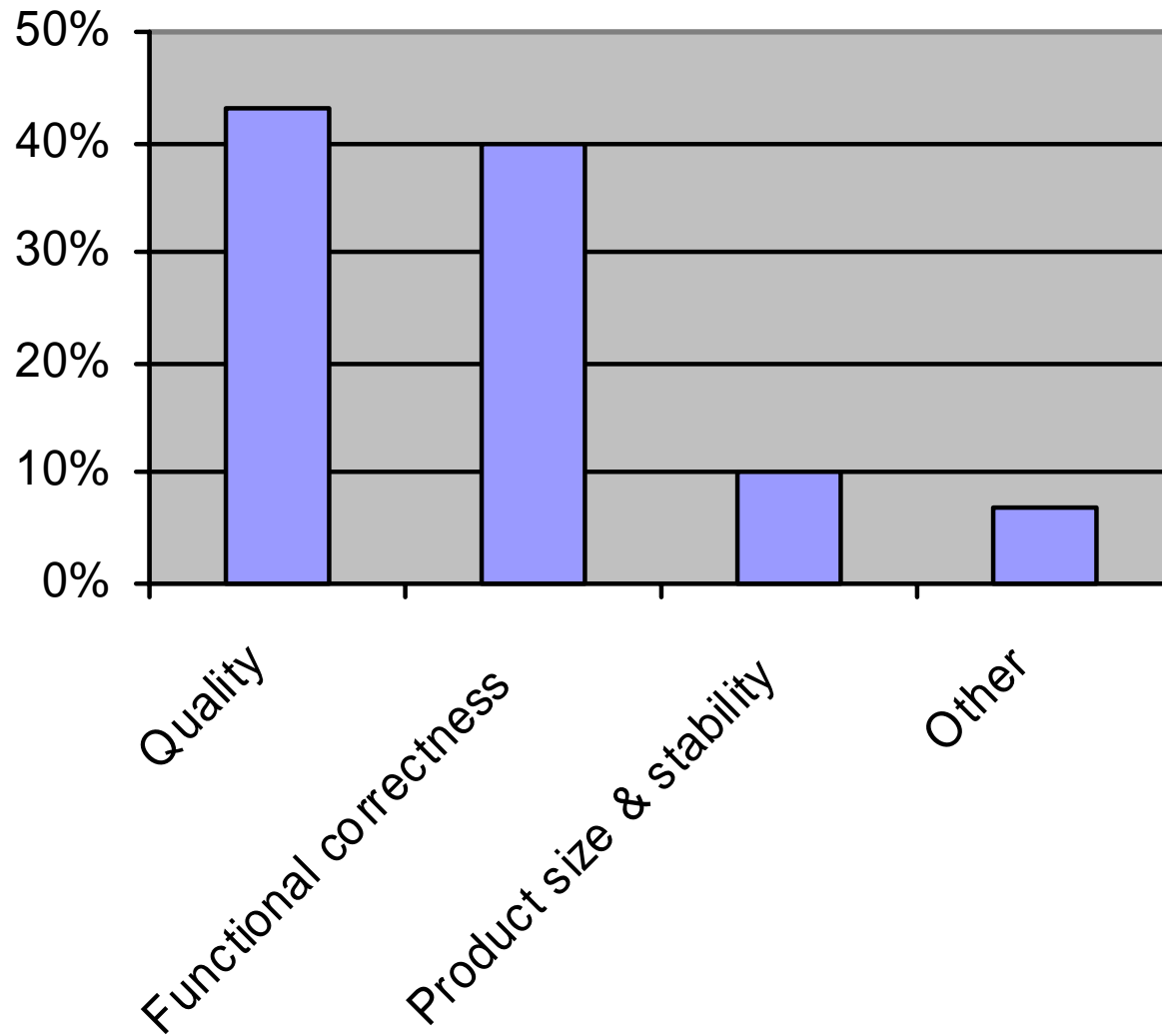
Detail: Process Performance₂

“Other” includes:

- Process compliance (5 instances)
- Tool shortage (4)
- Incremental capability (1)
- Personnel (1)



Product Quality & Technical Effectiveness



N = 187 coded findings



Detail: Product Quality & Technical Effectiveness

Does this mean that the appraised organizations had little difficulty measuring these attributes? And/or:

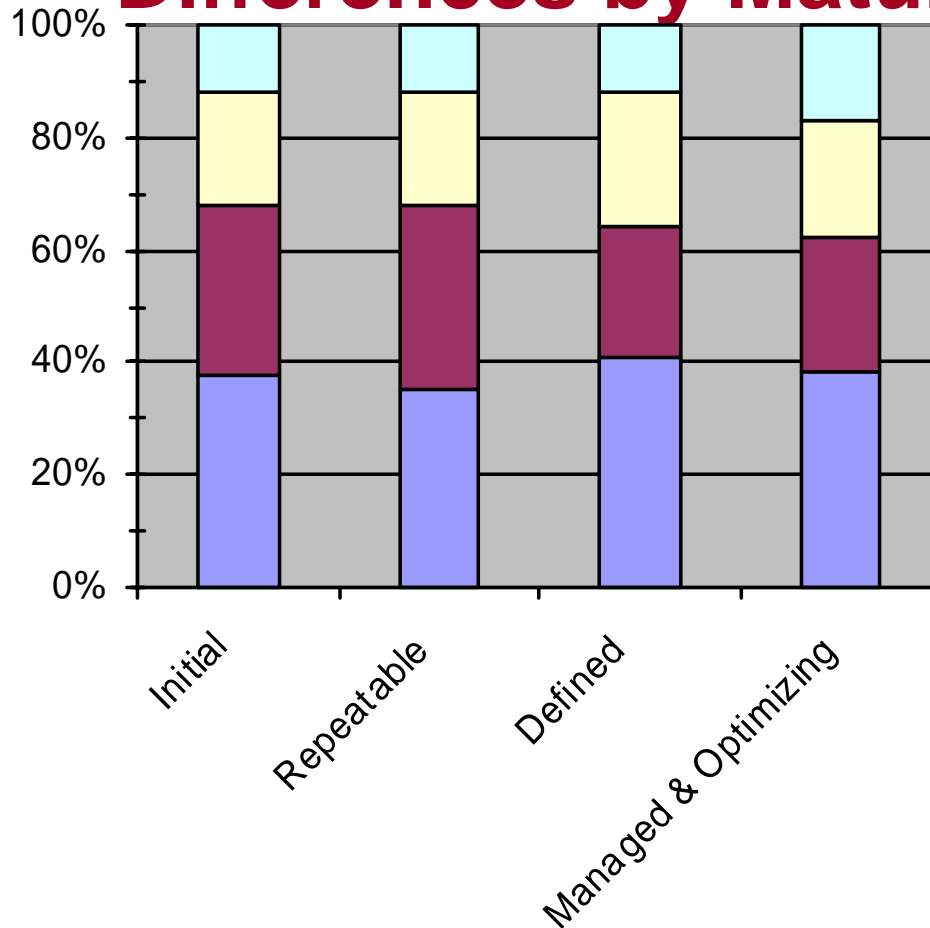
- Did they fail to try?
- Did the appraisal ignore such issues?

“Other” includes:

- Customer satisfaction (4 instances)
- Technical effectiveness (4)
- Reliability (3)
- Security (1)
- Supportability (1)
- Usability (1)
- Technical volatility (1)



Differences by Maturity Level?



All four groups remain problematic throughout

- Including the measurement process itself
 - Nature of difficulties may differ
 - But proper enactment & institutionalization remains a problem for higher maturity organizations
- Similar pattern for process performance
 - Particularly pertinent at maturity levels 4 and 5
 - But noticeable proportions also address similar issues in lower maturity organizations



Today's Talk

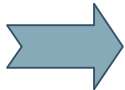
Refresher on CMM models and appraisal methods

Measurement in CMM and CMMI

Appraisal findings

The CIO survey

Summary, conclusions & future research





The CIO survey

Done as part of a UNC doctoral thesis

- Includes a short series of questions about difficulties encountered in implementing software measurement
- Administered to 174 public & private sector CIO's in January 2004

Public sector sample drawn from:

- National Association of State Chief Information Officers (NASCIO)
- International City/County Management Association (ICMA)
- US CIO Council
- N = 83
- 40% response rate

Random sample of 200 private sector Chief Information Officers

- Drawn from the Leadership Library database
- N = 95
- 51% response rate



CIO Questions

Extent to which organizations rely on measurement to guide their system development and maintenance efforts

Level of difficulty encountered in establishing & using a series of measurement classes required by the Clinger-Cohen Act

Answers characterized on a scale of 1 to 10

- Where 10 indicates highest reliance or difficulty respectively



CIO Survey Results

CIO's differ in reliance on measurement

- One fourth have a high degree of reliance on measurement (scores of 8 through 10 on the 10 point scale)
- 39% medium reliance (4 through 7)
- 36% low reliance (1 through 3)

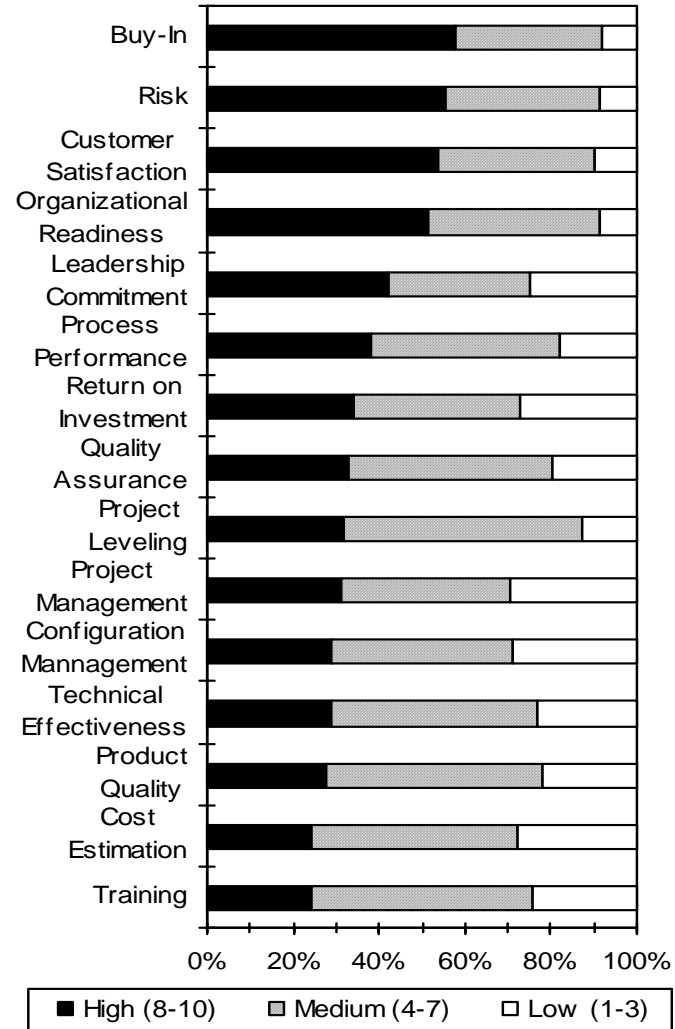
Difficulty encountered establishing & using measurement, particularly:

- Tracking buy-in
- Risk
- Customer satisfaction
- Organizational readiness
- Leadership commitment
- Process performance





An Eye Chart: Difficulty Using Measurement





CIO Comparisons by Sector

Perhaps not surprisingly

- Public sector CIO's reported lower reliance on measurement than did private sector
- $p < .01$

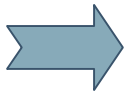
Public sector CIO's also reported greater difficulty in establishing measures for:

- Cost estimation
- Quality assurance
- Project management
- Product quality
- Technical effectiveness
- $p < .01$



Today's Talk

Refresher on CMM models and appraisal methods
Measurement in CMM and CMMI
Appraisal findings
The CIO survey
Summary, conclusions & future research





Issues & Interpretation₁



Questions asked of CIO differ from findings categories

- Many of survey categories refer to product or technology being developed
- As well as or instead of project or organizational processes *per se*
 - e.g., buy-in, organizational readiness, and leadership commitment
 - “Project leveling” refers to the existence of technologies and/or processes shared across projects



Issues & Interpretation₂

Areas that give the CIO's most difficulty not mirrored by appraisal findings

- May mean the two sets of organizations differ in difficulties faced
- But also may be a function of data collection & analysis methods
- Survey asks explicitly about topics not comparably covered by CMM based process appraisals

Derived survey measures combining responses from similar questions comparably to appraisal findings may yield more similar results

- E.g., tuples of survey replies about training, cost estimation, configuration management, project management, or quality assurance



Measurement Guidance₁

Findings results suggest several areas where better guidance is necessary for:

- Appraisers & appraisal teams
- Software and systems engineering organizations
- Improving work processes, delivered products & services

Large number of findings re inadequacies in measurement processes is particularly noteworthy

- As are problems with measurement of product characteristics

Relative similarities in appraisal findings across maturity levels suggest need to improve guidance throughout

- For managers, engineers & appraisers
- Perhaps particularly re weaknesses in using measurement to monitor & improve process performance



Measurement Guidance₂

CIO survey provides complementary results

Difficulties reported by the CIO's differ at first glance from appraisal findings results

- They also highlight the fact that any results are dependent on method & question context

Survey found notable difficulty in implementing measurement in all of the areas about which it queried

- Including areas similar to the appraisal findings

However, survey also identified problem areas not typically emphasized in process appraisals



Future Work₁

Could / should include:

Breakdowns of appraisal findings results by:

- Patterns within organizations
- Model structure
- Non model content

Breakdowns by PAIS finding tags

- Typically tagged by KPA but also by common feature and other general issues

Recoding of appraisal findings according to different categories

- Perhaps more tightly coupled with the CMMI Measurement and Analysis process area

Tests of inter-coder reliability

Further analysis of the CIO survey data.



Future Work₂

Also could / should include:

Lexical analyses based on natural language processing
Additional studies of appraiser & practitioner understanding of measurement content of CMMI

- Lexical analyses of qualitative survey data from both practitioners and appraisers

Analyses of CMMI appraisal findings

- Including synonyms in addition to “measure”

Analyses of appraisal findings of organizational strengths

- Many appear to be boilerplate restatements of model content
- Still, there are slightly more findings of strengths than weaknesses of all kinds, including those related to measurement



Other Implications

Appraisal findings

- Ought PAIS reporting procedures capture fuller information about finding content and context?
 - Any such work should begin as a research activity
 - With proper expectations & incentives for the appraiser corps
- Additional research could be done on using appraisals findings to guide process improvement

Inadequate measurement processes & product quality are found relatively often

- Appraisers often have a good appreciation about what can go wrong in the way organizations handle, or don't handle, measurement
- Still, the problem may be more widespread

Guidance can come in many forms

- Interpretive documents to augment the CMMI product suite
- Future model revisions
- Tutorials and courses



Carnegie Mellon
Software Engineering Institute

Contact Information

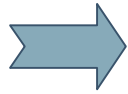
Dennis R. Goldenson
Software Engineering Institute
Pittsburgh, PA 15213-3890
1.412.268.8506
dg@sei.cmu.edu

Maureen Brown
University of North Carolina
Chapel Hill, NC 27599-3330
1.919.966.4347
brown@unc.edu



Carnegie Mellon
Software Engineering Institute

Back Pocket



How can we expedite things?
Maturing measurement capabilities



How Can We Do Better?

Measurement and Analysis is at CMMI Maturity Level 2

- Put there to get it right from the start
- Lots of favorable anecdotes, but...
 - Intent not yet well understood by process champions
 - And we still need better (measurement based) evidence

The bulk of the measurement content is at Maturity Level 3 & above ... mostly at levels 4 & 5

Why wait?

- Causal thinking is (or should be) the essence of statistics 101
- The problem is keeping the management commitment in an *ad hoc*, reactive environment
- But, it can be done...



Measurement Done Early and Well

Two examples (reported under non disclosure)

Level 1 organization used Measurement and Analysis:

- Significantly reduced the cost of quality in one year
- Realized an 11 percent increase in productivity, corresponding to \$4.4M in additional value
- 2.5:1 ROI over 1st year, with benefits amortized over less than 6 months

Level 2 organization used Causal Analysis and Resolution:

- 44 percent defect reduction following one causal analysis cycle
- Reduced schedule variance over 20 percent
- \$2.1 Million in savings in hardware engineering processes
- 95 percent on time delivery



Aligning Measurement & Information Needs

CMM based measurement always got done

- However much was required by appraisers...
- But less likely to be used if divorced from the real improvement effort

Organizations still struggle, even at higher Maturity Levels

- Need a marriage of domain, technical & measurement knowledge
- Yet, measurement often assigned to new hires with little deep understanding or background in domain or measurement

How can we do better?

- GQ(I)M when the resources & commitment are there
- Prototype when they aren't ... or maybe always
- May be easier in small settings because of close communications & working relationships



Performance Models

Called out explicitly in CMM and CMMI

- Especially at Maturity Levels 4 & 5
- But, what do they (usually) mean?
 - Often poorly understood
 - Little more than informal causal thinking

We (the measurement mafia) can do better

- In fact, some have done better...
- By applying modeling & simulation models to process improvement
 - Not common, but it has been & is being done
 - 10 years ago, as an integral part of one organization's process definition, implementation & institutionalization
 - The organization is gone now, but that's another (measurement) story



Modeling & Simulation

Analytic method can be applied in many domains

- Estimate when experimentation, trial & error are impractical
- By being explicit about variables & relationships, process definitions, business & technical goals & objectives

Use it to:

- Proactively inform decisions to begin, modify or discontinue a particular improvement or intervention
- By comparing alternatives & alternative scenarios

Of course, there's still a need for measurement...!

- To estimate model parameters based on fact
- To validate and improve the models



What's Next? (Or, what do I think should be next...?)

Can early attention to measurement really expedite organizational maturation?

- That's part of the rationale for Six Sigma too
- But it's not well, or at least widely, understood
 - How can we demonstrate the relationship?
 - What data & research designs do we need?

Cause and effect?

- Do the analyses early and well

Pay more attention to performance measures

- Including enterprise measures
- And including quality attributes beyond defects
(See ISO/IEC Working Group 6, ISO 25000)

And don't ignore (or wait to do) modeling and simulation

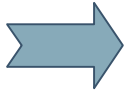


Carnegie Mellon
Software Engineering Institute

Back Pocket

How can we expedite things?

Maturing measurement capabilities





Carnegie Mellon
Software Engineering Institute

From Software Engineering Symposium 2000

Work with Khaled El Emam



The Prescribed Order: Items in Presumptive Maturity Level 2

- Schedule e.g., actual versus planned completion, cycle time (85%)
- Cost/budget e.g., estimate over-runs, earned value (77%)
- Effort e.g., actual versus planned staffing profiles (73%)
- Field defect reports (68%)
- Product size e.g., in lines of code or function points (60%)



The Prescribed Order: Items in Presumptive Maturity Level 3

- Test results or other trouble reports (81%)
- Data, documentation, and reports are saved for future access (76%)
- Organization has common suite of software measurements collected and/or customized for all projects or similar work efforts (67%)
- Results of inspections and reviews (58%)
- Customer or user satisfaction (56%)



The Prescribed Order: Items in Presumptive Maturity Level 4

- **Quality assurance and audit results** (54%)
- **Comparisons regularly made between current project performance and previously established performance baselines and goals** (44%)
- **Requirements stability** e.g., number of customer change requests or clarifications (43%)
- **“Other” quality measures** e.g., maintainability, interoperability, portability, usability, reliability, complexity, reusability, product performance, durability (31%)
- **Process stability** (31%)
- **Sophisticated methods of analyses are used on a regular basis** e.g., statistical process control, simulations, latent defect prediction, or multivariate statistical analysis (14%)
- **Statistical analyses are done to understand the reasons for variations in performance** e.g., variations in cycle time, defect removal efficiency, software reliability, or usability as a function of differences in coverage and efficiency of code reviews, product line, application domain, product size, or complexity (14%)



The Prescribed Order: Items in Presumptive Maturity Level 5

- Experiments and/or pilot studies are done prior to widespread deployment of major additions or changes to development processes and technologies (38%)
- Evaluations are done during and after full-scale deployments of major new or changed development processes and technologies (e.g., in terms of product quality, business value, or return on investment) (27%)
- Changes are made to technologies, business or development processes as a result of our software measurement efforts (20%)



Exceptions

Exceptions

- Level 5
 - Experiments and/or pilot studies (38%)
- Level 4
 - Sophisticated analyses (14%)
 - Statistical analyses of variations (14%)
- Level 3
 - Test results or other trouble reports (81%)
 - Data, documentation, and reports saved (76%)
- Level 2
 - Product size (60%)

May be due to

- Measurement error in this study
- Differences among organizational contexts
- Subtleties in “natural” order



Where Do the Exceptions Occur?

Of the possible comparisons with presumptively lower level items ...

Level 3

- 14% fail level 2 items

Level 4

- 6% fail level 3 items
- 4% fail level 2 items

Level 5

- 14% fail level 4 items
- 6% fail level 3 items
- 6% fail level 2 items