



Software Engineering Institute

The State of Software Measurement Practice: Results of 2006 Survey

Mark Kasunic

December 2006

TECHNICAL REPORT
CMU/SEI-2006-TR-009
ESC-TR-2006-009

Software Engineering Measurement and Analysis
Unlimited distribution subject to the copyright.



CarnegieMellon

This report was prepared for the

SEI Administrative Agent
ESC/XPK
5 Eglin Street
Hanscom AFB, MA 01731-2100

The ideas and findings in this report should not be construed as an official DoD position. It is published in the interest of scientific and technical information exchange.

This work is sponsored by the U.S. Department of Defense. The Software Engineering Institute is a federally funded research and development center sponsored by the U.S. Department of Defense.

Copyright 2006 Carnegie Mellon University.

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.

Use of any trademarks in this report is not intended in any way to infringe on the rights of the trademark holder.

Internal use. Permission to reproduce this document and to prepare derivative works from this document for internal use is granted, provided the copyright and "No Warranty" statements are included with all reproductions and derivative works.

External use. Requests for permission to reproduce this document or prepare derivative works of this document for external and commercial use should be addressed to the SEI Licensing Agent.

This work was created in the performance of Federal Government Contract Number FA8721-05-C-0003 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center. The Government of the United States has a royalty-free government-purpose license to use, duplicate, or disclose the work, in whole or in part and in any manner, and to have or permit others to do so, for government purposes pursuant to the copyright license under the clause at 252.227-7013.

For information about purchasing paper copies of SEI reports, please visit the publications portion of our Web site (<http://www.sei.cmu.edu/publications/pubweb.html>).

Table of Contents

Acknowledgments	vii
Abstract.....	ix
1 Introduction.....	1
2 Conducting the Survey.....	3
2.1 Characteristics of the Survey.....	3
2.2 Characteristics of the Study Population.....	3
2.2.1 Databases Used to Select the Population	3
2.2.2 Sampling Plan.....	4
2.3 Survey Response	4
3 Results	5
3.1 Population Demographics.....	5
3.2 Attitudes and Beliefs About Measurement Use	6
3.2.1 Level of Involvement with Measurement.....	6
3.2.2 Are the Purposes for Measurement Understood?	7
3.2.3 Does Measurement Help Performance?.....	8
3.2.4 Is Measurement Used to Understand Quality?	9
3.2.5 Are Documented Measurement Processes Used?.....	10
3.2.6 Are Measurement Definitions Understood and Consistent?	12
3.2.7 Do Measurable Criteria Exist for Products and Services?	13
3.2.8 Is Corrective Action Taken When Thresholds Are Exceeded?	15
3.3 Measurement Guidance Used	17
3.4 Measures That Are Reported.....	23
4 Summary Observations.....	29
4.1 Responses Patterns for Management Compared to Staff	29
4.2 Impact of the Organization's Size	30
4.3 Response Patterns for Industry Compared to Government.....	31
4.4 Response Pattern for United States Compared to Other Countries	32
4.5 Use of Measurement Information	32
4.6 Methods Used.....	32

4.7 Measures Reported	33
4.8 Summary	33
Appendix A: Survey Questionnaire	35
Appendix B: Sampling Plan.....	43
Appendix C: Survey Response Rates and Drop-Out Rate	49
Appendix D: Detailed Information About Respondents.....	51
References.....	53

List of Figures

Figure 1: Involvement with Measurement.....	6
Figure 2: Understand Purpose of Data Collected	7
Figure 3: Understand Purpose of Data Collected, by Job Type.....	7
Figure 4: Measurement Helps Team Perform Better	8
Figure 5: Measurement Helps Team Perform Better, by Job Type	9
Figure 6: Measurement Used to Understand Quality	9
Figure 7: Measurement Used to Understand Quality, by Job Type	10
Figure 8: Use of Documented Processes for Collecting Data	10
Figure 9: Use of Documented Processes for Collecting Data, by Job Type	11
Figure 10: Documented Process Used to Report Data to Management	11
Figure 11: Documented Process Used to Report Data to Management, by Job Type	12
Figure 12: Measurement Definitions Are Understood and Consistent.....	12
Figure 13: Measurement Definitions Are Understood and Consistent, by Job Type	13
Figure 14: Measurable Criteria Exist for Products and Services	14
Figure 15: Measurable Criteria Exist for Products and Services, by Job Type.....	14
Figure 16: How Often Is Corrective Action Taken?	15
Figure 17: How Often Is Corrective Action Taken? by Job Type.....	16
Figure 18: Action-Oriented Response to Measurement Information	17

Figure 19: Measurement Methods Used	18
Figure 20: Other Measurement Methods Used	19
Figure 21: Respondents Using CMMI M&A Process Area	20
Figure 22: Respondents Using Goal-Driven Method.....	20
Figure 23: Respondents Using GQM	21
Figure 24: Respondents Using PSP/TSP	21
Figure 25: Respondents Using PSM	22
Figure 26: Respondents Using ISO 15939.....	22
Figure 27: Respondents Using “Other” Methods.....	23
Figure 28: Measures That Are and Are Not Reported	24
Figure 29: Frequency of Reporting Capability/Requirements Stability	25
Figure 30: Frequency of Reporting Effort Against Tasks	25
Figure 31: Frequency of Reporting Defects Identified.....	25
Figure 32: Frequency of Reporting Defects Removed	26
Figure 33: Frequency of Reporting Code Growth	26
Figure 34: Frequency of Reporting Schedule Progress	27
Figure 35: Frequency of Reporting Identified Risks	27
Figure 36: Other Measures Reported.....	28
Figure 37: Survey Response Rates	49
Figure 38: Drop-Out Rate by Question.....	50
Figure 39: Types of Organizations for Which Respondents Worked.....	51
Figure 40: Types of Jobs Respondents Had	52
Figure 41: Number of Full-Time Employees.....	52

List of Tables

Table 1:	Sample Size	4
Table 2:	Percentage of Respondents Who Responded “Frequently” to the Listed Questionnaire Items.....	31
Table 3:	Percentage of Government vs. Industry Respondents Who Strongly Agree or Agree to Listed Questionnaire Items	32
Table 4:	Percentage of Respondents Who Strongly Agree or Agree, by Country	32

Acknowledgments

Thanks to my Software Engineering Measurement and Analysis (SEMA) colleagues who contributed their ideas for this survey, including Robert Ferguson, Dennis Goldenson, Dave Zubrow, Wolf Goethert, Jim McCurley, Michael Zuccher, Laura Malone, and Robert Stoddard. This work benefited from their reviews and feedback.

Also, Linda Parker Gates and Erin Harper provided helpful feedback. I thank them for their contributions. Special thanks to Peter Capell for the detailed review and helpful feedback he provided on multiple work products associated with this effort. Thanks to Erin Harper and Suzanne Couturiaux for their excellent editing support.

Connie Sapienza, Jim McCurley, and Mike Zuccher provided invaluable assistance in organizing the database information used in this effort. Thanks to Laura Malone and Michael Zuccher for the extra effort required to implement the survey incentive offer associated with the Software Engineering Information Repository (SEIR).

Thanks to Dave Zubrow for management support of this effort.

Finally, thank you to the individuals who took the time to assist us with this research by responding to our survey.

Abstract

In February 2006, the Software Engineering Measurement and Analysis Initiative at the Carnegie Mellon[®] Software Engineering Institute (SEI) conducted the first in a series of yearly studies to gauge the state of the practice in software measurement. To conduct this study, a structured, self-administered survey consisting of 17 questions was distributed to a random sample of software practitioners who had contacted the SEI during 2004 and 2005.

The results of this study, which are revealed in this technical report, offer these benefits: they can be used to indicate (1) what measurement definition and implementation approaches are being adopted and used by the community, (2) the most prevalent types of measures being used by organizations that develop or acquire software, and (3) what behaviors are preventing the effective use of measurement (so that these barriers can be addressed). In addition, when the studies are conducted on a periodic basis, the results can indicate trends over time.

1 Introduction

The Software Engineering Measurement and Analysis (SEMA) Initiative helps organizations develop and evolve useful measurement and analysis practices. In February 2006, SEMA conducted its first survey to gauge the state of the practice in software measurement.

A structured, self-administered 17-item questionnaire was distributed to a random sample of individuals from a population of 15,180 software practitioners that had contacted the Carnegie Mellon[®] Software Engineering Institute (SEI) during the time period 2004-2005.

When reviewing the survey results in this report, please keep in mind that the population used for this study included only those individuals who had contacted the SEI. It is possible that the results were biased by the respondents' knowledge of the SEI or experience with the SEI's products and services.

The objectives of this survey were to characterize

- the degree to which software practitioners use measurement when conducting their work
- the perceived value of measurement
- approaches that are used to guide how measures are defined and used
- the most common types of measures used by software practitioners

This work provides several benefits. The results of the survey

- provide an indication of the measurement implementation approaches being used by the community
- point to behaviors that are preventing the effective use of measurement—thereby allowing these barriers to be addressed
- provide an indication of how well measurement practices are being transitioned into use by the community

Finally, if surveys are conducted on a periodic basis, the results of multiple surveys can indicate trends, such as the rate at which effective measurement practices are taking root.

[®] Carnegie Mellon is registered in the U.S. Patent and Trademark Office by Carnegie Mellon University.

2 Conducting the Survey

This section describes the survey design and the population that was studied. Information about the survey response is also provided.

2.1 Characteristics of the Survey

A structured, self-administered questionnaire was created and made available on the World Wide Web and in paper form. The questionnaire was designed to be short (17 questions) and easy to complete with questions phrased in closed-ended format. Several questions allowed for short, open-ended responses. The questionnaire is shown in Appendix A.

The methodology used for the design, development, and implementation of this survey is documented in the SEI guidebook titled *Designing an Effective Survey* [Kasunic 05]. Candidate respondents were offered the following incentives to participate:

- platinum membership to the Software Engineering Information Repository (SEIR) that provides access to documents otherwise unavailable through regular membership
- early access to the survey results

2.2 Characteristics of the Study Population

2.2.1 Databases Used to Select the Population

The surveyed population consisted of software practitioners who were entered into one of three SEI databases during the years 2004 and 2005. The databases that were used contained entries for individuals who

- contacted the SEI and were entered into the customer relations database during 2004-2005
- registered to gain access to the SEI's Software Engineering Repository during 2004-2005
- became an SEI Member during 2004-2005

When interpreting the results of this survey, it is important to remember that the survey results cannot be generalized beyond the population used in this study without risk.

2.2.2 Sampling Plan

Stratified sampling was conducted by viewing the population as three separate subpopulations because we wanted to learn whether there would be differences in response rates for the different subpopulations and whether there were significant differences in response profiles.

Respondents were selected by random sampling. The sample size was specified for a margin of error of $\pm 2.5\%$ with 95% confidence. See Table 1 for an explanation of the sample sizes used. The *adjusted calculated sample size* column shows an adjustment to the calculated sample size based on an estimated 30% response outcome. The *actual sample size* column shows the actual sample size that was realized due to invalid email addresses and ineligible respondents that we learned of after the survey invitation was distributed. A detailed sampling plan is provided in Appendix B.

Table 1: Sample Size

Subpopulation	Population Size	Calculated Sample Size	Adjusted Calculated Sample Size	Actual Sample Size
Customer Relations	6,398	603	2,010	1,670
SEI Members	1,242	434	1,242	951
SEIR Registrants	7,540	612	2,040	1,539
Total	15,180	1,649	5,292	4,160

2.3 Survey Response

There were 2143 survey responses. An initial survey question screened out those who do not work for or support a software organization. Based on this question, 183 respondents were screened from further participation in the survey.

The response outcome was 2,109 respondents (50.7%). This percentage includes individuals who completed all or a portion of the questionnaire (including the 183 respondents that were screened from further participation because they do not work for or support a software organization). The entire questionnaire was completed by 1,764 respondents (42.4%). This outcome was calculated using the standards outlined in *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys* [AAPOR 00].

There were only slight differences in outcome between the subpopulations. A graph showing the response rate of each subgroup and their drop-out rate over the course of the questionnaire is available in Appendix C. Also, there were no significant differences in the response profiles to the questionnaire based on subpopulation. Therefore, the results are reported for the combined subpopulations.

3 Results

This section summarizes the responses to the survey questions and identifies some response patterns. Response patterns are reported for the entire sample, and, in many cases, patterns are presented based on the organizational role of the respondent. The section contains the following information:

- demographic information about the people who participated in the survey (Section 3.1)
- participants' responses to questions regarding their attitudes and beliefs about measurement use (Section 3.2)
- which methods participants reported using to identify, collect, and analyze measurement data (Section 3.3)
- the participants' responses when asked to indicate the measures they report and how often they are reported (Section 3.4)

Key observations are discussed in Section 4. Additional cross-tabulations of the data were performed to determine whether there were differences in response to the questions based on the respondents'

- organization size
- type of organization (Department of Defense [DoD] and government versus commercial)
- country affiliation (United States versus other countries)

Any significant differences based on these demographic categories are also reported in Section 4, Summary Observations.

3.1 Population Demographics

Individuals from 84 countries responded, with 53.1% of respondents from the United States, 11.9% from Europe, 8.4% from India, and 1.7% from China. The sample included individuals from all organizational levels, including executive management, project management, and staff. Those employed at commercial enterprises made up 72.6% of the respondents, while 27.4% were from the DoD and other government agencies. Graphs showing detailed information about the type of organizations represented, the types of jobs respondents had, and the size of the respondents' companies are available in Appendix D.

3.2 Attitudes and Beliefs About Measurement Use

Eight of the survey questions were designed to gauge the respondents' level of involvement with measurement and their perception of the value of measurement activities. The responses to these questions are shown in Sections 3.2.1- 3.2.8.

3.2.1 Level of Involvement with Measurement

Respondents were asked to describe their involvement with measurement by selecting a single response from the following options:

- provider of measurement-based information
- user (consumer) of measurement-based information
- both a provider and user (consumer) of measurement-based information
- other (please specify)

The responses to this question are shown in Figure 1.

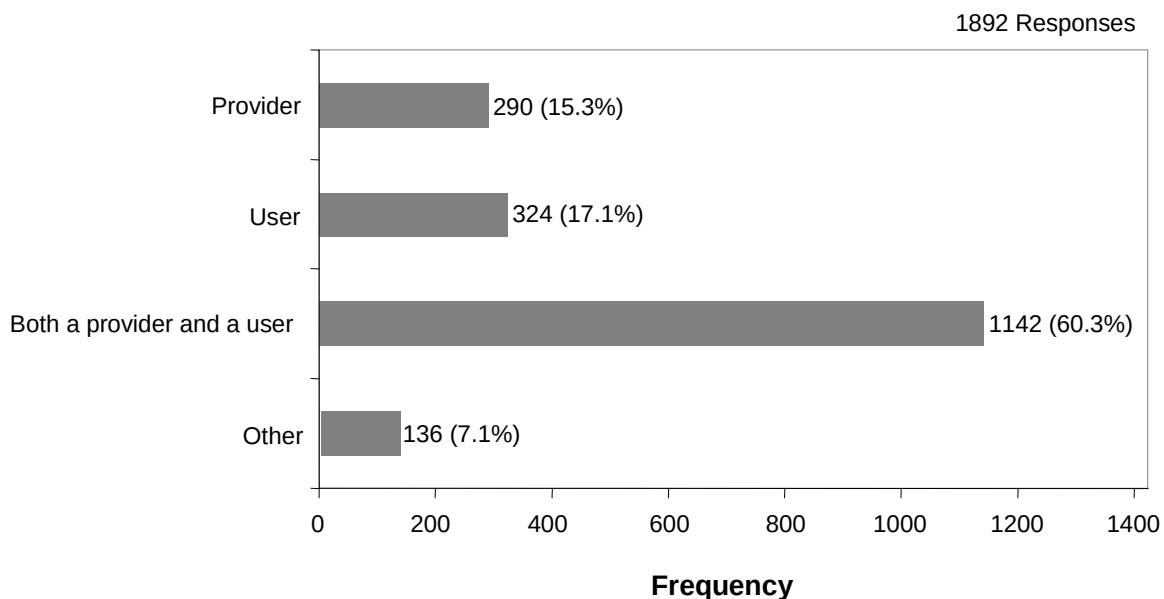


Figure 1: Involvement with Measurement

Most respondents reported that they both provided and used measurement information. Of those who identified themselves only as users, most were in management roles, while those who identified themselves only as providers were primarily programmers, analysts, and engineers. Respondents who identified themselves as “other” were individuals who commented that they did not use measurement or that they were responsible for implementing measurement programs in their organizations.

3.2.2 Are the Purposes for Measurement Understood?

Respondents were asked how often they understood the purpose for the data they collected or reported. Their responses are shown in Figure 2.

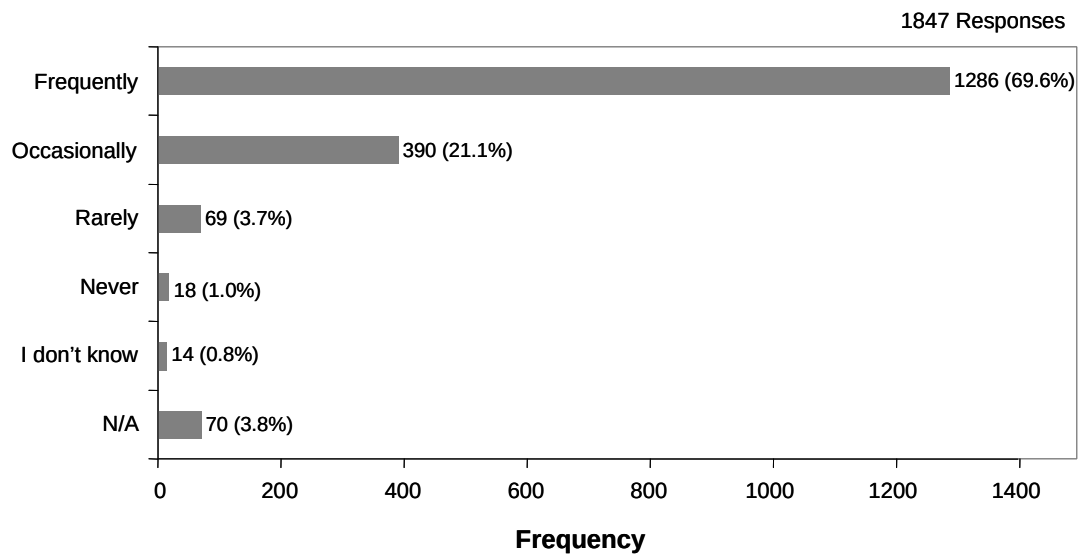


Figure 2: Understand Purpose of Data Collected

As shown in Figure 3, there were statistically significant differences (95% confidence) between staff and management responses to this question. While approximately 75% of the program managers and executives reported that they frequently understood why data was being collected, only 52% of the programmers and 60% of the engineers chose that response.

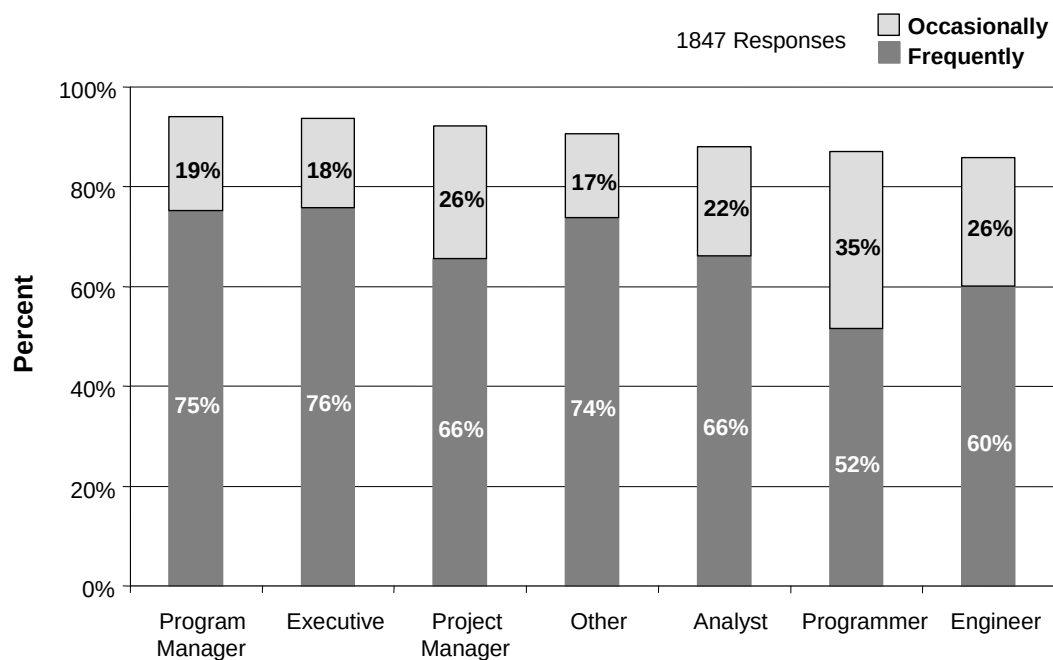


Figure 3: Understand Purpose of Data Collected, by Job Type

When comparing responses between the defense industry and the commercial industry, there was a statistically significant difference in response (with confidence of 95%): 80.0% of individuals from government agencies indicated that they understood the purpose of the data they collect, while 84.5% of individuals from commercial industries indicated that they understood the data's purpose.

3.2.3 Does Measurement Help Performance?

To determine whether respondents thought measurement helped their team's performance, they were asked to rate their level of agreement with the following statement: *Generally speaking, I believe that using measurement-based data helps my team to perform better than without using it.* Their responses are shown in Figure 4.

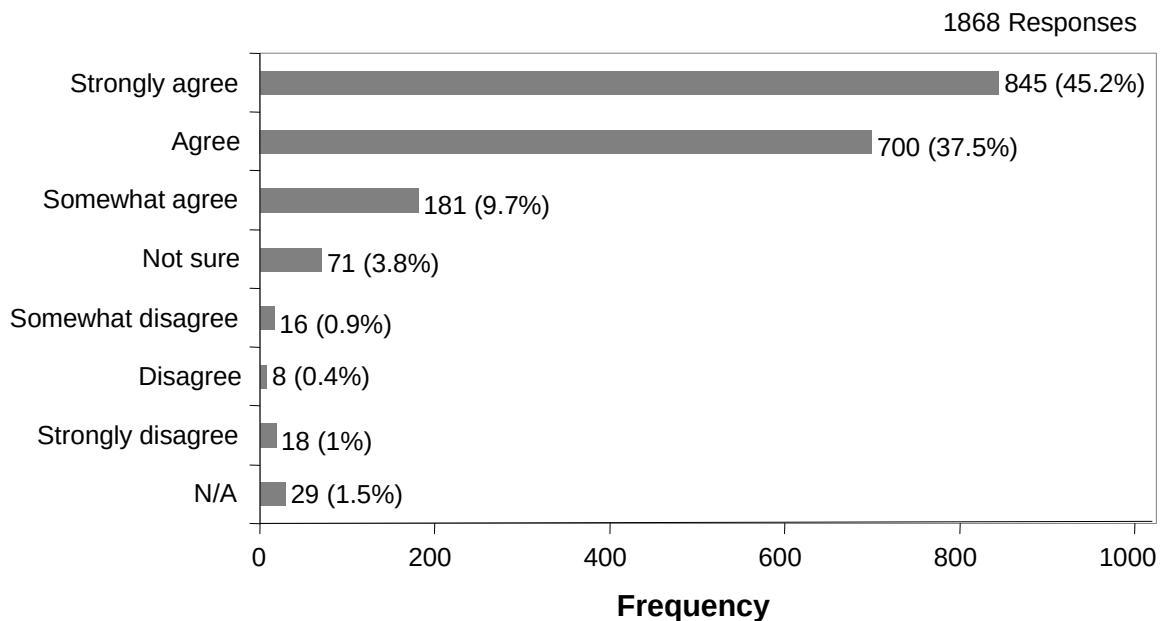


Figure 4: Measurement Helps Team Perform Better

The majority of the respondents felt that measurement was beneficial to their teams, with 82.7% strongly agreeing or agreeing with the statement. As shown in Figure 5, statistically significant differences were once again noted between management and staff responses to this question. For example, 96% of program managers and executives agreed that measurement was beneficial, while only 84% of programmers agreed.

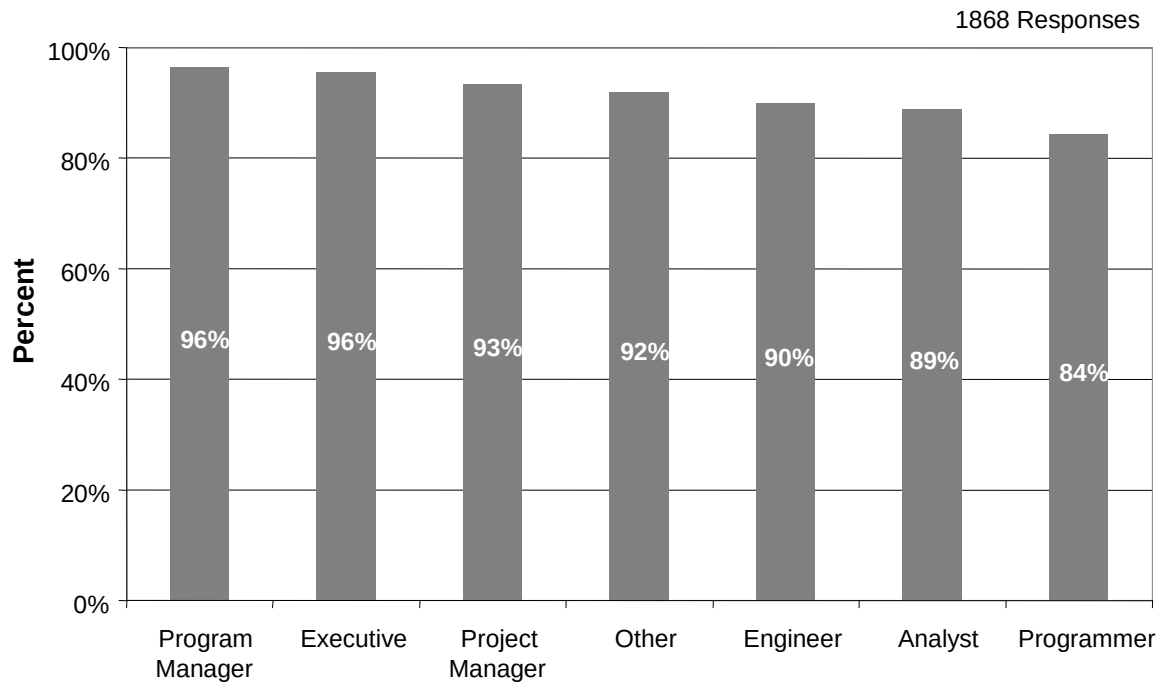


Figure 5: Measurement Helps Team Perform Better, by Job Type

3.2.4 Is Measurement Used to Understand Quality?

Respondents were asked how often they use measurement to understand the quality of the products or services on which they work. Their responses are shown in Figure 6.

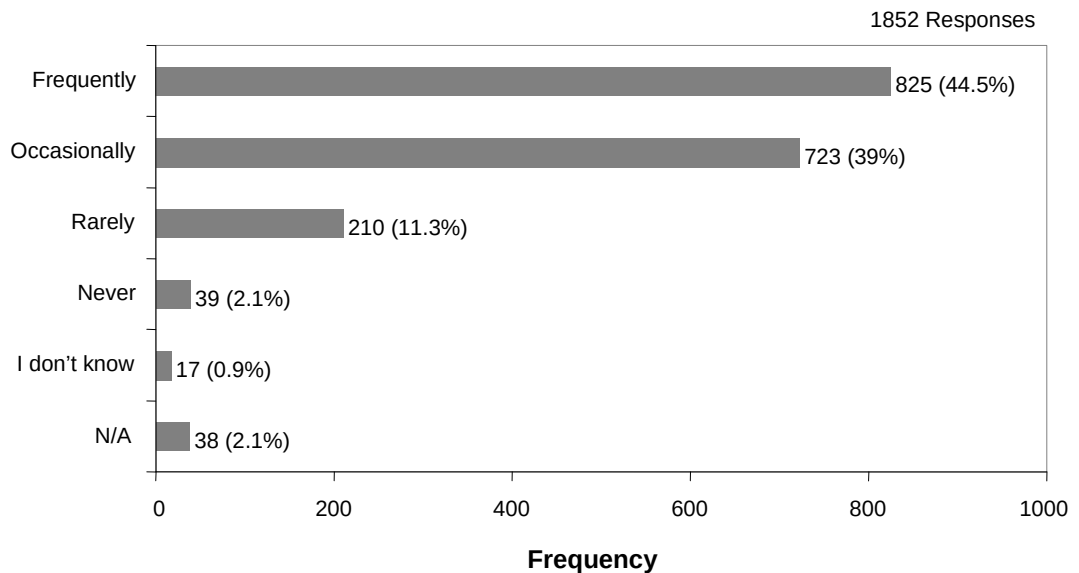


Figure 6: Measurement Used to Understand Quality

Figure 7 summarizes, by role, the responses to the following statement: *I use measurement to understand the quality of the products and/or services that I work on.*

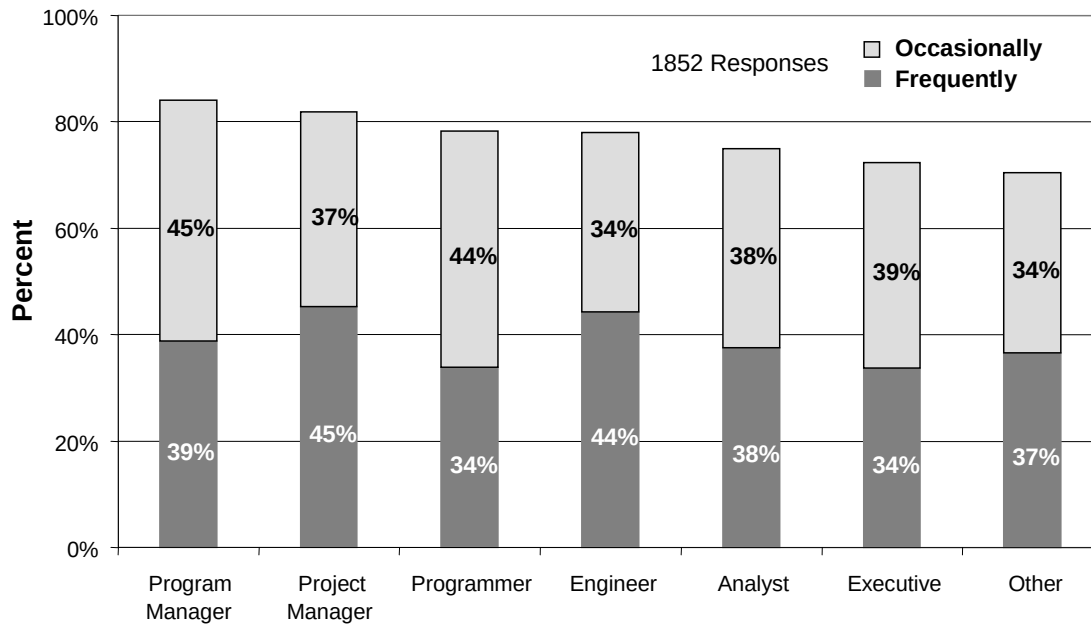


Figure 7: *Measurement Used to Understand Quality, by Job Type*

3.2.5 Are Documented Measurement Processes Used?

To determine the extent to which documented processes are used to collect data, respondents were asked to rate their level of agreement with the following statement: *My team follows a documented process for collecting measurement data.* Their responses are shown in Figure 8.

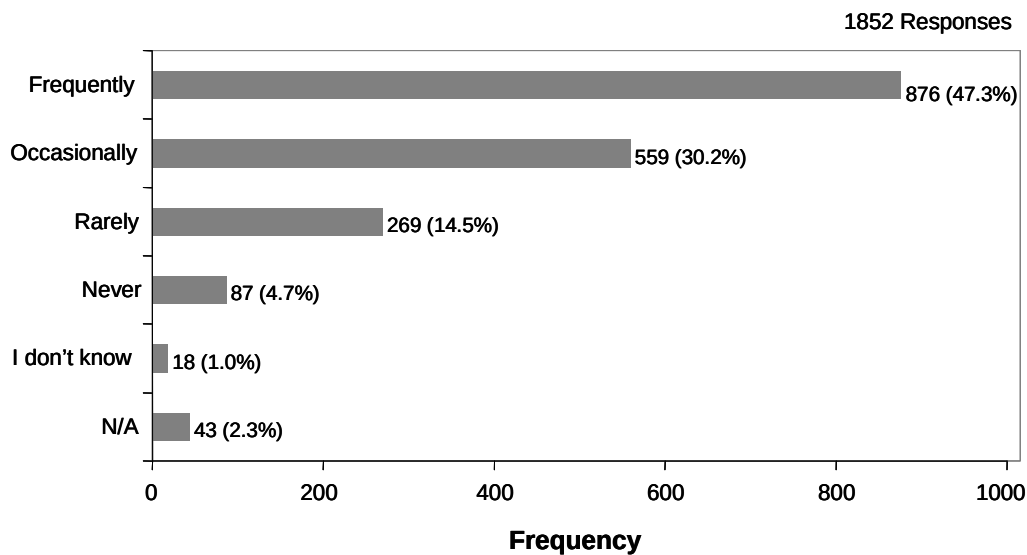


Figure 8: *Use of Documented Processes for Collecting Data*

As shown in Figure 9, the “other” group was the second most likely to use documented processes to collect measurement data. More than 72% of the respondents who selected “other” were in roles that typically promote the development and use of documented processes, such as process improvement and quality assurance type roles.

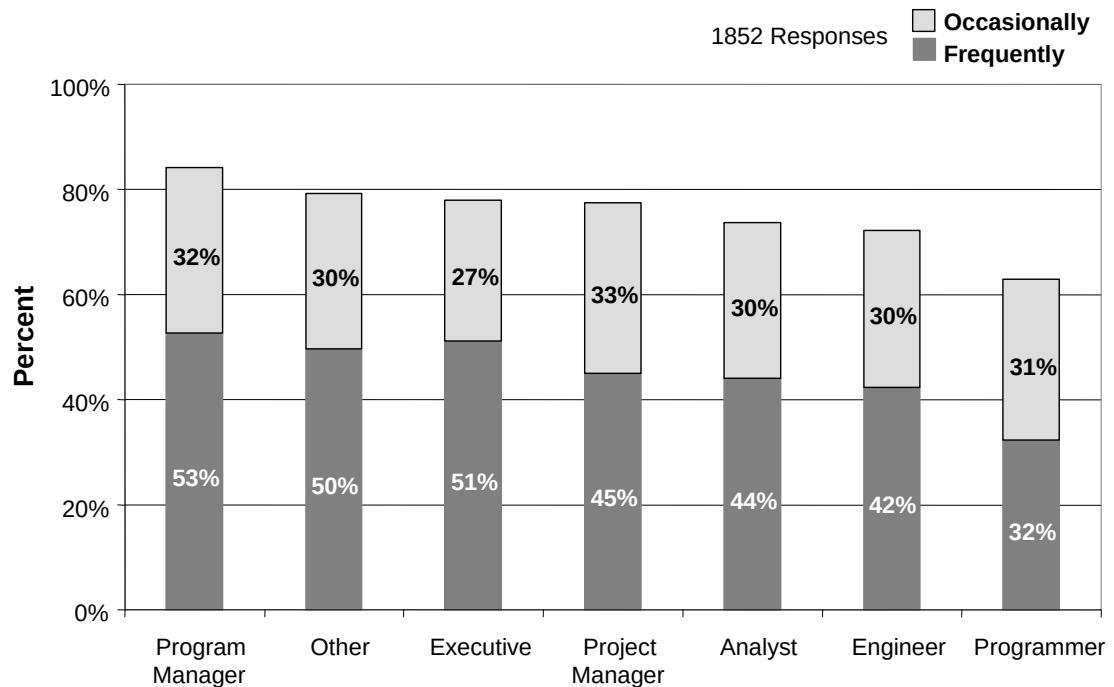


Figure 9: Use of Documented Processes for Collecting Data, by Job Type

Respondents were then asked if their teams followed a documented process for reporting measurement data to management. Their responses are shown in Figure 10.

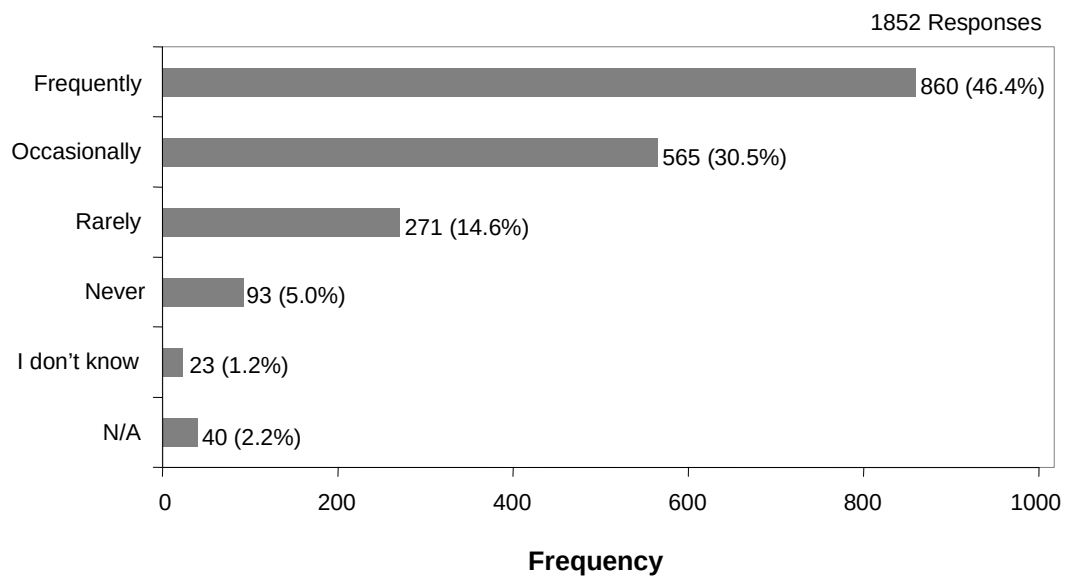


Figure 10: Documented Process Used to Report Data to Management

Those in management and “other” roles were more likely than analysts, engineers, and programmers to use a documented process, as shown in Figure 11.

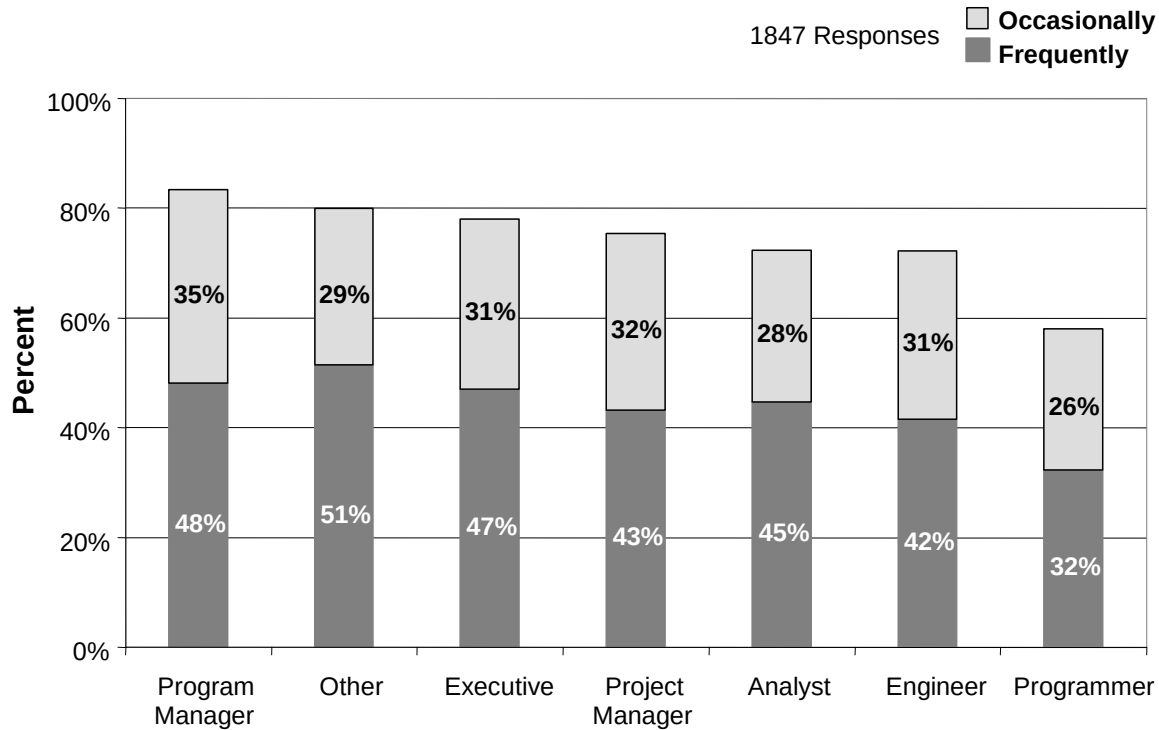


Figure 11: Documented Process Used to Report Data to Management, by Job Type

3.2.6 Are Measurement Definitions Understood and Consistent?

Respondents were asked to rate how well they agreed with the following statement: *The definitions of measures that are used in my organization are commonly understood and consistent.* Their responses are shown in Figure 12.

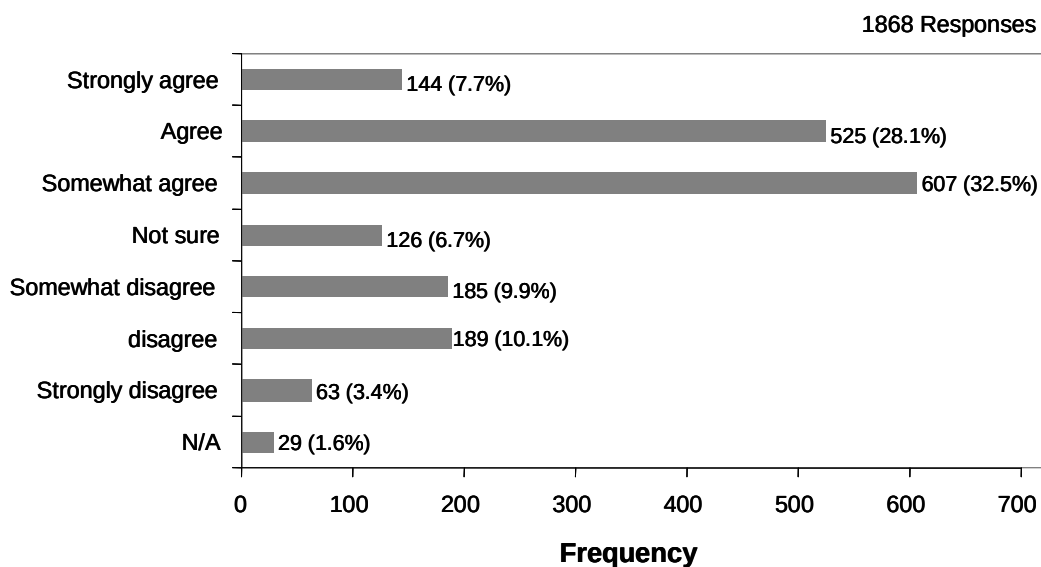


Figure 12: Measurement Definitions Are Understood and Consistent

Figure 13 shows the results cross-tabulated by role. The bars indicate the percentage of respondents who agree or strongly agree that measurement definitions are understood and consistent. Once again, there was a statistically significant difference (with confidence of 95%) between staff and management responses, as shown in Figure 13.

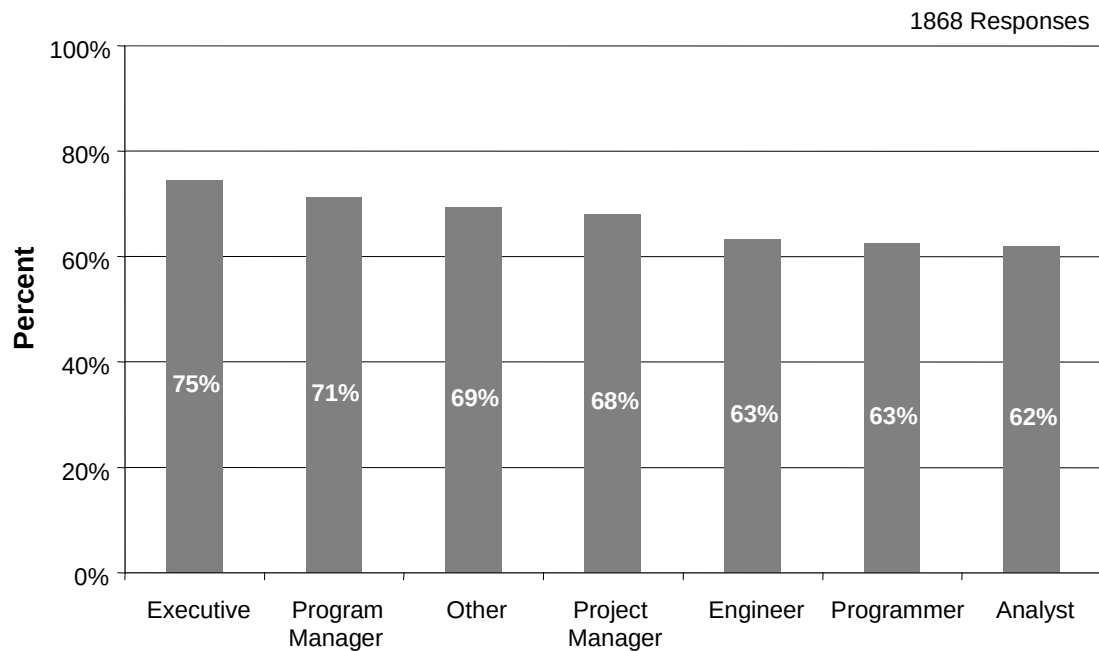


Figure 13: *Measurement Definitions Are Understood and Consistent, by Job Type*

3.2.7 Do Measurable Criteria Exist for Products and Services?

Respondents were asked to rate their agreement with the following statement: *There exist measurable criteria for the products and services to which I contribute.* Their responses are shown in Figure 14.

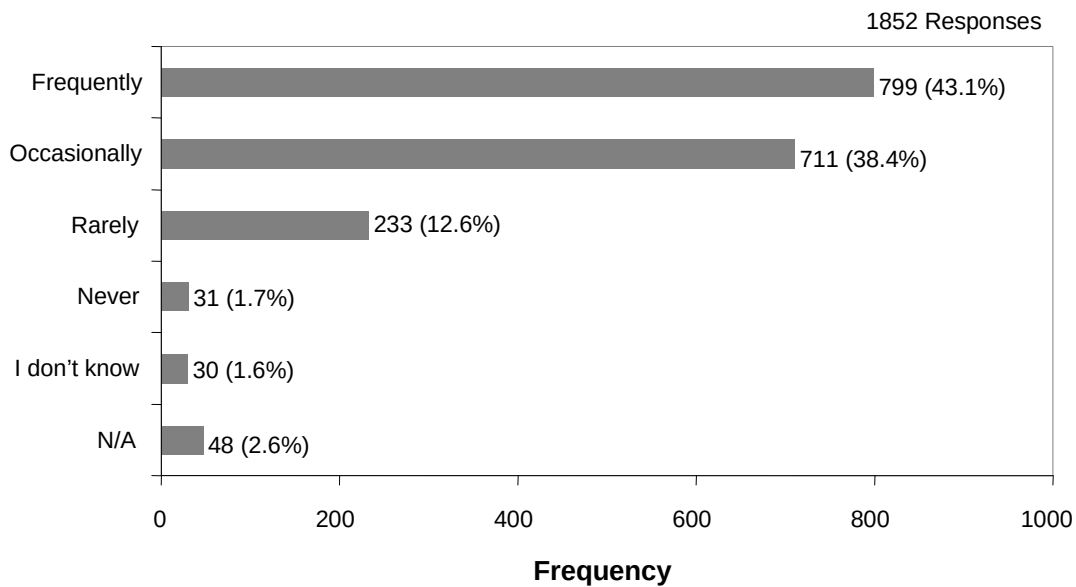


Figure 14: Measurable Criteria Exist for Products and Services

Figure 15 indicates that there are statistically significant differences in response (with 95% confidence) between management and staff.

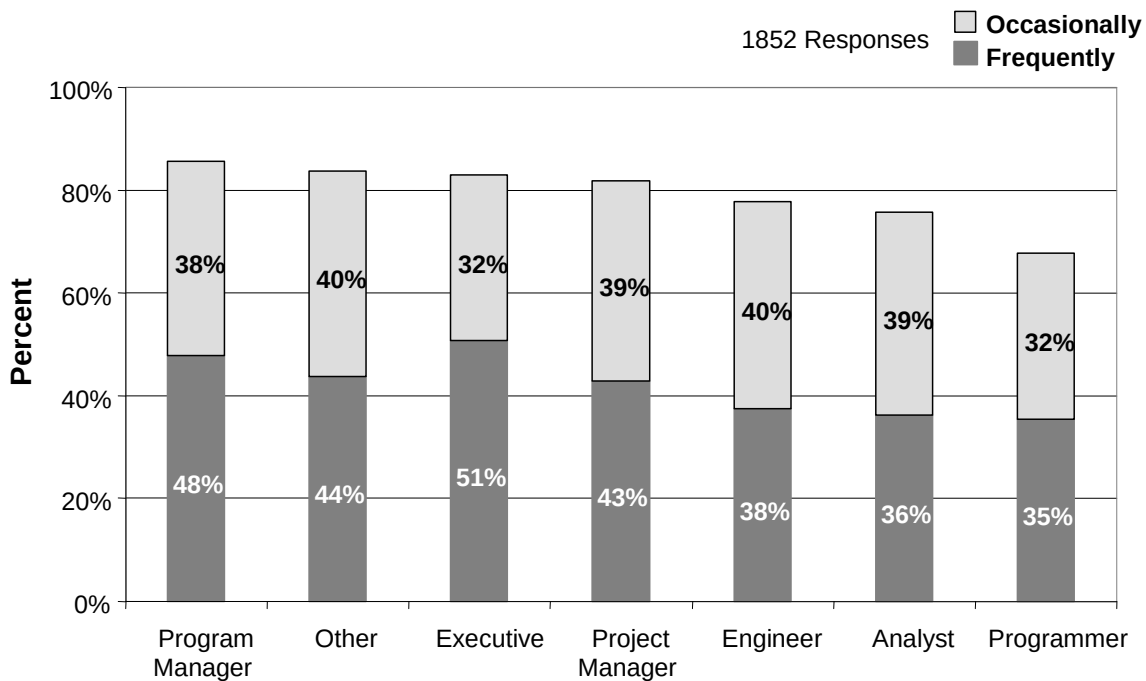


Figure 15: Measurable Criteria Exist for Products and Services, by Job Type

3.2.8 Is Corrective Action Taken When Thresholds Are Exceeded?

To help indicate how respondents used measurement, they were asked how often corrective action was taken when measurement data indicated that a threshold had been exceeded. Their responses are shown in Figure 16.

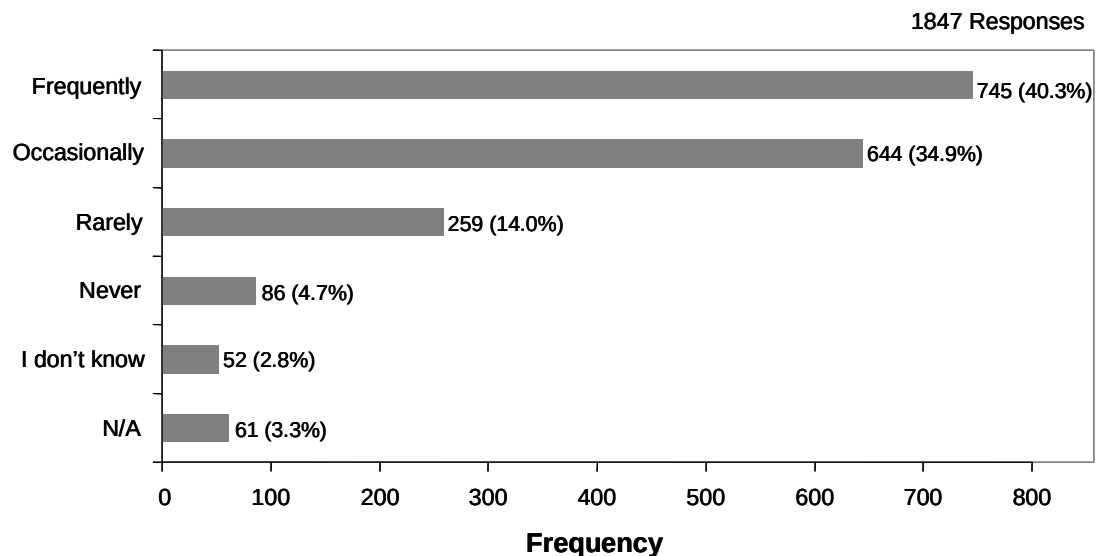


Figure 16: How Often Is Corrective Action Taken?

As shown in Figure 17, there are statistically significant differences (with 95% confidence) between the responses of management (program manager, executive, and project manager) and staff (engineer, analyst, and programmer).

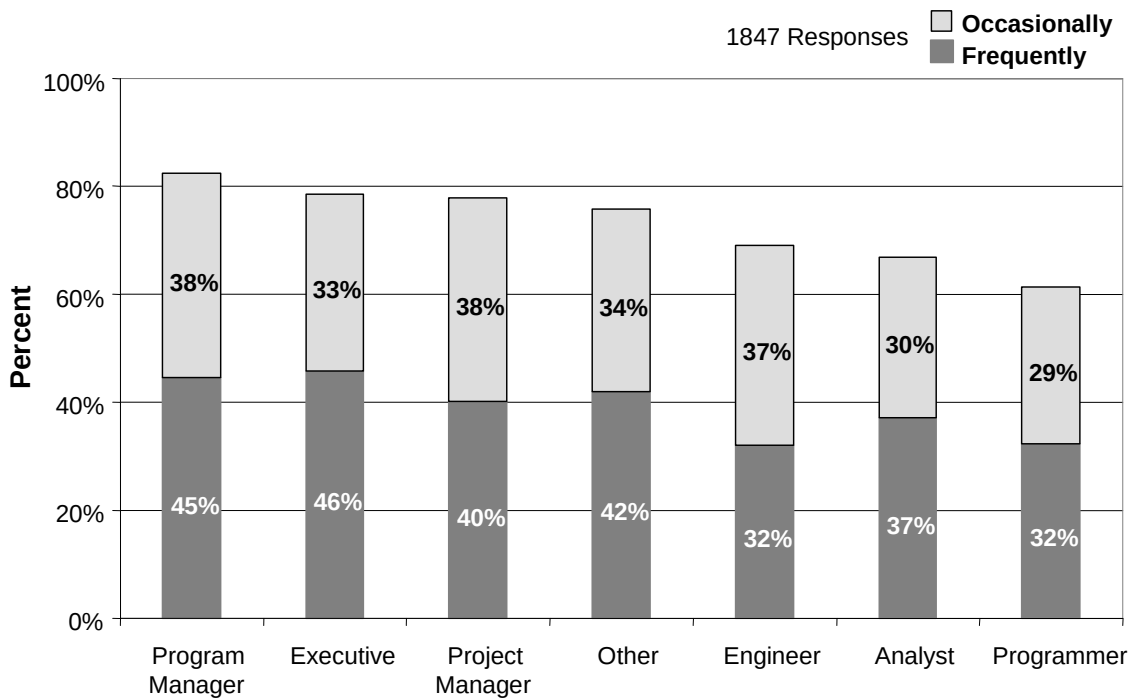


Figure 17: How Often Is Corrective Action Taken? by Job Type

It is interesting to compare the response shown in Figure 17 to the response to the question about measurable criteria that was shown in Figure 15. When compared to staff, management more often believes that corrective action is taken when a measurement threshold has been exceeded. However, when this same information is compared to earlier responses about whether measurable criteria exist, it appears that project managers and staff are consistent in the sense that they both believe that corrective action is taken for instances where measurable criteria have been identified. However, as shown in Figure 18, upper management appears less consistent, reporting that corrective action is not always taken even when measurable criteria exist.

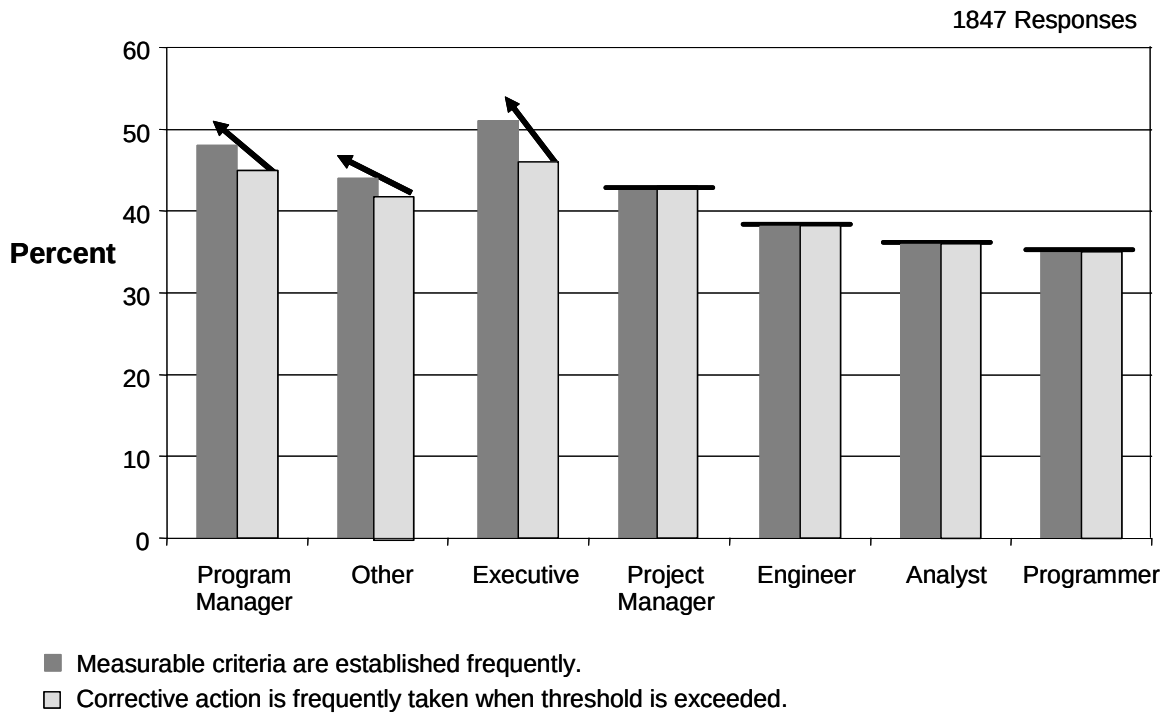


Figure 18: Action-Oriented Response to Measurement Information

3.3 Measurement Guidance Used

This portion of the survey was used to gauge which methods were most often used to identify, collect, and analyze measurement data.

Respondents were asked to identify the methods they used by selecting one or more responses from the list below:

- I do not use a defined method.
- Goal-Driven Software Measurement Method
- Practical Software and Systems (PSM) Method
- GQM (Goal-Question-Metric)
- Capability Maturity Model[®] Integration (CMMI[®]) Measurement and Analysis (M&A) process area
- Personal Software ProcessSM (PSPSM) & Team Software ProcessSM (TSPSM)

[®] Capability Maturity Model and CMMI are registered in the U.S Patent and Trademark Office by Carnegie Mellon University.

SM Personal Software Process, Team Software Process, PSP, and TSP are service marks of Carnegie Mellon University.

- International Standards Organization (ISO) 15939 (Information Technology – Software Engineering – Software Measurement Process)
- Other (please specify)

Their responses are shown in Figure 19.

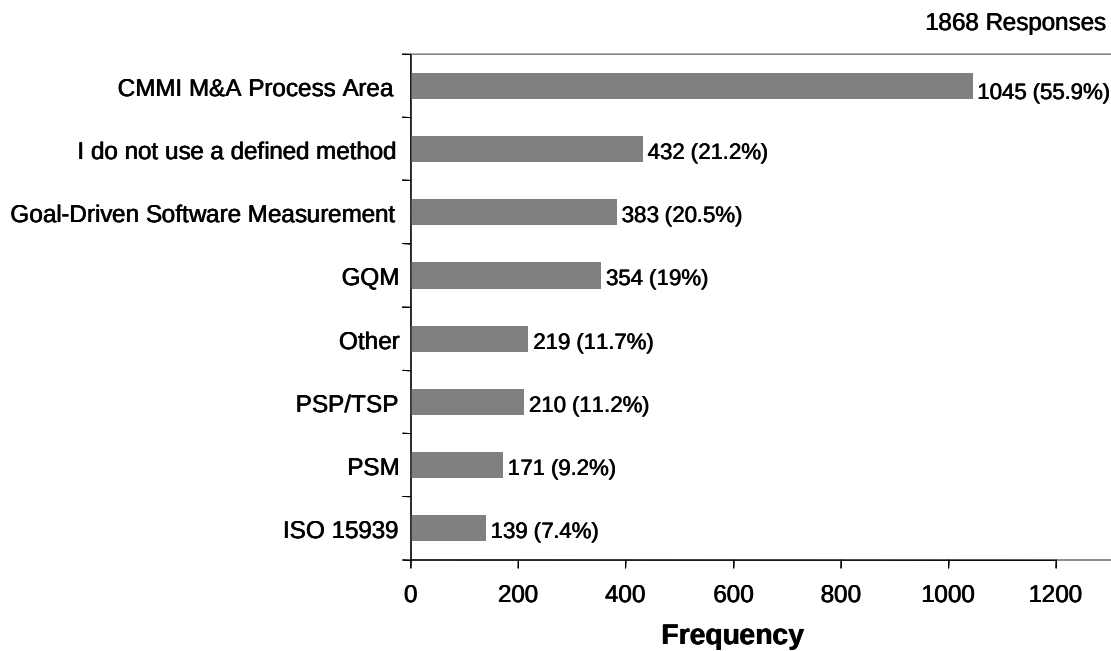


Figure 19: Measurement Methods Used

There were 219 respondents who selected “other” and specified a measurement method. These responses were clustered using affinity grouping [Brassard 96], and the categories are shown in Figure 20. The percentages in Figures 19 and 20 (combined) exceed 100% because many people reported using more than a single measurement method.

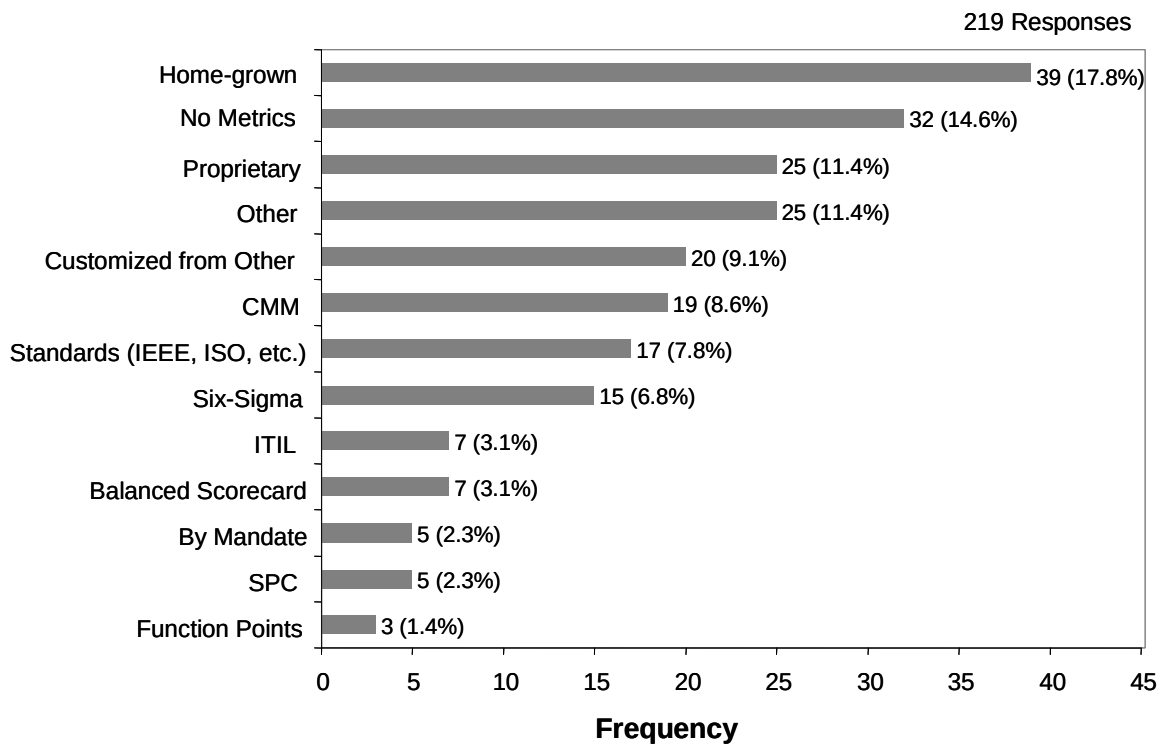


Figure 20: Other Measurement Methods Used¹

Figures 21-27 could be misleading if one assumes that only a single measurement guidance method was used by an organization. As shown in these figures, most organizations use more than one method to identify, collect, and analyze measurement data.

Figures 21-27 illustrate the use of multiple methods by showing the percentage of respondents who used a given measurement method in conjunction with other methods. Consider for example Figure 21. The left-most bar in Figure 21 shows that 511 respondents stated that they use the CMMI Measurement & Analysis process area as their measurement method and that this is the only method that they used. The next bar to the right shows that 256 of 354 respondents (or 72.3% of those that use GQM) stated that they use the CMMI M&A and GQM. Therefore, the percentage above the GQM display bar indicates that 72.3% of those who reported using GQM also use CMMI M&A. Each of the display bars in Figures 21-27 should be interpreted this way.

¹ In Figure 20, ITIL is Information Technology Infrastructure Library and SPC is Statistical Process Control.

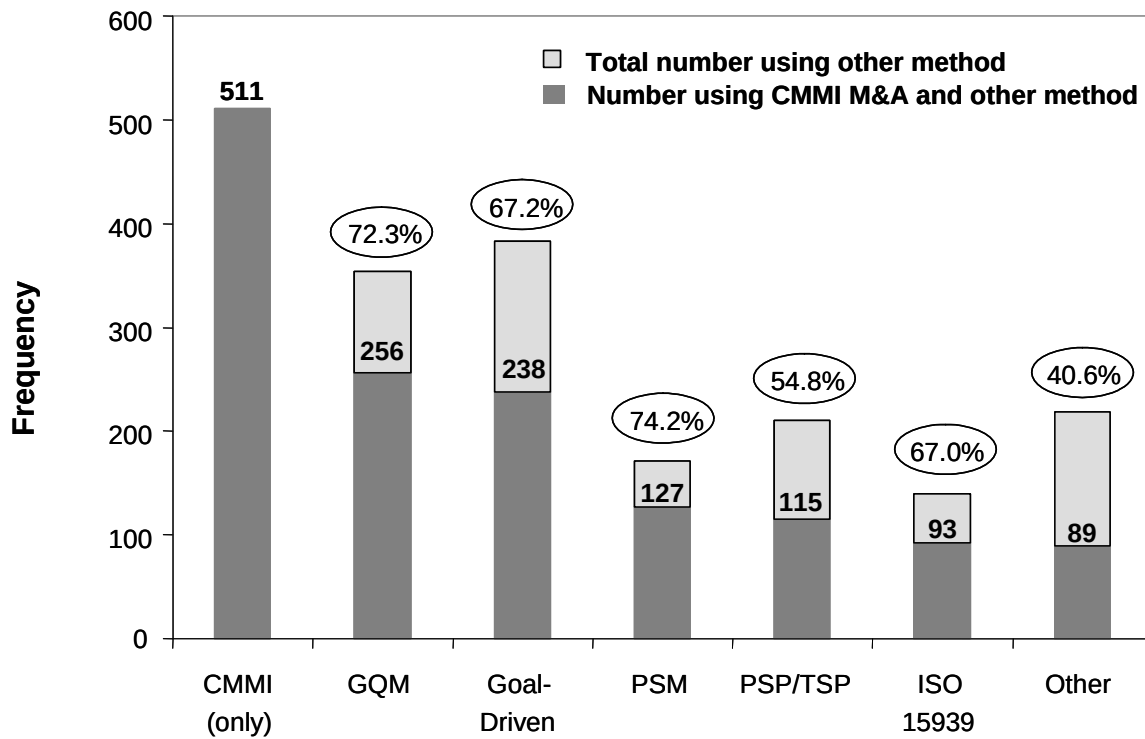


Figure 21: Respondents Using CMMI M&A Process Area

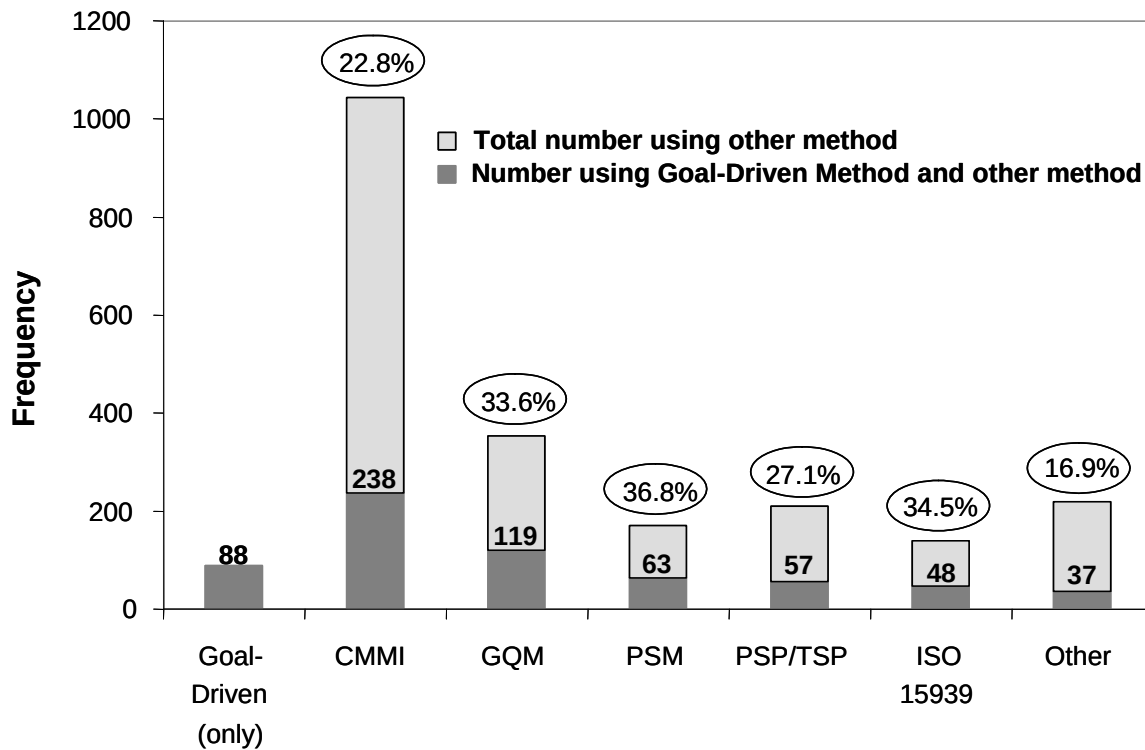


Figure 22: Respondents Using Goal-Driven Method

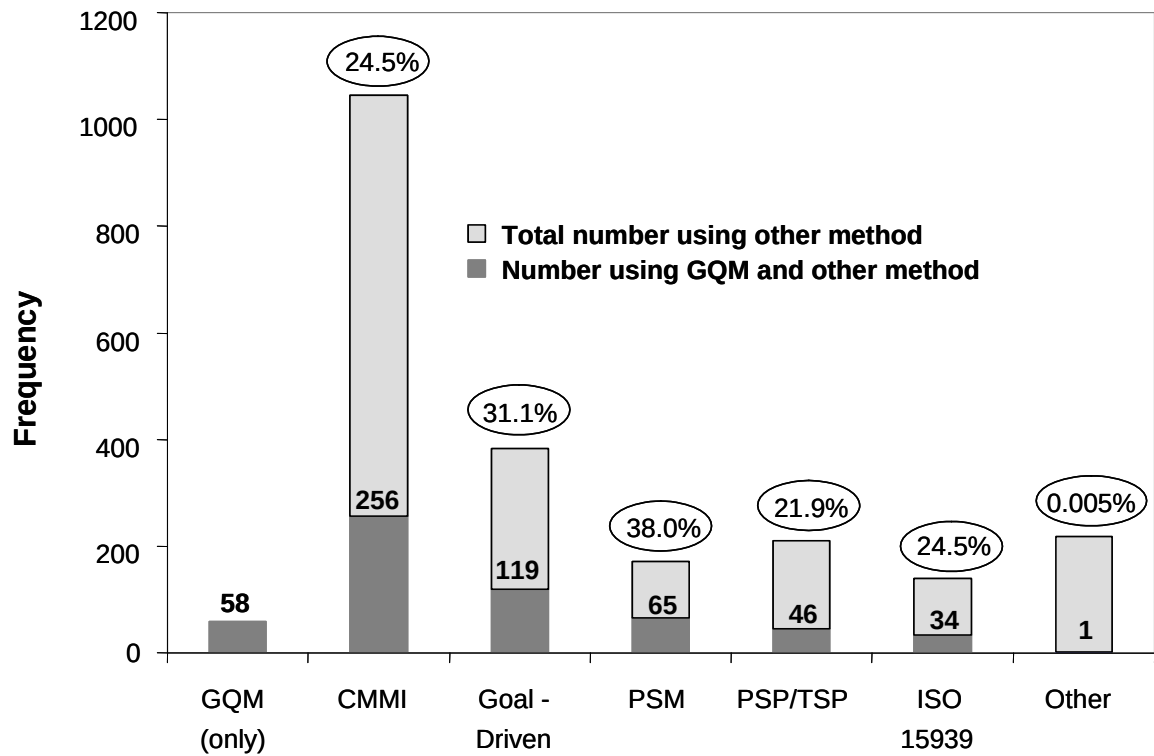


Figure 23: Respondents Using GQM

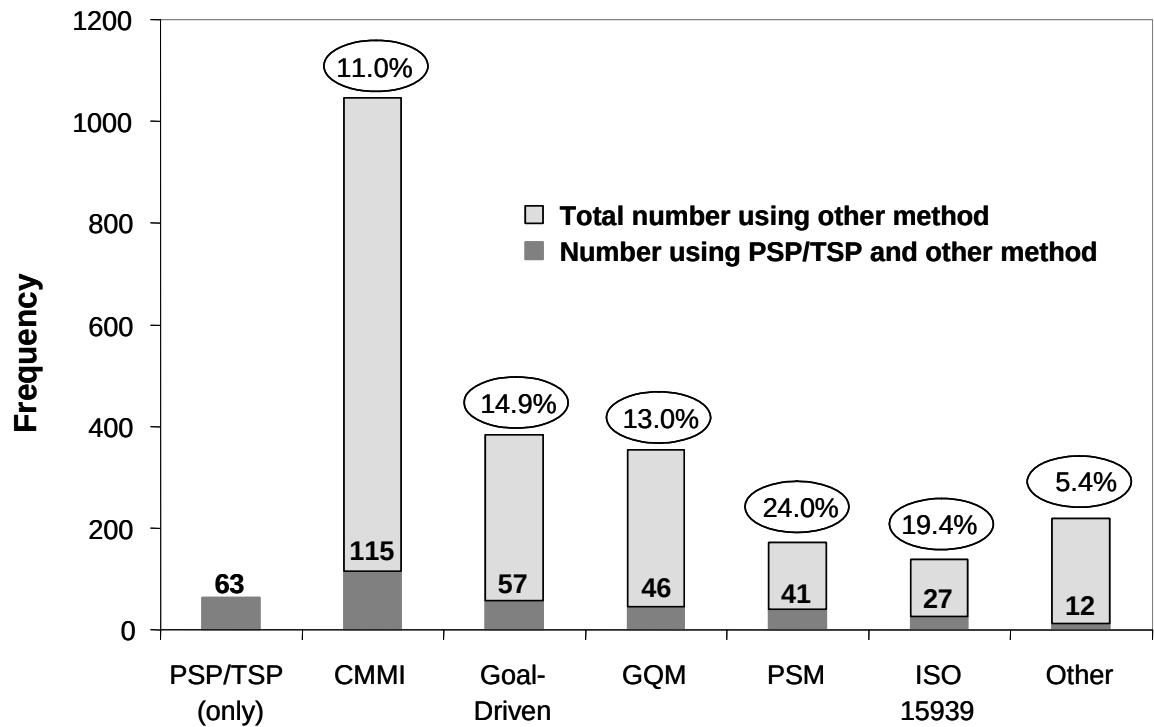


Figure 24: Respondents Using PSP/TSP

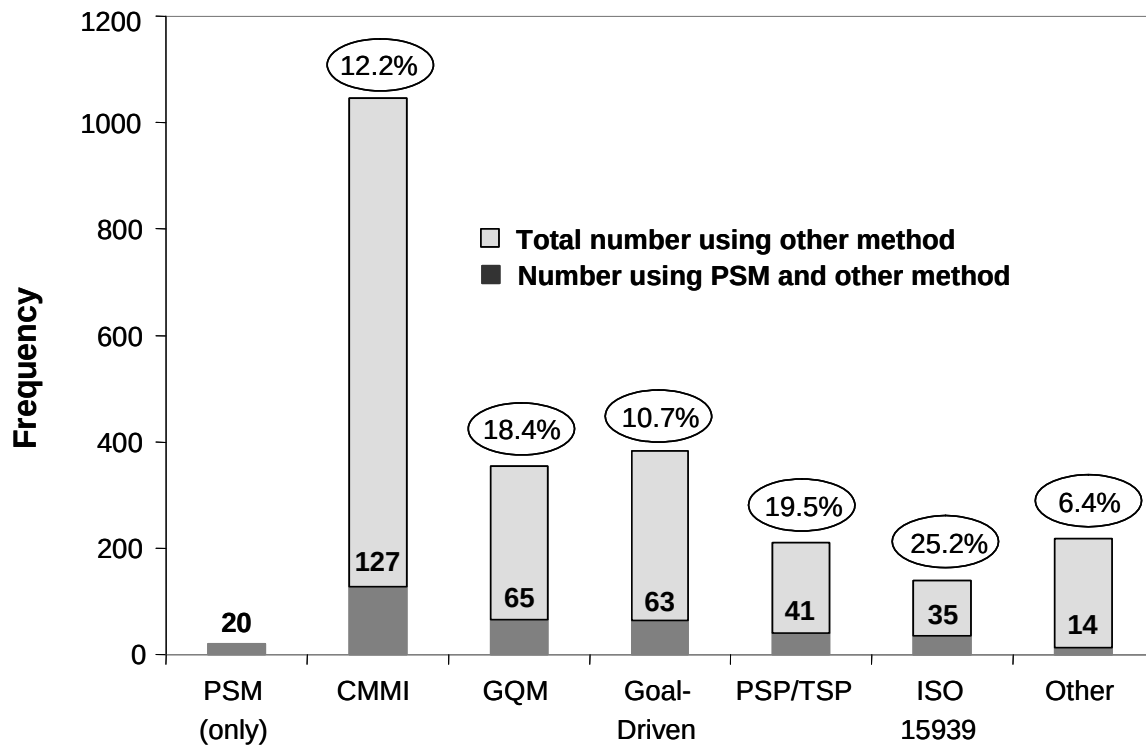


Figure 25: Respondents Using PSM

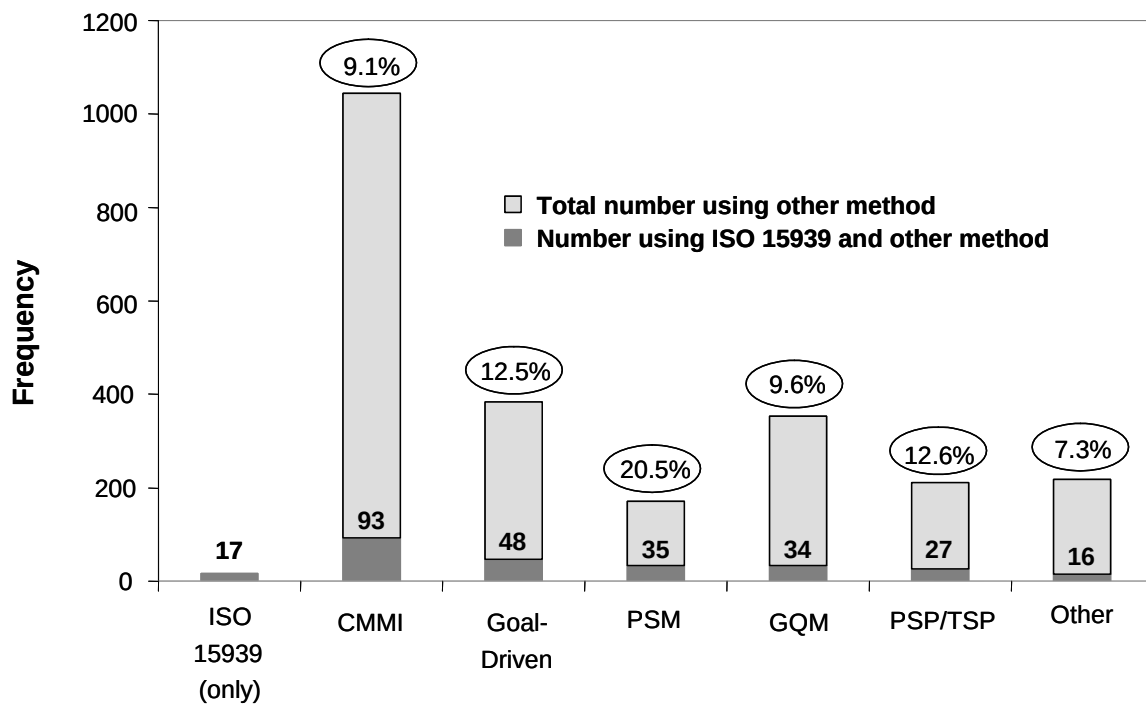


Figure 26: Respondents Using ISO 15939

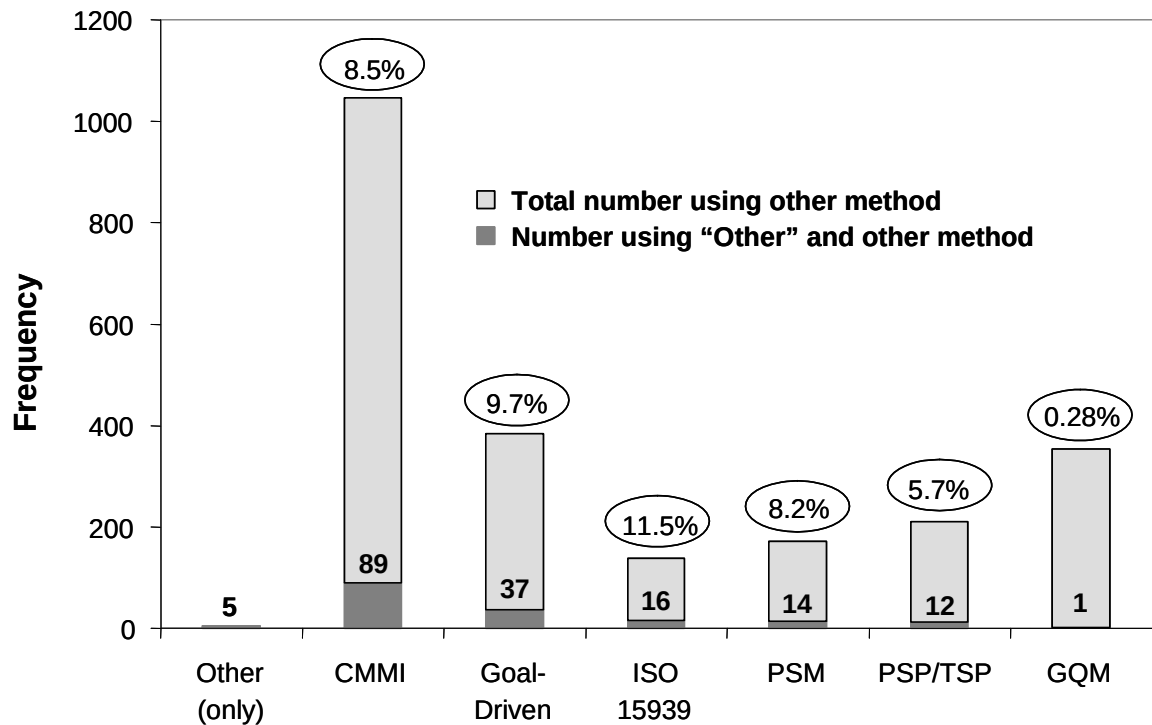


Figure 27: Respondents Using "Other" Methods

Figures 21-27 illustrate that most organizations rely on multiple methods for guiding the implementation of their measurement programs. They provide an indication of how organizations combine guidance methods to achieve their measurement program goals.

3.4 Measures That Are Reported

The survey questionnaire included questions to determine what measures are reported most often. A list of measures was provided, and respondents were asked to rate how often they reported each measure. The listed measures included

- capability/requirements stability (creep)
- effort applied to tasks
- defects identified
- defects removed
- code growth
- schedule progress
- risks identified

Respondents were asked to identify any additional measures that were not listed but were considered key to their organizations. They were also asked to indicate how often these other measures were reported (for example, "time spent in peer review [weekly]").

Figure 28 shows the number of respondents who indicated that they report (or do not report) each measure. For example, 1461 individuals responded that they report risks, while 156 individuals responded that they do *not* report risks. The percentages reflect the proportion that report each of the measures.

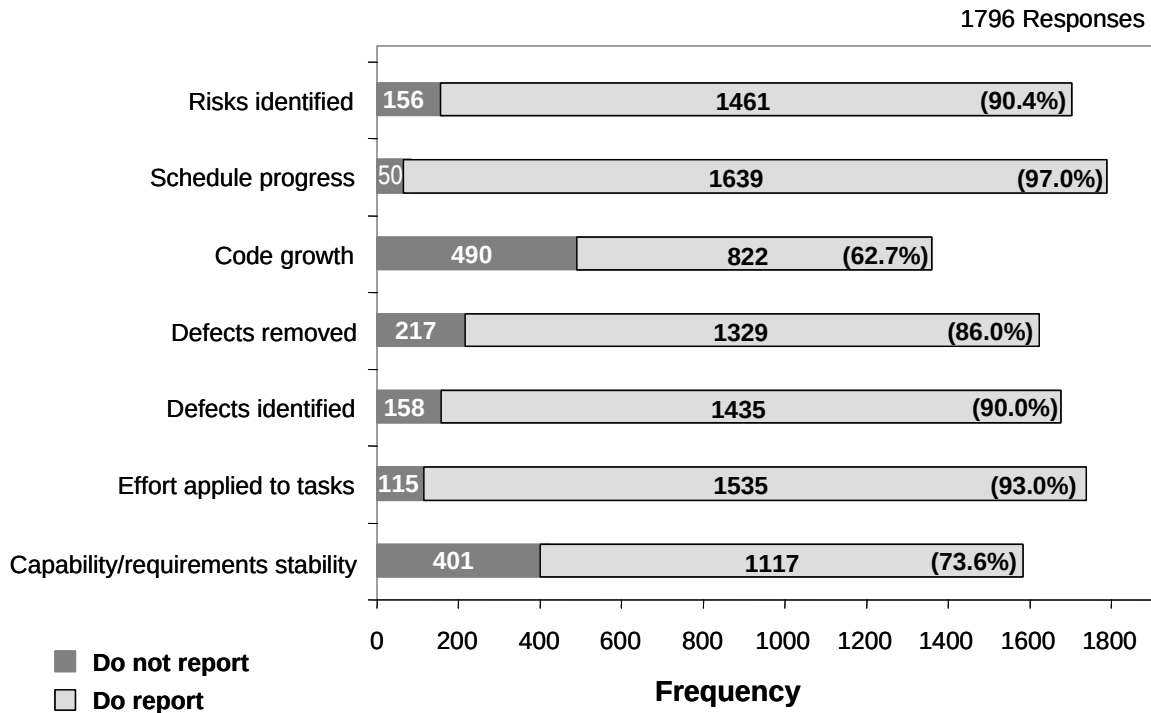


Figure 28: Measures That Are and Are Not Reported

Figures 29-35 display how frequently respondents stated that each measure was reported. In general, it appears that most measures are reported on a weekly or monthly basis.

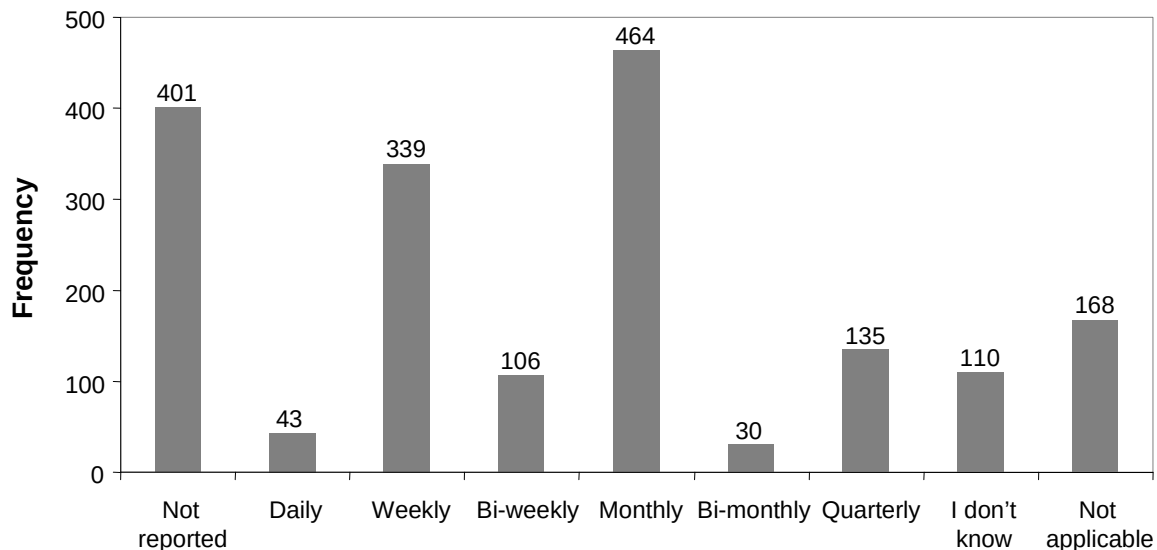


Figure 29: Frequency of Reporting Capability/Requirements Stability

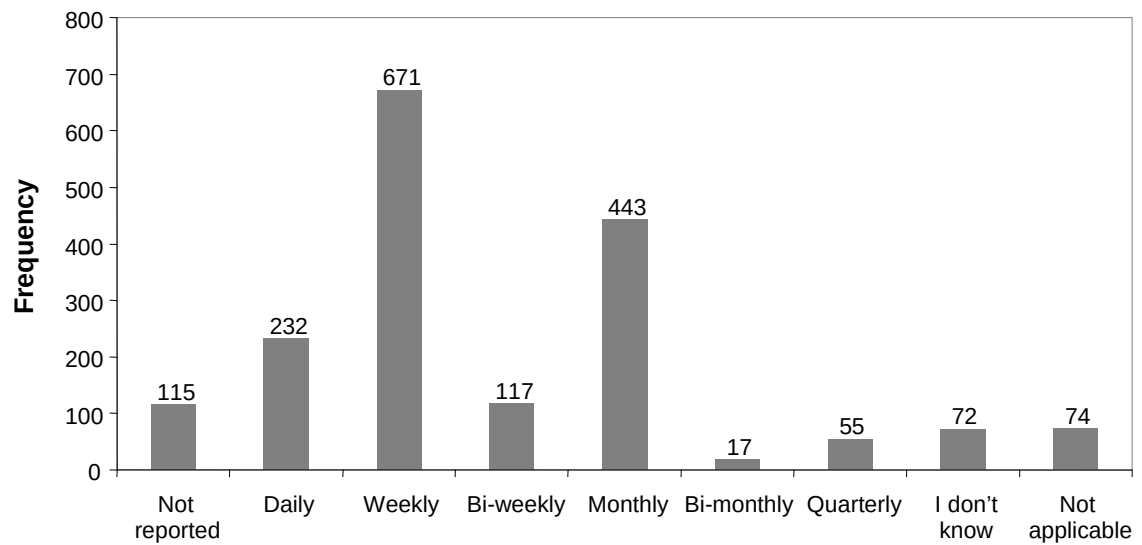


Figure 30: Frequency of Reporting Effort Against Tasks

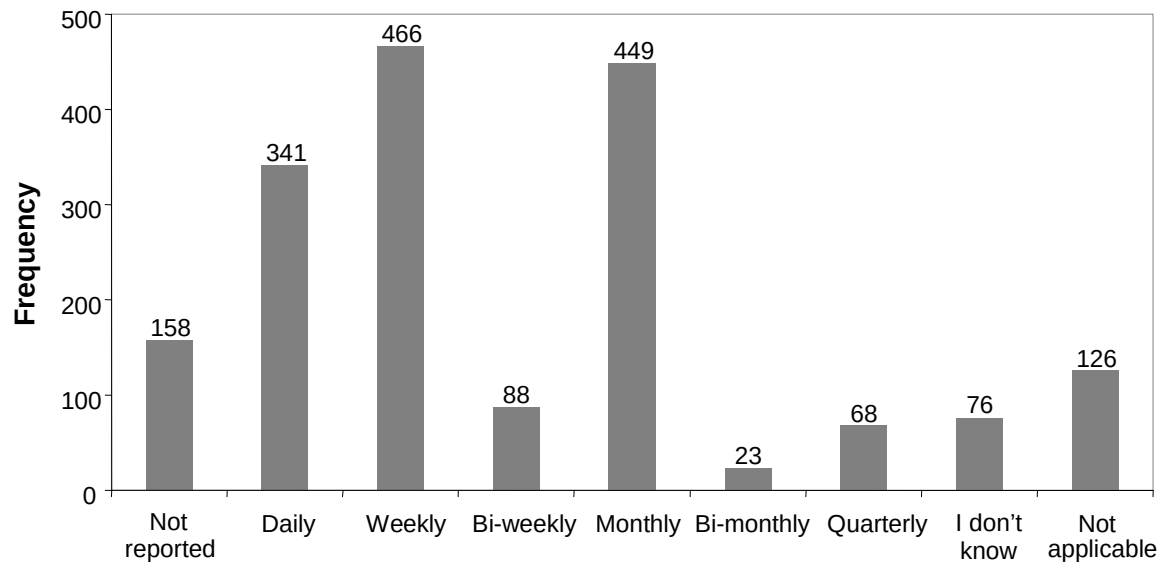


Figure 31: Frequency of Reporting Defects Identified

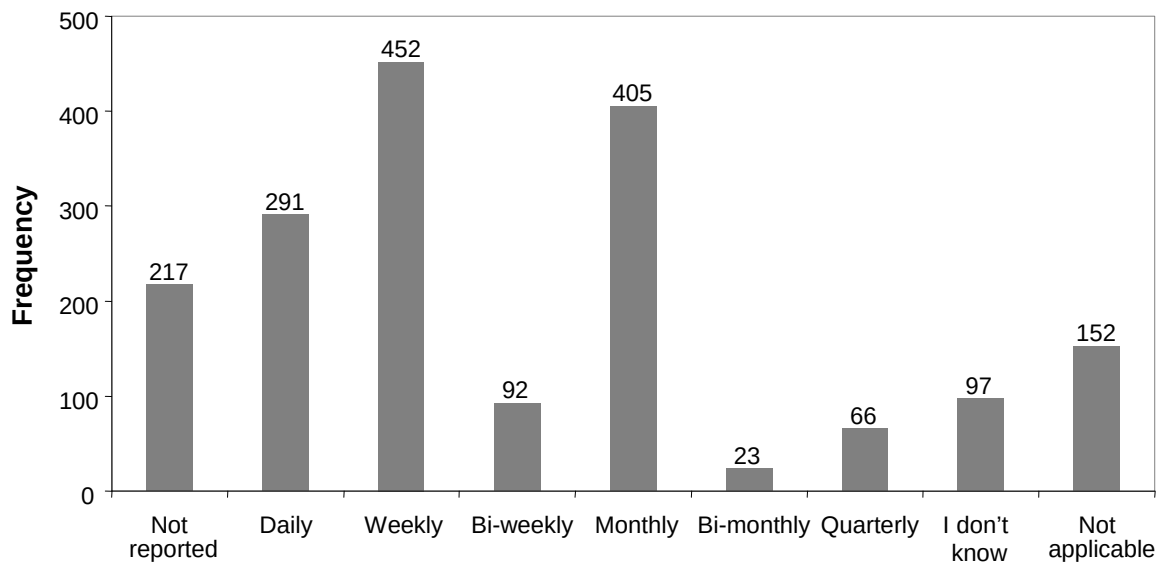


Figure 32: Frequency of Reporting Defects Removed

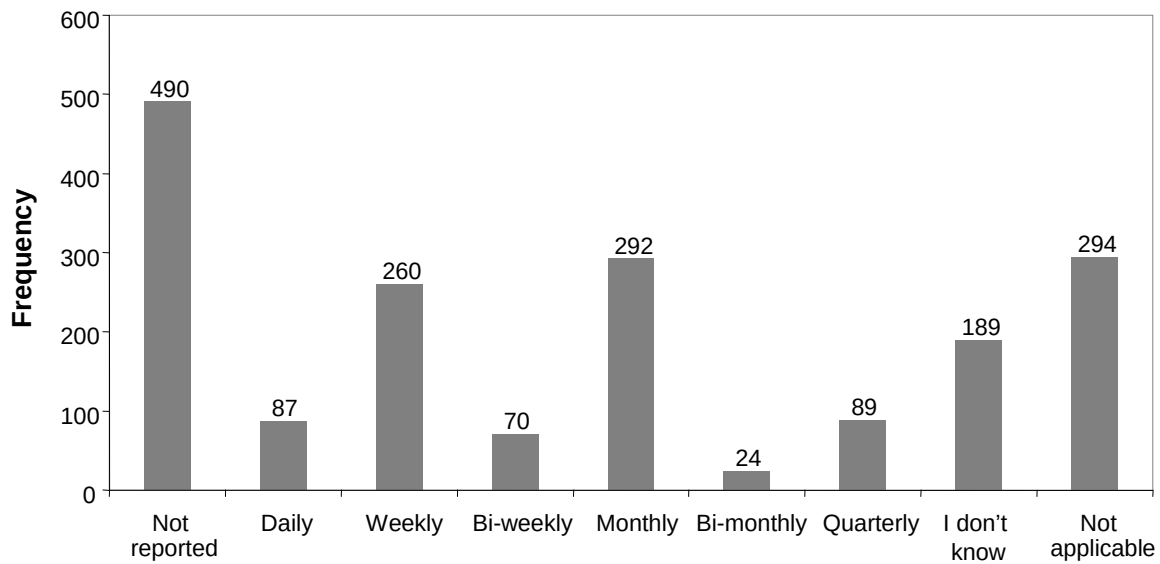


Figure 33: Frequency of Reporting Code Growth

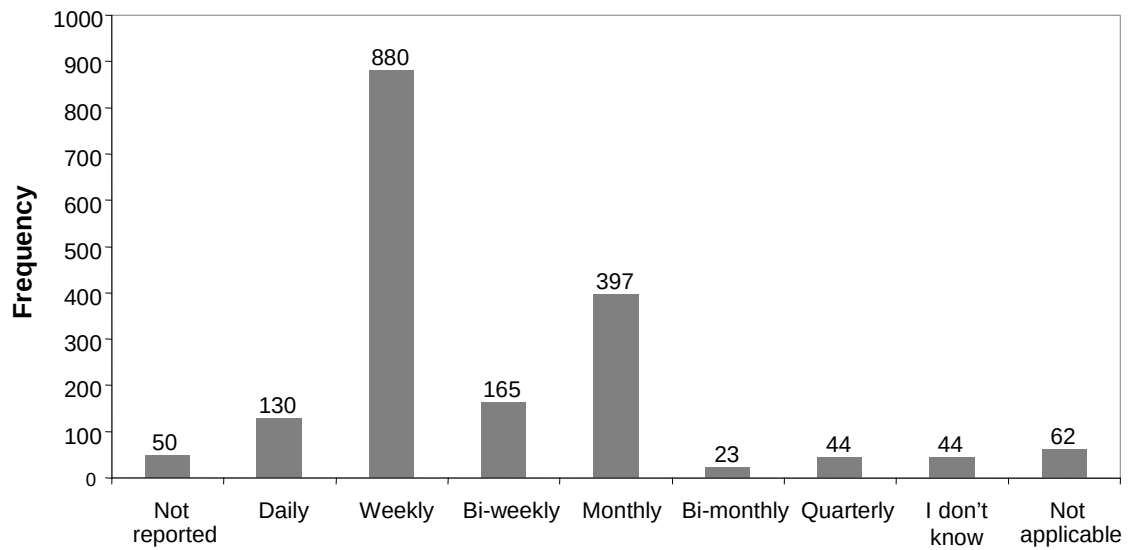


Figure 34: Frequency of Reporting Schedule Progress

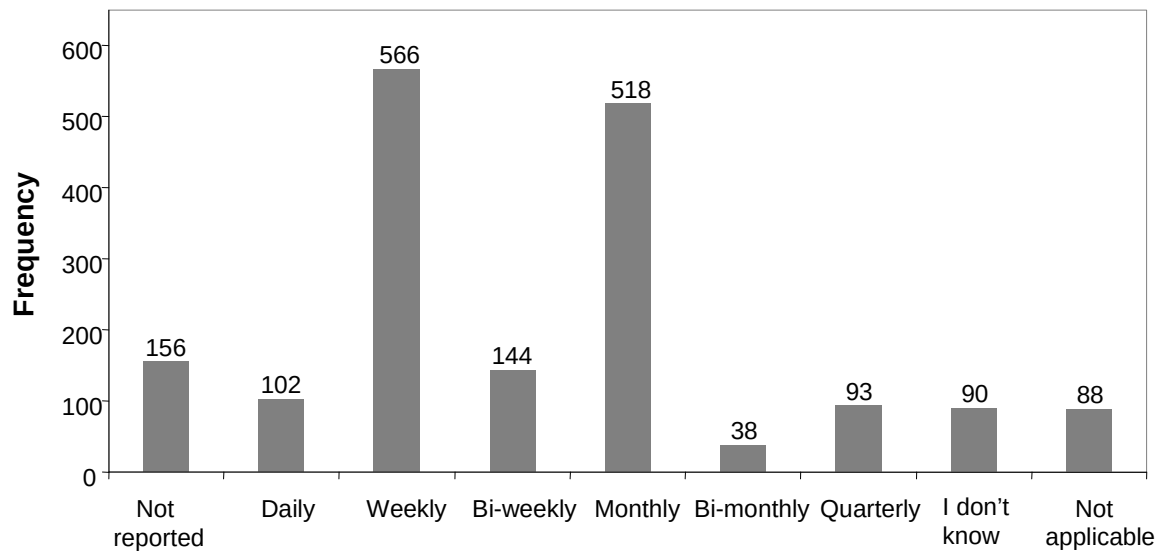


Figure 35: Frequency of Reporting Identified Risks

The respondents were asked to identify additional measures that they report, and they specified a large number of measures. These measures were categorized by affinity grouping, and the categories are shown in Figure 36. Percentages are shown for the top three measurement categories that were indicated.

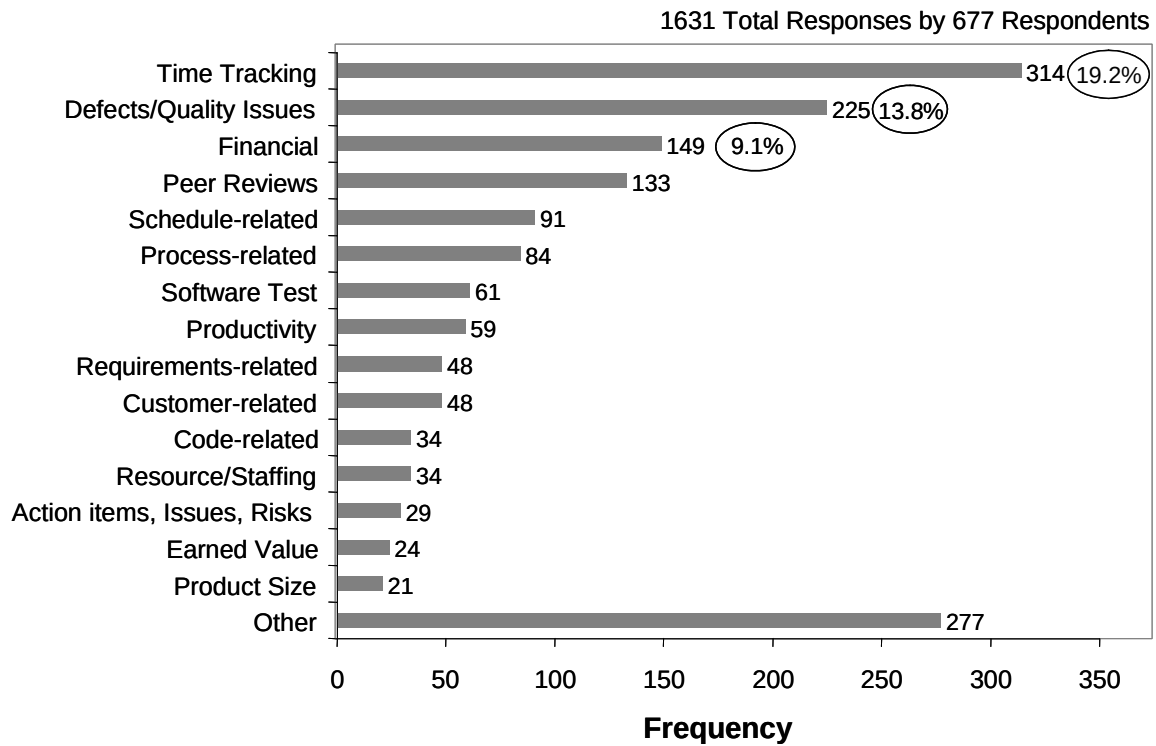


Figure 36: Other Measures Reported

4 Summary Observations

This section describes some key observations and general trends based on the statistical analysis of the results.

One might naturally ask if the current state of measurement practice represents a healthy state. From the perspective of SEI's Software Engineering Measurement and Analysis (SEMA) Group, there is still a significant gap between the current and desired state of measurement practice. This is evidenced by a number of response profiles that appear to indicate a lack of effective implementation and follow-through when it comes to measurement practices. For example, consider Figure 14 on page 14. This figure summarizes the responses to the statement, "There exist measurable criteria for the products and services to which I contribute." Note that only 43% indicated that this was *frequently* the case, while 52.7% responded that this was the case occasionally, rarely, or never. Certainly, this outcome does not represent a best practice benchmark. One could look at many of the questionnaire response profiles in a like manner. Generally speaking, based on the results of this survey, we believe that there is still much that needs to be done so that organizations use measurement effectively to improve their processes, products, and services.

4.1 Responses Patterns for Management Compared to Staff

The survey results indicate that when it comes to measurement within their organization, management and staff have different perspectives.

Respondents who identified themselves as executives, program managers, or project managers were considered "management." Engineers, analysts, and programmers were considered "staff." In general, there were significant differences in response patterns between management and staff.

Statistical tests of significance demonstrated that the differences were significant with confidence of at least 99% in almost all cases (and 99.9% in some cases). Hypothesis tests for equality of proportions and the Chi-Square test for significance were performed.

When compared to staff, management responded more strongly that

- they understand the purposes for measurement
- measurement helps teams perform better
- they often follow a documented process for collecting and reporting measurement data
- measurement definitions are commonly understood and consistent in their organization
- measurable criteria exist for their products and services
- corrective action is taken when a measurement-based threshold has been exceeded

The difference between the response profiles of management and staff seems to indicate a lack of congruent communication between those in the two roles. One possibility for this is that management is responsible for policy that leads to the instantiation of a measurement program. When individuals are responsible for the success of a system, they may have an overly optimistic view of how the system is actually perceived and implemented by those who are not involved in decisions that led to the policy. In any case, the data indicate the need for mechanisms that will lead to a shared and aligned understanding among all individuals within an organization.

The only questionnaire item that did not exhibit a significant difference in response between management and staff was the following statement (questionnaire item 11): *I use measurement to understand the quality of the products and/or services that I work on.*

4.2 Impact of the Organization's Size

The size of the respondent's organization appeared to have an impact on the response pattern associated with many of the questionnaire items. Table 2 shows the percentage of respondents who responded "frequently" to the listed questionnaire items for the three size categories.² Note that the degree of frequency tended to increase with the size of the organization.

² Possible responses to these questionnaire items included (a) frequently, (b) occasionally, (c) rarely, (d) never, (e) I don't know, and (f) N/A.

Table 2: Percentage of Respondents Who Responded “Frequently” to the Listed Questionnaire Items

Questionnaire Item	Number in Organization		
	≤100	101-499	≥500
There exist measurable criteria for the products and services to which I contribute.	38.9%	39.2%	49.6%
I use measurement to understand the quality of the products and/or services that I work on.	38.4%	42.0%	52.8%
My team follows a documented process for collecting measurement data.	42.3%	46.2%	53.1%
My team follows a documented process for reporting measurement data to management.	37.0%	46.4%	54.7%
Corrective action is taken when measurement data indicate that a threshold has been exceeded.	35.1%	41.1%	46.2%
I understand the purposes for the data I collect or report.	65.7%	71.6%	72.1%

4.3 Response Patterns for Industry Compared to Government

The response patterns for individuals who work for industry were compared to those of individuals who work for government. While there were statistically significant differences for the responses for some questionnaire items, the differences between the two groups were not significant for others. Table 3 shows the percentage of respondents who agree or strongly agree to the questionnaire items for each category. Note that there is a significant difference in the level of agreement between government and industry for these items. For all other questionnaire items, the differences in the responses between government and industry respondents were insignificant.

Table 3: Percentage of Government vs. Industry Respondents Who Strongly Agree or Agree to Listed Questionnaire Items

Questionnaire Item	Government	Industry
Generally speaking I believe that using measurement-based data helps my team to perform better than without using it.	80.0%	84.5%
The definitions of measures that are used in my organization are commonly understood and consistent.	31.9%	37.1%

4.4 Response Pattern for United States Compared to Other Countries

Questionnaire responses were compared between respondents from the United States and all other countries combined. As shown in Table 4, there were only two questionnaire items where there were significant differences in the responses.

Table 4: Percentage of Respondents Who Strongly Agree or Agree, by Country

Questionnaire Item	United States	Other Countries
Generally speaking I believe that using measurement-based data helps my team to perform better than without using it.	80.10%	85.90%
The definitions of measures that are used in my organization are commonly understood and consistent.	31.30%	42.40%

4.5 Use of Measurement Information

The survey results indicate that measurement information is not always being used effectively because it is not being acted upon. It is notable and somewhat alarming that only 40.3% of all respondents reported that corrective action is taken when a measurement threshold is exceeded. Close to 20% of respondents reported that corrective action is rarely or *never* taken when a measurement threshold is exceeded.

Measurement doesn't help much unless the information is acted upon.

4.6 Methods Used

The CMMI M&A process area was identified as the measurement method used most often to identify, collect, and analyze measurement data. Approximately 56% of respondents reported using the CMMI M&A process area, while 27.4% of all respondents reported that the CMMI M&A process area was the *only* method that they used (see Figure 21 on page 20).

As mentioned in the introduction of this report, the population for this study was composed of individuals who had contacted the SEI. Therefore, these individuals may have had prior knowledge of SEI products and services. Proper interpretation of the results must take into account the possibility that the results were biased in favor of SEI products and services.

Approximately 41% of all respondents stated that they used only a single method for identifying, collecting, and analyzing measurement data, while 59% use two or more methods.

There were 432 respondents (21.1%) who reported that they do not use any measurement method.

4.7 Measures Reported

Schedule and the time applied to a task were the measures that respondents indicated reporting most often. Most respondents (97%) indicated that schedule progress was the measure most often reported, while 93% indicated that effort applied to task was reported. In addition, some respondents listed other measures that they report, and 19.2% of these were related to time tracking.

Code growth, capability, and requirements stability were the measures that respondents indicated reporting the least often, with 27.3% not reporting code growth and 22.3% not reporting capability and requirements stability.

The frequency of reporting measurement information varied depending on the measurement. However, most respondents indicated that measures are reported on a weekly, monthly, or daily basis.

4.8 Summary

The long-term vision of the Software Engineering Measurement and Analysis Initiative is that organizations involved in the development and acquisition of software-intensive systems effectively and efficiently manage and improve their projects, processes, and enterprises through the use of quantitative and statistical techniques. This study is the first in a series of yearly studies that are intended to explore the state of software measurement practice. The results of this particular study can be used to provide

- an indication of what measurement definition and implementation approaches are being adopted and used by the community
- a benchmark indicating the most prevalent types of measures being used by organizations that develop or acquire software
- an indication of behaviors that are preventing the effective use of measurement (so that these barriers can be addressed)

- an indication of trends when conducted on a periodic basis (e.g., the degree to which effective measurement practices are taking root over time)
- an indication of how well measurement practices are being transitioned into use by the community

Appendix A: Survey Questionnaire

State of Software Measurement Practice Survey- 2006

Thank you for participating in the State of Software Measurement Practice survey organized by the Software Engineering Institute.

The survey comprises 17 brief questions that will take you approximately 10 minutes to complete. Please answer all questions (except #17 which is optional). **We must receive your completed questionnaire by February 28.**

Your privacy will be safeguarded and your answers will not be linked to you or your company in any way. The overall results from the survey will be presented at various conferences and through journal articles.

Upon completing the questionnaire, you will receive platinum membership to the SEI's Software Engineering Information Repository (SEIR). Platinum membership provides access to documents not available through normal access to the repository. You will also be provided early access to the survey results.

If you have any questions or difficulties with the survey please contact:

Mark Kasunic
mkasunic@sei.cmu.edu

Software Engineering Institute (SEI)
Carnegie Mellon University

When you have completed this questionnaire,
please fax to:

Mark Kasunic
Fax Number: 412-268-5758

Definition: By **software organization**, we mean an organization that develops, maintains, or acquires software or software-intensive systems.

1. Do you work for or support a software organization as we have defined it above?

- ☐ Yes
- ☐ No

Note: We have designed the survey to include individuals that work for or support a software organization. Therefore, if your response to question #1 was “No”, you do not need to complete the remainder of the questionnaire.

Write your email address on the last sheet and fax the entire questionnaire to Mark Kasunic so that you can receive your complimentary gifts for participating in this survey.

2. Please select the description that best describes the role you play in your organization. (Select a single response.)

- ☐ Executive manager
- ☐ Program manager
- ☐ Project manager
- ☐ Engineer
- ☐ Programmer
- ☐ Analyst
- ☐ Other (please specify)

3. How is your organization best described? (Select a single response.)

- ☐ Commercial shrink-wrap
- ☐ Custom software development
- ☐ "In-house" or proprietary development or maintenance
- ☐ Defense contractor
- ☐ Other government contractor
- ☐ Department of Defense or military organization
- ☐ Other government agency
- ☐ Other (please specify)

4. Please select the response that best describes your involvement with measurement. (Select a single response.)

- ☐ I am a provider of measurement-based information
- ☐ I am a user (consumer) of measurement-based information
- ☐ I am both a provider and user (consumer) of measurement-based information
- ☐ Other (please specify

5. List the country that is the primary location of your organization.

6. Indicate the number of full-time employees in your software organization. (Select a single response.)

- ☐ 25 or fewer
- ☐ 26-50
- ☐ 51-75
- ☐ 76-100
- ☐ 101-200
- ☐ 201-300
- ☐ 301-500
- ☐ 501-1000
- ☐ 1001-2000
- ☐ more than 2000

Please select a response indicating your level of agreement with the following statements.
Select a single response for each numbered item.

7. Generally speaking, I believe that using measurement-based data helps my team to perform better than without using it.

- ☐ Strongly Agree
- ☐ Agree
- ☐ Somewhat Agree
- ☐ Not sure
- ☐ Somewhat Disagree
- ☐ Disagree

- ☐ Strongly Disagree
- ☐ N/A

8. The definitions of measures that are used in my organization are commonly understood and consistent.

- ☐ Strongly Agree
- ☐ Agree
- ☐ Somewhat Agree
- ☐ Not sure
- ☐ Somewhat Disagree
- ☐ Disagree
- ☐ Strongly Disagree
- ☐ N/A

9. Select each method that you use to identify, collect, and analyze your measurement data.

(You may select more than one response.)

- ☐ I do not use a defined method
- ☐ Goal-Driven Software Measurement Method
- ☐ Practical Systems and Software (PSM) Method
- ☐ GQM (Goal-Question-Metric)
- ☐ CMMI Measurement and Analysis Process Area
- ☐ Personal Software Process (PSP) & Team Software Process (TSP)
- ☐ ISO 15939 (Information Technology – Software Engineering – Software Measurement Process)
- ☐ Other (please specify)

Please select the response that best describes you or your team. Select a single response for each numbered item.

10. There exist measurable criteria for the products and services to which I contribute.

Frequently	Occasionally	Rarely	Never	I don't know	N/A
☐	☐	☐	☐	☐	☐

11. I use measurement to understand the quality of the products and/or services that I work on.

Frequently	Occasionally	Rarely	Never	I don't know	N/A
☐	☐	☐	☐	☐	☐

12. My team follows a documented process for collecting measurement data.

Frequently	Occasionally	Rarely	Never	I don't know	N/A
☐	☐	☐	☐	☐	☐

13. My team follows a documented process for reporting measurement data to management.

Frequently	Occasionally	Rarely	Never	I don't know	N/A
☐	☐	☐	☐	☐	☐

In the following question the term, "threshold," indicates a target or boundary that when exceeded is evidence that a risk or problem exists.

14. Corrective action is taken when measurement data indicate that a threshold has been exceeded. (Select a single response.)

Frequently	Occasionally	Rarely	Never	I don't know	N/A
☐	☐	☐	☐	☐	☐

15. I understand the purposes for the data I collect or report. (Select a single response.)

Frequently	Occasionally	Rarely	Never	I don't know	N/A
☐	☐	☐	☐	☐	☐

16. Please indicate measures that you report and approximately how often they are reported by selecting a response to the right of the listed measure. (Select a single response for each row.)

	Not reported	Daily	Weekly	Bi-weekly	Monthly	Bi-monthly	Quarterly	I don't know	N/A
Capability/requirements stability (creep)	☐	☐	☐	☐	☐	☐	☐	☐	☐
Effort applied to tasks	☐	☐	☐	☐	☐	☐	☐	☐	☐
Defects identified	☐	☐	☐	☐	☐	☐	☐	☐	☐
Defects removed	☐	☐	☐	☐	☐	☐	☐	☐	☐
Code growth	☐	☐	☐	☐	☐	☐	☐	☐	☐
Schedule progress	☐	☐	☐	☐	☐	☐	☐	☐	☐
Risks identified	☐	☐	☐	☐	☐	☐	☐	☐	☐

17. If there are additional measures that are considered key to your organization which are not listed in question #16, list them in the space provided below. Please indicate how often they are reported by enclosing this information within parentheses.

[Example Response: "Time spent in peer review (Weekly)"]

1.	
2.	
3.	
4.	
5.	

Write your email address in the provided space so that we can provide you with your complimentary gifts for participating in this survey.

Your email address:	
---------------------	--

Thank you

Appendix B: Sampling Plan

The following sampling plan was used for this survey.

Sampling Frames

This study will employ three different sampling frames and compare responses among the various frames. The sampling frames include the following:

1. individuals who have contacted SEI Customer Relations and were therefore entered into their database
2. individuals who are part of the SEI Membership Program
3. individuals who have registered on the Web-based Software Engineering Information Repository (SEIR)

For each of the databases, we have chosen to use records entered from January 1, 2004 to December 31, 2005. Using the more recent records from the databases significantly reduces the number of questionnaires that are likely to be returned due to invalid email addresses.

About the Customer Relations Database

Anyone who contacts the SEI is entered into this database. Other entries are made when the SEI hosts an event such as a workshop. Individuals attending courses are not entered.

This database was started in the early 1990s. There are approximately 30,000 valid records, 6,600 of which are from individuals outside the United States.

About the SEI Member Records

SEI Member records are a subset of the Customer Relations database. SEI Customer Relations reports that these records are up-to-date and accurate. There are approximately 2000 active SEI members. About 50% of these active members are employed by the Department of Defense.

About SEIR Registrant Records

The SEIR is a Web-based repository of information related to software engineering, software acquisition, and systems engineering. Any individual may register at <https://seir.sei.cmu.edu/seir/seir-home.html> to access the information available on the Web site. The SEIR database contains about 32,865 records, entered from June 1997 to January 2006.

Customer Relations Database Preparation

The SEI Customer Relations database records were exported to an MS Excel file by staff members in Customer Relations.

The following actions were taken on this file:

1. Records from Jan. 1, 2004 through Dec. 31, 2005 were cut and pasted into another worksheet.
2. Records were eliminated from this subset if they contained
 - invalid email addresses
 - email addresses with .sei.cmu.edu extensions
 - email addresses with .andrew.cmu.edu extensions
 - email addresses with .edu extensions
 - organization names indicating affiliation with Carnegie Mellon
3. SEI Member records from the list in Step 2 were cut and pasted into a different worksheet.

SEIR Database Preparation

The SEIR registration records were exported to an MS Excel file by staff members in SEMA.

We considered the possibility that there would be redundant records in the SEIR database and the Customer Relations (CR) database. To ensure that an individual would not be contacted twice about the survey, records that appeared in both databases were removed.

Records entered between Jan.1, 2004 to Dec. 31, 2005 were cut and pasted into a new worksheet.

Database Size Used for Sampling

The following table lists the record sizes of the three distilled databases for entries made between January 2004 and December 2005.

Database	Number of Records
Individuals in Customer Relations database	6398
SEI Members	1242
SEIR Registrants	7540

Stratified Sampling

Rather than combining the records into a single sampling frame, we chose to conduct stratified sampling on each subpopulation.

Sample Size Formula

Cochran developed a formula for calculating the sample size from a population [Cochran 77].

$$n_0 = \frac{z^2 pq}{e^2}$$

where:

n_0 is the sample size

z is a point on the abscissa of the standard normal curve that specifies the confidence level

p is an estimated proportion of an attribute that is present in the population

q is equal to $(1-p)$

e specifies the desired level of precision where $e = 1 - \text{precision}$

Sample Size Calculations

This table provides sample sizes for various values of the parameters. In each case, the conservative approach of assuming maximum variation ($p=q=0.5$) is assumed.

Confidence level	z	e	Sample size
99%	2.58	0.01	16,641
99%	2.58	0.05	666
95%	1.96	0.01	9,604
95%	1.96	0.05	384

Target Sample Size

For the purposes of this study, it seems reasonable to use a sample size of 666 that provides a precision of 95% with a confidence level of 99%.

Adjustment for Small Population Size

The fpc measures how much extra precision is achieved when the sample size becomes close to the population size. When the sample size is at least 10% of the population size, then one can employ the finite population correction (fpc).

By using the fpc, a revised sample size can be calculated using the following formula:

$$n_R = \frac{n_o}{1 + \frac{(n_o - 1)}{N}}$$

where:

n_R is the revised sample size based on the fpc

N is the population size

n_o is the original sample size

The adjusted sample size for each of the subpopulations is shown in the table below.

Population	Population size	Original sample size	Adjusted sample size
Customer Relations records	6398	666	603
SEI Member records	1242	666	434
SEIR Registrants	7540	666	612

Adjusting for Invalid Records and Non-Responses

The following table was constructed to determine the actual sample size that would be needed given various response outcomes.

If the response rate is ...	Then the actual sample size should be ...		
	for CR database	for Member list	for SEIR list
100%	603	434	612
75%	804	579	816
50%	1206	868	1224
40%	1508	1085	1530
30%	2010	1447	2040
20%	3015	2170	3060

Actual Sample Size

It is difficult to predict what a response outcome might be for a given survey. For planning/estimating purposes, a 30% response rate will be assumed for this study. Therefore, the following sample sizes will be used for each subpopulation.

CR database	SEI Member	SEIR
2010	1447	2040

The SEI Member list comprises 1242 records. Therefore, the sample from that database will actually be a census (that is, all individuals will be included in the sample).

Response rates will be monitored throughout the survey access period and adjustments may be made to compensate for poor response rate.

Random Samples

Candidate respondents will be selected from each subpopulation randomly. This will be accomplished by

1. assigning a unique, randomly generated number to each record (using MS Excel's "RAND" function)
2. sorting the records based on the column containing the random number
3. selecting the records from the top of the sorted list (where the number of records selected equals the sample size for each subpopulation)

Adjustments to the Original Sampling Approach: Invalid Contact Information

When email addresses are imported into the Web-based survey tool, the tool detects and screens out invalid email addresses. Invalid email addresses were identified for all three databases.

To compensate for invalid email addresses, additional randomized records were added to meet the desired sample size numbers. This was possible for the Customer Relations database and the SEIR database. However, the SEI Member sample size was reduced in number. The final sample sizes are shown below:

CR database	SEI Member	SEIR
2010	1177	2040

The invitation emails were sent on Friday, February 10 at 1:40 pm. A large number of email bounce-backs occurred (approximately 700). These were likely due to changes in people's jobs, incorrect database entry, or self-selected changes to personal email addresses or providers.

Appendix C: Survey Response Rates and Drop-Out Rate

Figure 37 shows the response rate of each of the three groups during February. Figure 38 shows the number of respondents who dropped out of the questionnaire, by question.

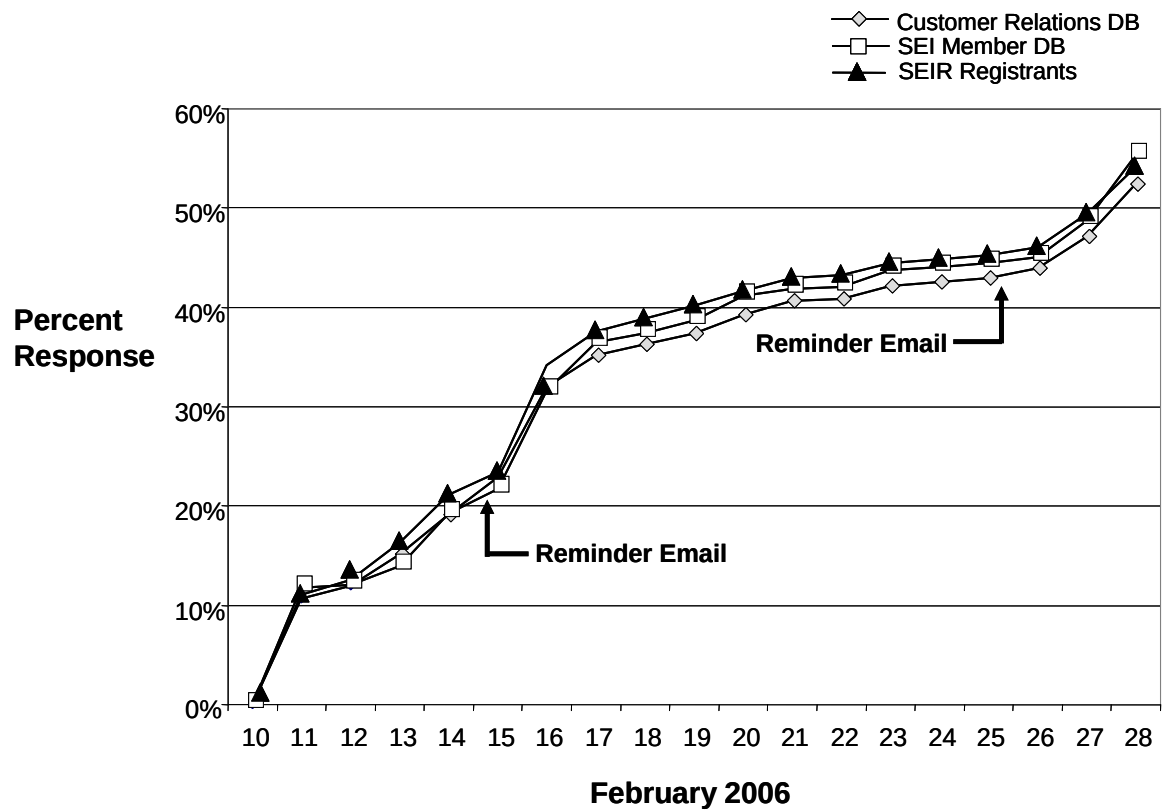


Figure 37: Survey Response Rates

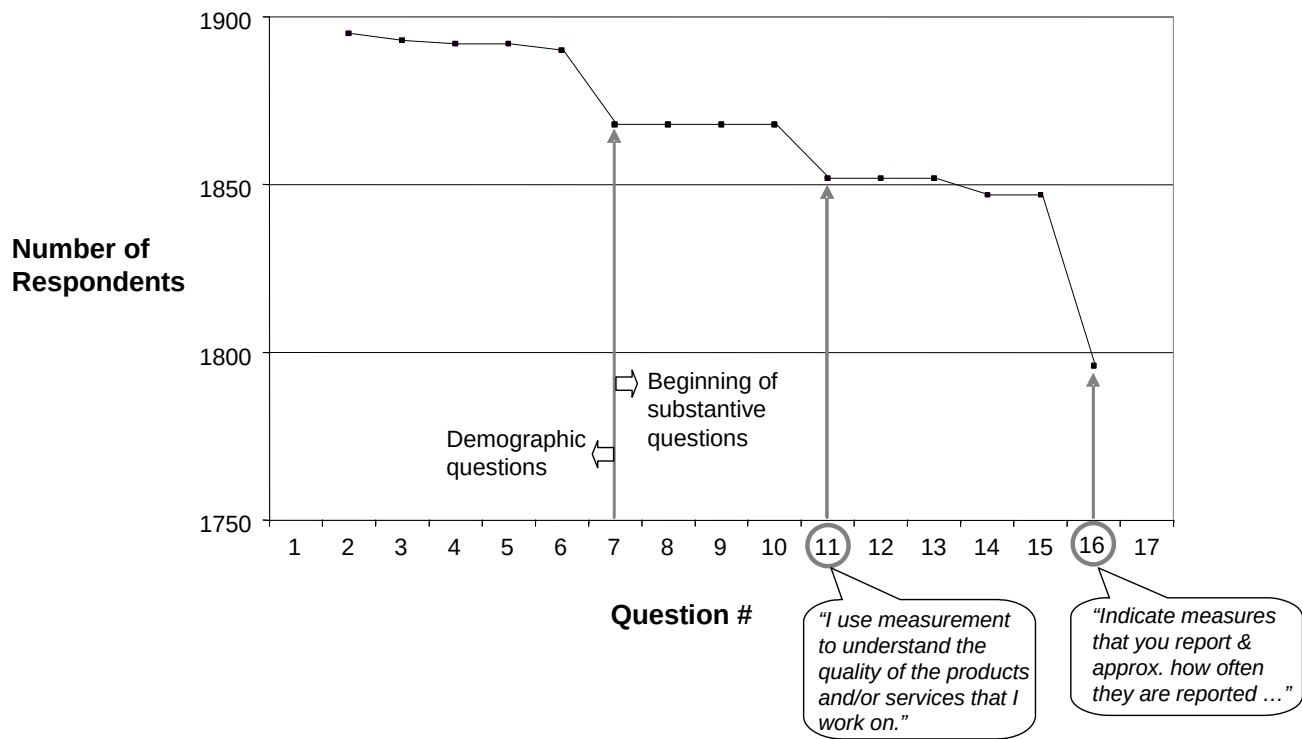


Figure 38: Drop-Out Rate by Question

Appendix D: Detailed Information About Respondents

This appendix contains graphs showing detailed information about the types of organizations represented, the jobs of the respondents, and the size of the respondents' companies.

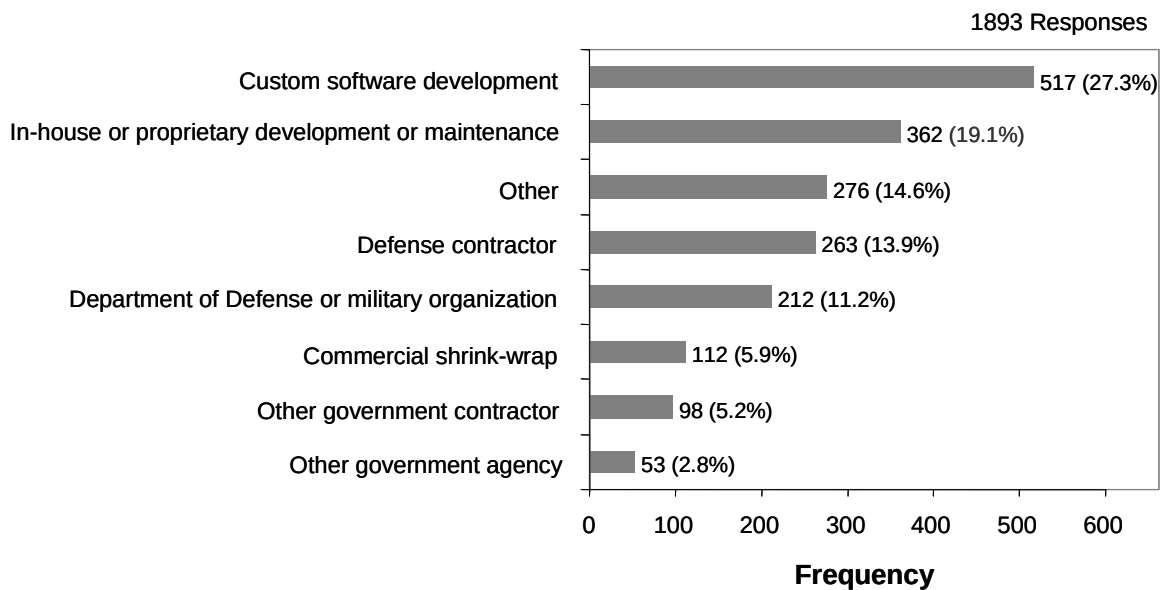


Figure 39: Types of Organizations for Which Respondents Worked

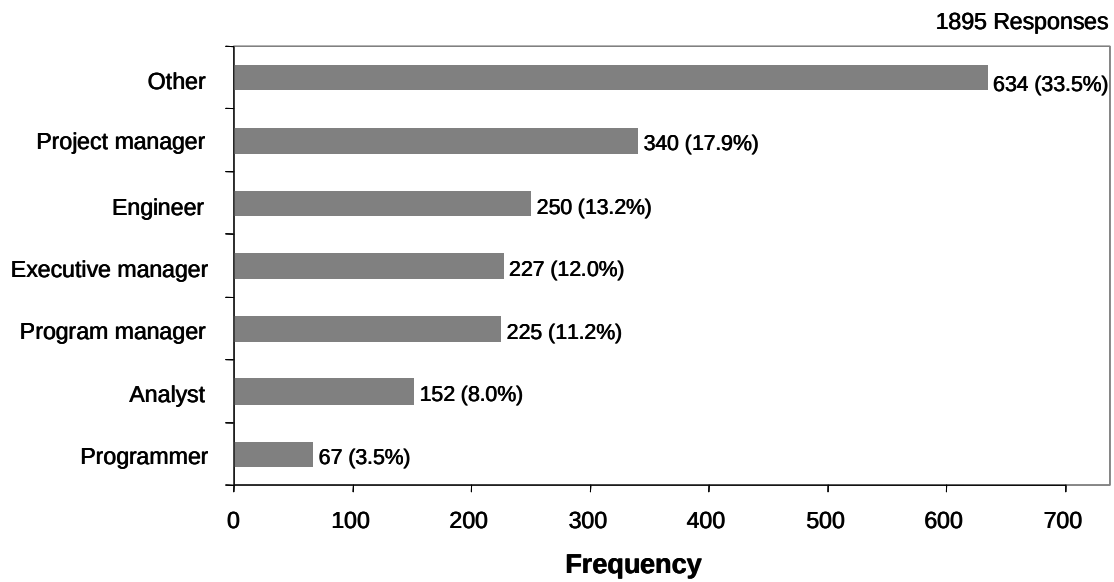


Figure 40: Types of Jobs Respondents Had

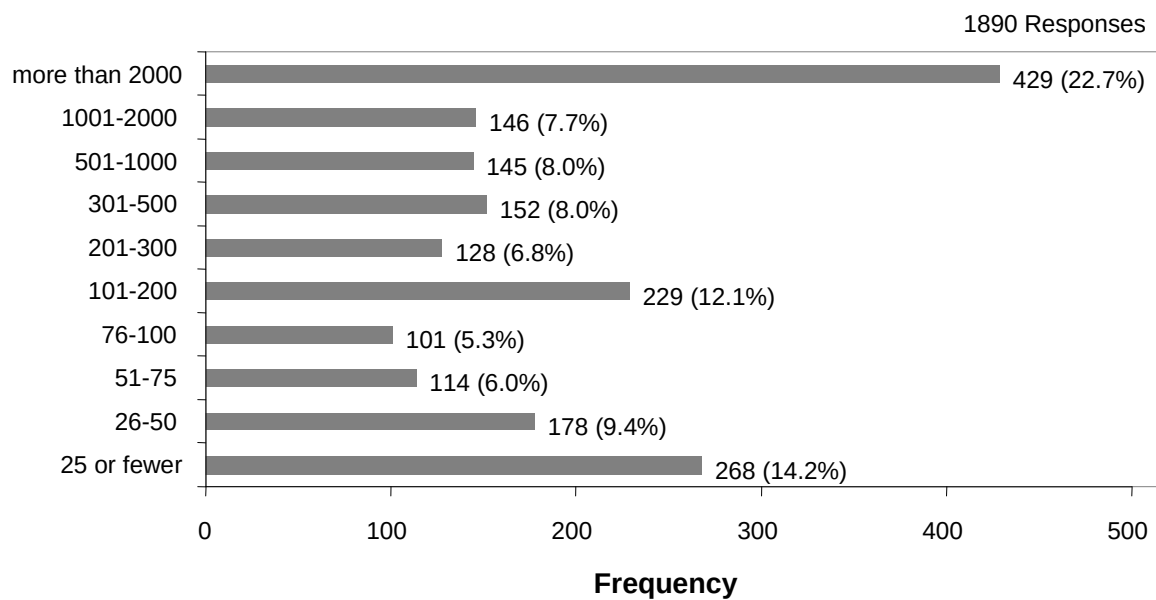


Figure 41: Number of Full-Time Employees

References

URLs are valid as of the publication date of this document.

- [AAPOR 00]** The American Association for Public Opinion Research (AAPOR). *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*. Ann Arbor, Michigan: AAPOR. <http://www.aapor.org/pdfs/newstandarddefinitions.pdf> (2000).
- [Brassard 96]** Brassard, M. *The Memory Jogger Plus + Featuring the Seven Management and Planning Tools*. Salem, NH: Goal/QPC, 1996.
- [Cochran 77]** Cochran, W. G. *Sampling Techniques*, 3rd Ed. New York, NY: John Wiley and Sons, Inc., 1977.
- [Kasunic 05]** Kasunic, Mark. *Designing an Effective Survey* (CMU/SEI-2005-HB-004, ADA441817). Pittsburgh, PA: Software Engineering Institute, Carnegie Mellon University, 2005. <http://www.sei.cmu.edu/publications/documents/05.reports/05hb004.html>.

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.				
1. AGENCY USE ONLY (Leave Blank)		2. REPORT DATE December 2006		3. REPORT TYPE AND DATES COVERED Final
4. TITLE AND SUBTITLE The State of Software Measurement Practice: Results of 2006 Survey			5. FUNDING NUMBERS FA8721-05-C-0003	
6. AUTHOR(S) Mark Kasunic				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Software Engineering Institute Carnegie Mellon University Pittsburgh, PA 15213			8. PERFORMING ORGANIZATION REPORT NUMBER CMU/SEI-2006-TR-009	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) HQ ESC/XPK 5 Eglin Street Hanscom AFB, MA 01731-2116			10. SPONSORING/MONITORING AGENCY REPORT NUMBER ESC-TR-2006-009	
11. SUPPLEMENTARY NOTES				
12A DISTRIBUTION/AVAILABILITY STATEMENT Unclassified/Unlimited, DTIC, NTIS			12B DISTRIBUTION CODE	
13. ABSTRACT (MAXIMUM 200 WORDS) In February 2006, the Software Engineering Measurement and Analysis Initiative at the Carnegie Mellon® Software Engineering Institute (SEI) conducted the first in a series of yearly studies to gauge the state of the practice in software measurement. To conduct this study, a structured, self-administered survey consisting of 17 questions was distributed to a random sample of software practitioners who had contacted the SEI during 2004 and 2005. The results of this study, which are revealed in this technical report, offer these benefits: they can be used to indicate (1) what measurement definition and implementation approaches are being adopted and used by the community, (2) the most prevalent types of measures being used by organizations that develop or acquire software, and (3) what behaviors are preventing the effective use of measurement (so that these barriers can be addressed). In addition, when the studies are conducted on a periodic basis, the results can indicate trends over time.				
14. SUBJECT TERMS software development, software measurement, state of the practice, survey			15. NUMBER OF PAGES 67	
16. PRICE CODE				
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT UL	