An exploratory study of why organizations do not adopt CMMI

Mark Staples [a,c,*], Mahmood Niazi [a,b,1], Ross Jeffery [a,c,2], Alan Abrahams [d,3], Paul Byatt [e,4], Russell Murphy [f,5]

[a] *National ICT Australia, Australian Technology Park, Eveleigh, NSW 1430, Australia*
[b] *School of Computing and Mathematics, University of Keele, Staffordshire ST5 5BG, United Kingdom*
[c] *Computer Science and Engineering, The University of New South Wales, Sydney NSW 2052, Australia*
[d] *Object Consulting Pty Ltd, Level 5, 451 Little Bourke Street, Melbourne, VIC 3000, Australia*
[e] *14 Agonis Court, Mount Martha, VIC 3934, Australia*
[f] *Dowling Consulting, Level 1, 1102 Toorak Road, Camberwell, VIC 3124, Australia*

Abstract

This paper explores why organizations do not adopt CMMI (Capability Maturity Model Integration), by analysing two months of sales data collected by an Australian company selling CMMI appraisal and improvement services. The most frequent reasons given by organizations were: the organization was small; the services were too costly, the organization had no time, and the organization was using another SPI approach. Overall, we found small organizations not adopting CMMI tend to say that adopting it would be infeasible, but do not say it would be unbeneficial. We comment on the significance of our findings and research method for SPI research.
© 2006 Elsevier Inc. All rights reserved.

1. Introduction

Software engineering researchers "seek better ways to develop and evaluate software" (Shaw, 2003). Software Process Improvement (SPI) has been a long-standing approach promoted by software engineering researchers,

---

* Corresponding author. Address: National ICT Australia, Locked Bag 9013, Alexandria, NSW 1435 Australia. Tel.: +61 2 8374 5549; fax: +61 2 8374 5520.

E-mail addresses: Mark.Staples@nicta.com.au (M. Staples), mkniazi@cs.keele.ac.uk (M. Niazi), Ross.Jeffery@nicta.com.au (R. Jeffery), alana@objectconsulting.com.au (A. Abrahams), pbyatt@pacific.net.au (P. Byatt), russell.murphy@dowlingconsulting.biz (R. Murphy).

[1] Tel.: +44 1782 583081.
[2] Tel.: +61 2 8374 5512.
[3] Tel.: +61 3 8615 4574.
[4] Tel.: +61 3 5977 0238.
[5] Tel.: +61 3 9809 1111.

intended to help organizations develop higher-quality software more efficiently. Process capability maturity models such as CMM, CMMI (Chrissis et al., 2003) and ISO/IEC 15504 (SPICE) are SPI frameworks for defining and measuring processes and practices that can be used by software-developing organizations. However, the population of organizations that have adopted process capability maturity model SPI is only a part of the entire population of software-developing organizations. Why do (perhaps most) software-developing organizations not use CMMI? Researchers have to date gathered almost no empirical evidence about this issue – almost all published articles about CMMI describe the post hoc experiences of organizations that have already chosen CMMI, who have already successfully implemented CMMI, and have already gained benefit from CMMI. Staples and Niazi (2006) discovered, after systematically reviewing almost 600 papers related to CMM(I), that apart from anecdote and expert opinion, there has been

almost no prior published evidence about the experiences, needs, and constraints of those organizations who have decided to not adopt CMMI. There has been a call to understand business drivers for SPI "...to make SPI methods and technologies more ... widely used" (Conradi and Fuggetta, 2002), and as a first step we should understand why existing SPI approaches are not chosen by many software-developing organizations. This paper presents some preliminary empirical evidence about why CMMI is not chosen by organizations.

CMMI is the successor to CMM and is consistent with the international standard ISO/IEC 15504. The most well-known representation of CMMI is the "staged" representation, which has five "levels" of process maturity for organizations. Level 2 is the first level that defines a collection of process capabilities that largely focus on supporting process areas, but also includes some project management and engineering process areas. Higher levels have all the requirements of lower levels, but add requirements for other increasingly sophisticated process areas, including those in process management. Organizations can assess their capability to operate at a specific CMMI process maturity level by undergoing an "appraisal". There are three classes of CMMI appraisal: A, B, and C. (CMMI, 2001) Class A appraisals are very costly, time-consuming, and resource-intensive, but provide the highest levels of assurance about their findings. Class B appraisals are less costly, time-consuming and resource-intensive as they use fewer appraisers, appraise fewer projects, and rely on fewer forms of evidence. Class C appraisals are the cheapest and easiest to perform, and can approach the simplicity of a structured questionnaire. All appraisals result in a report of findings about an organization's capability, but only Class A appraisals can result in a publicly-reportable "rating" of the organization's CMMI level.

CMM, CMMI, and related process capability models have been long studied. Many papers have reported the costs and benefits (Herbsleb et al., 1994; Goldenson and Herbsleb, 1995) to organizations of using process capability models for SPI, including intangible benefits (Hyde and Wilson, 2004). Some earlier papers have discussed organizational motivations for adopting these approaches (Jung et al., 2001).

There have been no studies prior to ours about why organizations have decided not to adopt process capability maturity SPI. An earlier survey (Brodman and Johnson, 1994) reported problems experienced by organizations implementing CMM, but the sample population of that survey consisted of small organizations servicing the Defense industry that were familiar with CMM, and most of which had already used CMM. Previous studies of process capability models have reported motivators (Baddoo and Hall, 2002) and de-motivators (Baddoo and Hall, 2003) for individual practiioners, but these studies did not address organizational (de-)motivation. Organizations are more than the sum of their individual members. Although within an organization decisions are made by individuals,

these decisions are made on behalf of the organization and largely in its interests, rather than purely in the interests of the individual decision-makers.

A common concern about CMM, CMMI, and related approaches is their relevance and applicability for small organizations (Brodman and Johnson, 1994). Case studies reporting small organizations' experience with CMM (Batista and Dias de Figueiredo, 2000) invariably discuss the peculiar difficulties that small organizations have of using and benefiting from CMM, and attempts have been made to provide guidance about using CMM tailored for small organizations (Paulk, 1998).

This paper reports an empirical study of reasons why software-developing organizations do not adopt CMMI. The study analyses statistically two months of sales data and contact notes collected by a company selling CMMI Level 2 Class B and C appraisal and improvement services in Australia. This represents the "entry level" for CMMI, as Level 2 is the first level with defined process areas, and Class B and C appraisals are the least expensive forms of appraisal. The offering did not have an option for Class A appraisal, but as discussed below we also excluded from our study organizations that wanted a Class A appraisal (e.g. so as to obtain a CMMI rating). We examined relationships between reasons, organizational size, and organizational type (i.e. head or branch office, of a national or multinational organization).

We have called our study an "exploratory" one because it is the first to empirically examine the question, and because it uses a methodology that is novel for software engineering research. The use of sales data and contact notes gathered by commercial organizations for empirical software engineering research has inherent threats to validity which are not yet well-known in the research community, and for which adequate controls are not yet well-established. We do attempt to describe some of these threats, how we have attempted to control these threats, and opportunities for methodological improvements in similar future studies. For our study, we believe that the remaining residual threats to validity are acceptable, because we have shed some light on an important issue in software engineering research that has not previously been empirically investigated, and because we acknowledge that our study is an exploratory one.

In Section 2 we describe our research questions and hypotheses, describe the method used to collect and analyse the sales data, report on the outcomes of reliability checks performed while classifying this data, and describe some aspects of the organizations in our study. Section 3 reports our main findings on the reasons given by organizations for not adopting CMMI, and on relationships between reasons and the size and type of organizations. Section 4 discusses how our findings relate to our hypotheses, comments on the broader significance of our findings for software engineering research, and discusses some threats to the validity of our findings. Section 5 concludes with a summary and some directions for future work.

2. Method

Our research questions and hypotheses were:

- *Why do organizations decide not to adopt CMMI?* Cost-weighted risk-adjusted benefits over time (return on investment – ROI) is often thought to motivate organizational decision-making. We believe that in a decision not to adopt CMMI, certain individual factors will be seen to dominate the ROI calculation. Specifically, we hypothesized that cost, applicability, and time to benefit would be the most frequent reasons for such not adopting CMMI.
- *How are these reasons related to the size of organizations?* We hypothesized that small and medium sized organizations would have a different pattern of response than large organizations. For example we expect that smaller organizations are more resource-constrained and require benefits to be returned within shorter periods.
- *How are these reasons related to organizational type (i.e. head or branch office, national or multinational)?* We hypothesized that there would be a significant relationship between reasons and organizational type. For example, branch offices might have less autonomy to adopt new improvement initiatives.

The following subsections describe the steps of our method.

2.1. Sales data

Three salespeople (not all full-time) methodically recorded sales data and contact notes over two months (May and June 2004) of concentrated sales effort. The three salespeople were also the consultants used to subsequently deliver the CMMI packaged services. The three consultants/salespeople were part of a 10-person company, with other employees being overhead (accountant and CEO), and training sales and marketing staff. The company's purpose for commencing on the period of effort was to contact as many companies as possible and to try to establish why, despite marketing efforts through various media, there had been limited interest in and take-up of the CMMI services. Two packaged services were offered for CMMI Level 2 appraisal and improvement. An initial approach was made for Class B services, but if the potential client had reservations about cost or effort, Class C services were offered as an alternative. The cost of these services was consistent with daily market rates for specialist consulting services (over eight days for Class B appraisals, and two days for Class C appraisals), but Class B services were also partly subsidized by a state government grant. These consultants were CMMI and SPI experts capable of leading and supporting a Class B or Class C CMMI appraisal, but they were not SCAMPI Lead Appraisers authorised to lead a Class A CMMI appraisal.

The Class B service consisted of six days of appraisal followed by two days of improvement planning. About half of the full duration of eight working days was conducted at the client organization's offices, with the remainder conducted by the consultant off-site working on analysis and reporting. The appraisal part consisted of a one day seminar for all development and management staff, followed by a series of two-to-three hour workshops for each of the seven CMMI Level 2 process areas. The workshops consisted of face-to-face interviews and document reviews relating to each process area, and involved the participation of process area specialists and/or development leaders or project managers. The appraisal reported a "gap analysis" with respect to the criteria of CMMI Level 2 which identified key strengths and areas for improvement. The improvement part included a half-day presentation to management and the development team.

The Class C service was two day's duration. It could be conducted remotely, using an on-line questionnaire supported by telephone. It resulted in an appraisal report which, like the Class B service, identified key strengths and areas for improvement relative the CMMI Level 2 process areas.

The two month sales and market research effort was run according to a verbal plan, which was followed and reviewed weekly in joint meetings of the salespeople. The contacted organizations were identified from a variety of sources: third-party company databases sourced from government and industry bodies; news articles; and the salespeoples' personal contacts. The selection of organizations was performed to maximize the likelihood of the sale of services, and so was driven by a mix of both opportunism and also methodical judgments of business opportunity. The salespeople recorded information in a spreadsheet about each contact event (phone call, meeting, fax, or email), including the date, the salesperson, the organization contacted, the individual contacted, the current and next status of the sales opportunity, and notes about the current and next status. There was no explicit or implicit performance measurement or incentive related to recording any specific reason. The raw spreadsheet contains data for 429 contact events with 274 Australian organizations. There was an average of 1.57 contact events per organization (median = 1, mode = 1), with a standard deviation of 1.09.

2.2. Categorizing sales opportunities

Using only the sales data and contact notes, two researchers independently categorized every organization as "Sale", "No Sale" or "In Progress" as at the end of the recording period. We were guided mostly by descriptive information in the contact notes, but also considered status indications ("Contact", "Prospect", or "Closed") recorded for contact events.

These two independent categorizations of organizations had 95% agreement, and using Cohen's Kappa (Cohen, 1960) achieved a satisfactory level of inter-rater reliability (kappa = 0.88, which is greater than commonly applied minimum criteria of 0.70). The 13 individual disagreements

were resolved in a joint meeting between two researchers. Of the 274 organizations, 73 were finally categorized as being "No Sale" organizations. (Most other organizations were categorized as being "In Progress".)

2.3. Classifying reasons

Using only the sales data and contact notes for "No Sale" organizations, two researchers independently derived a list of reasons for not buying CMMI Level 2 Class B or C appraisal and improvement services. These reasons were obtained by quoting or paraphrasing text present in the sales data and contact notes. The two researchers derived lists of 15 and 12 reasons respectively, with initial agreement on 6 reasons. The differences were resolved in a joint meeting of the researchers, resulting in the final list of 16 reasons as shown in Table 1.

Using this list of reasons, the two researchers then independently classified the reasons given by every "No Sale" organization, and in a joint meeting agreed on the list of reasons given by each organization. Note that organizations could give multiple reasons, and so in Table 1 the number of reasons shown is greater than the number of organizations.

2.4. Organizations under analysis

Our research questions concern decisions made by organizations not to adopt CMMI. As such, we only consider organizations that were a "No Sale" within the sales period, i.e. we excluded "Sale" and "In Progress" organizations. There were 73 "No Sale" organizations, but not all of these were relevant to our research questions, as described below.

We first excluded 24 organizations whose reasons included either "No software development", "Not in business", "Already using capability maturity model", or "Want Class A rating", i.e. gave any reason in the "N/A" reason group of Table 1. These organizations are not within the target population of our research questions, which concern software-developing organizations that neither use process capability maturity model SPI such as CMM, CMMI, or SPICE, nor want CMMI Class A appraisal services. Note that the organizations found to be "not in business" were initially contacted because of their presence in an industry database that was the source of some of the contacts used in the sales process.

Of the remaining 49 organizations, we excluded 9 that gave only one or both of the reasons "May consider SPI later" or "No interest" (i.e. gave only reasons in the "Insubstantive" reason group of Table 1). These "reasons" have no power to explain the rationale behind an organization's decision. On one hand, these reasons may represent a missed opportunity for further probing into organizational decision-making, but on the other hand such probing in this setting is unlikely to provide an accurate response from people who may simply want to "get rid" of the calling salesperson.

We were left with a final sample population of 40 software-developing organizations neither using process capability maturity model SPI nor wanting a CMMI Class A appraisal, that within the two month recording period decided not to buy CMMI Level 2 Class B or C appraisal and process improvement services, and gave at least one substantive reason for their decision.

When organizations make a decision about adopting a new practice, they must consider "Can we use it?" (i.e. is it feasible that the organization could adopt the practice), and "Should we use it?" (i.e. is there a compelling business case for us to adopt the practice). We grouped the reasons given by organizations for not buying these services into categories of "we could not use it" and "we should not use it", as shown in Table 1. Note that the multiple reasons given by organizations could span or be contained within each of these categories.

2.5. Determining size and type

For each of the 40 organizations to be analysed, a researcher determined both their size (i.e. number of employees) and type (i.e. head or branch office, national or multinational) using various sources of information: the sales data and contact notes; the organizations' websites; and industry databases.

When determining size, we established the number of employees for each organization, and categorized them as SMALL, MEDIUM, or LARGE. Our definition of size

Table 1
Grouping reasons from all "No Sale" organizations into "could not" and "should not"

Group	Group frequency (of 73 orgs)	Reason	Reason frequency (of 73 orgs)
N/A	24	No software development	11
		Not in business	8
		Already using capability maturity model	3
		Want Class A rating	2
Insubstantive	37	No interest	21
		May consider SPI later	17
Could not	33	Small organization	17
		Too costly	15
		No time	10
		Not applicable to our projects	2
Should not	17	Using other SPI	8
		No clear benefit	4
		Potential benefits not wanted	3
		Already know gaps	2
		No customer demands	2
		Risk of poor certification damaging business	1

Table 2
Definition of organizational size categories in terms of number of employees

Size category	More or equal to	Less or equal to
SMALL	0	19
MEDIUM	20	199
LARGE	200	+

categories is given in Table 2, and is based on the Australian Bureau of Statistics definition of small and medium enterprises. (Trewin, 2002) Note that our measure of size is of the organization that was the potential client for the services. That is, the entire business unit, but not specifically the (smaller) size of the software-developing part of the business unit. Thus we may have mistaken small software development groups for larger ones, but are unlikely to have mistaken large software development groups for small ones. See Section 4.4 for a discussion on this and other threats to validity in our study related to the size of organizations.

When we could not establish the number of employees, we categorized the size as UNKNOWN. Based on our searches, we speculate that most of the UNKNOWN organizations would in fact be SMALL, with perhaps a few MEDIUM. The summary statistics of size categories in terms of number of employees for the 40 organizations under analysis is given in Table 3.

Note that in our sample there are organizations on the "breaks" between size categories. We have conducted a sensitivity analysis on every test of significance of relationships between size categories and other attributes, and have only reported a relationship to be significant if it was not sensitive to the categorization of organizations on these break-points between categories.

The adoption of an organizational process such as CMMI can sometimes be mandated by higher parts of the organization removed from the business unit that must adopt the process. We were interested to know how organizational type affected the reasons given for not adopting CMMI. When determining organizational type we established the location of the head office, whether organizations had multiple office locations, and whether the organization had international offices. Using this information we classified organizations according to the definitions in Table 4. This table also shows the number of each type of organizations under analysis.

Table 3
Summary statistics of size in terms of employee numbers for organizations under analysis

Size category	N	Minimum	Median	Mean	Maximum	Std Dev.
SMALL	15	1	5	7.8	18	6.53
MEDIUM	9	20	45	48.3	95	28.6
LARGE	4	200	245	313	560	169
UNKNOWN	12	?	?	?	?	?

Table 4
Definition of organizational type categories, and number of each type of organization under analysis

Type	Definition	Number
Single unit	Single office location; No international offices	22
Head of national	Multiple office locations; Head office contacted; No international offices	6
Branch of national	Multiple office locations; Branch office contacted; No international offices	0
Head of multinational	Multiple office locations; Head office contacted; International offices	9
Branch of multinational	Multiple office locations; Branch office contacted; International offices	3

3. Analysis and results

3.1. Reason frequency

The frequency-ordered list of substantive reasons for not buying CMMI Level 2 Class B or C appraisal and process improvement services given by the 40 organizations under analysis appears in Table 5. Each organization could give many reasons.

For the reason "Using other SPI", three organizations said they used ISO9001, three organizations said they used agile approaches, one organization said they used the Microsoft Framework, and one organization did not specify a method.

When we group reasons as shown in Table 1, and count the number of organizations who gave a reason in either or both of the groups "could not" or "should not", we get Table 6. (Note that Table 6 shows number of organizations, whereas Table 1 shows number of reasons. One

Table 5
Number of organizations under analysis giving each reason

Reason	Frequency (of 40)	%
Small organization	17	43
Too costly	14	35
No time	10	25
Using other SPI	8	20
No clear benefit	4	10
Potential benefits not wanted	3	8
No customer demands	2	5
Not applicable to our projects	2	5
Already know gaps	2	5
Risk of poor certification damaging business	1	3

Table 6
Number of organizations under analysis giving "could not" reasons, "should not" reasons, or both sorts of reason

Reason group	Frequency (of 40)	%
Only "could not" reason(s)	23	58
Only "should not" reason(s)	8	20
Both "could not" and "should not" reasons	9	23

organization only gave two "could not" reasons, which counts twice in Table 1, but only once in Table 6.)

3.2. Reason vs. size

A chart of the percent of each size category giving each reason is shown in Fig. 4. We used Fisher's exact test (using FEXACT (Mehta and Patel, 1986) in the R package, version 2.1.1 (R, 2005)) to test if there was a significant relationship for organizations between their size and the reasons they gave for not adopting CMMI. We have reported the significance of relationships excluding the UNKNOWN size category, but for every significance test involving size categories we have conducted a sensitivity analysis under four other treatments: leaving UNKNOWN as its own category; and adding UNKNOWN to each other size category. We have commented in this paper on any difference in the resulting judgments. When examining these relationships, we first looked at the overall relationship between size categories and reasons. To understand the source of any significance, we subsequently collapsed SMALL and MEDIUM categories to focus on reasons given by LARGE organizations. Similarly, we collapsed MEDIUM and LARGE categories to focus on SMALL organizations.

We tested the significance between size and each of the four most frequent reasons: "small organization", "too costly", "no time", and "using other SPI". There were no significant associations except for the reason of being a "small organization" (Fisher's exact test, $p = 0.022$).[6] However, this reason is inherently not independent of size, and so this significance is not surprising. Nonetheless, 67% of the SMALL organizations gave this reason, compared to only 15% of MEDIUM + LARGE organizations. Specifically, 10 had fewer than or equal to 8 employees, 1 had 20 employees, 1 had 520 employees (we suspect this was a large organization with only a small software development team), and 5 were UNKNOWN. This partly validates our definition of SMALL organizations as those being less than 20 employees. More importantly, it shows that small organizations tended to consider their small size to be a barrier to the adoption of CMMI. However, it is not clear if this is because of cost, resource constraints, or because they consider CMMI to be inapplicable to small organizations.

For the grouped reasons "should not" and "could not", a chart showing the percent of organizations for each size category giving reasons in each group is shown in Fig. 1.

Overall, there was a significant relationship between the size of organizations and giving only "should not", only "could not", or both sorts of reasons ($p = 0.001$, Fisher's exact test). For "should not" reasons, there was a significant overall relationship to size ($p = 0.001$, Fisher's exact test). When we collapsed MEDIUM and LARGE catego-
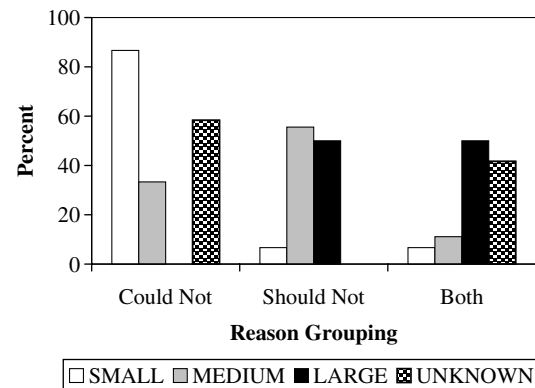

Fig. 1. Percent of organizations under analysis of each size group giving reason(s) in either or both group.

ries, we saw a significant (negative) relationship between SMALL organizations and "should not" reasons ($p = 0.002$, Fisher's exact test). Only 13% of the SMALL organizations gave this sort of reason, compared to 77% of MEDIUM + LARGE organizations. However, when we collapsed SMALL and MEDIUM categories, we could not claim there is a significant relationship between LARGE organizations and "should not" reasons, as the judgment of significance provided by Fisher's exact test was sensitive to the values at the breaks between our size categories.

For "could not" reasons, there was a significant overall relationship to size ($p = 0.017$, Fisher's exact test).[7] When we collapsed MEDIUM and LARGE categories, we saw a significant (positive) relationship between SMALL organizations and "could not" reasons ($p = 0.011$, Fisher's exact test).[8] 93% of the SMALL organizations gave this sort of reason, compared to 46% of MEDIUM + LARGE organizations. However, when we collapsed SMALL and MEDIUM categories, we could not claim there was a significant relationship between LARGE organizations and "could not" reasons, as the judgment of significance provided by Fisher's exact test was sensitive to including or excluding UNKNOWN as a category, and when adding UNKNOWN to SMALL + MEDIUM was sensitive to the breaks between our size categories.

3.3. Size vs. type

Before we discuss the relationship between reason and type in Section 3.4, we will first see how size and type are related. (Section 4.1 discusses all of these relationships.) The distribution of organizations under study as a percent of each of our size categories for each organizational type is shown in Fig. 2.

---

[6] This is sensitive to adding UNKNOWN to MEDIUM for our primary definition of breaks between our size categories. We do not take this limited sensitivity to invalidate the main judgment.
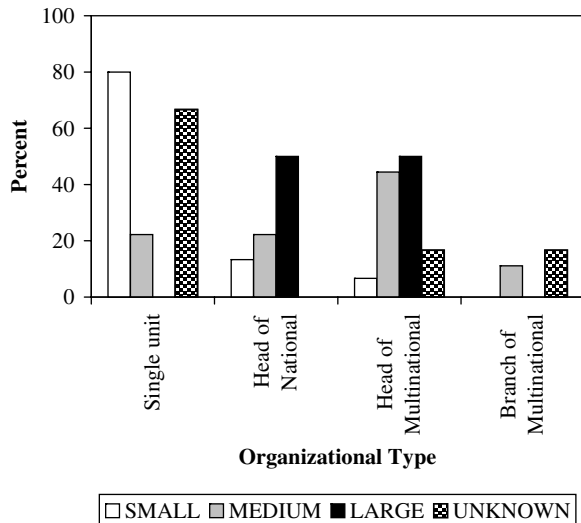
[7] This is sensitive to the treatment of adding UNKNOWN to MEDIUM. We do not take this partial sensitivity to invalidate the main judgment.

[8] This is sensitive to adding UNKNOWN to LARGE + MEDIUM. We do not take this limited sensitivity to invalidate the main judgment.

Fig. 2. Percent of organizations under analysis of each size for each type.

Fig. 3. Percent of each organizational type for organizations under analysis giving reason(s) in either or both group.

Size and type are not independent, e.g. because an organization with many branches must also have many employees. Indeed, we did find a significant relationship between size and type ($p = 0.004$, Fisher's exact test).[9] When we collapsed MEDIUM and LARGE categories, we saw a significant relationship between organizational type and SMALL organizations ($p = 0.002$, Fisher's exact test).[10] 80% of the SMALL organizations are "single unit" types, whereas 77% of the MEDIUM + LARGE organizations are instead head offices of either national or multi-national organizations. However, when we collapsed SMALL and MEDIUM categories, we could not see a significant relationship between organizational type and LARGE organizations ($p = 0.087$, Fisher's exact test).

3.4. Reason vs. type

We first note that as discussed in Section 3.3 above, size and type are not independent, and as 80% of SMALL organizations are "Single unit" organizations, there is a confounding effect whereby the reasons given by SMALL organizations will also be seen for "Single unit" organizations.

The distribution of organizations by organizational type for each reason is given in Fig. 5. We tested the significance between organizational type and each of the four most frequent reasons: "small organization", "too costly", "no time", and "using other SPI". We did not find that any of these relationships to individual reasons were significant.

The distribution of organizations by organizational type for each of the grouped reasons is given in Fig. 3. We did not find a significant association between organizational
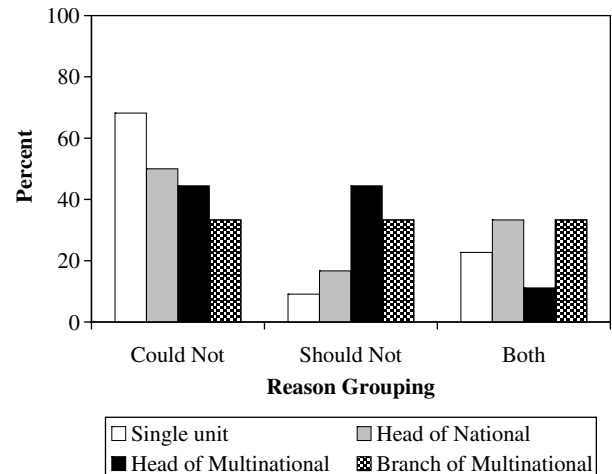
type and giving only "should not", only "could not", or both sorts of reasons ($p = 0.441$, Fisher's exact test). Moreover, we did not find a significant association between organizational type and giving either "should not" ($p = 0.496$, Fisher's exact test) or "could not" reasons ($p = 0.102$, Fisher's exact test).

4. Discussion

4.1. Hypotheses

Why do organizations decide not to adopt CMMI? We had hypothesized that cost, applicability, and time to benefit would be the most frequent reasons for such decisions. The sales data showed that the most common reasons were: the organization was small; the services were too costly, the organization had no time to use the services, and the organization was already using other SPI approaches. The reason of being a "small organization" may be a reflection of perceived inapplicability of CMMI, or it may be a reflection of constrained budget or available time. Under either interpretation it provides some support for our hypothesis. Our hypothesis concerning cost as a common reason was also supported by our data. "Time" was revealed as a common concern, but instead of our hypothesized length of time to benefit, the common reason was the lack of available time to adopt CMMI. Overall, organizations do not adopt CMMI because they think that it would be infeasible to do so, i.e. they give "we could not" reasons. One common reason that we had not foreseen was the competition from other forms of SPI already being used within organizations. We think it likely that these were seen either to satisfy the same needs as CMMI or to be incompatible with it.

How are the reasons given by organizations related to their size? We hypothesized that small and medium sized organizations would have a different pattern of response than large organizations. Our analysis shows instead that SMALL organizations have a different pattern of response

---

[9] This is sensitive to adding UNKNOWN to either LARGE or MEDIUM. We do not take this partial sensitivity to invalidate the main judgment.

[10] This is sensitive to adding UNKNOWN to MEDIUM + LARGE. We do not take this partial sensitivity to invalidate the main judgment.
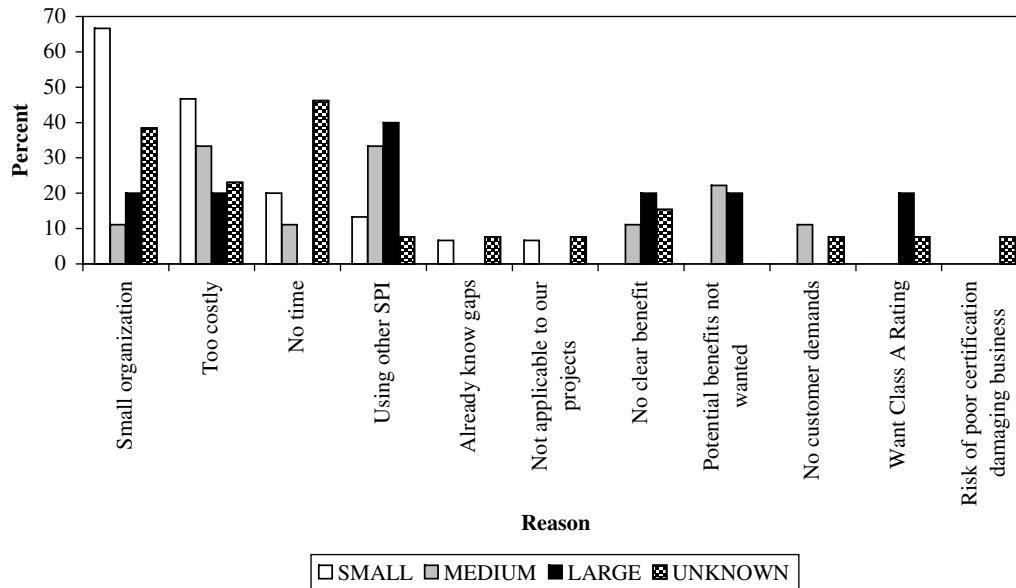
Fig. 4. Percent of each size group for organizations under analysis giving each reason.
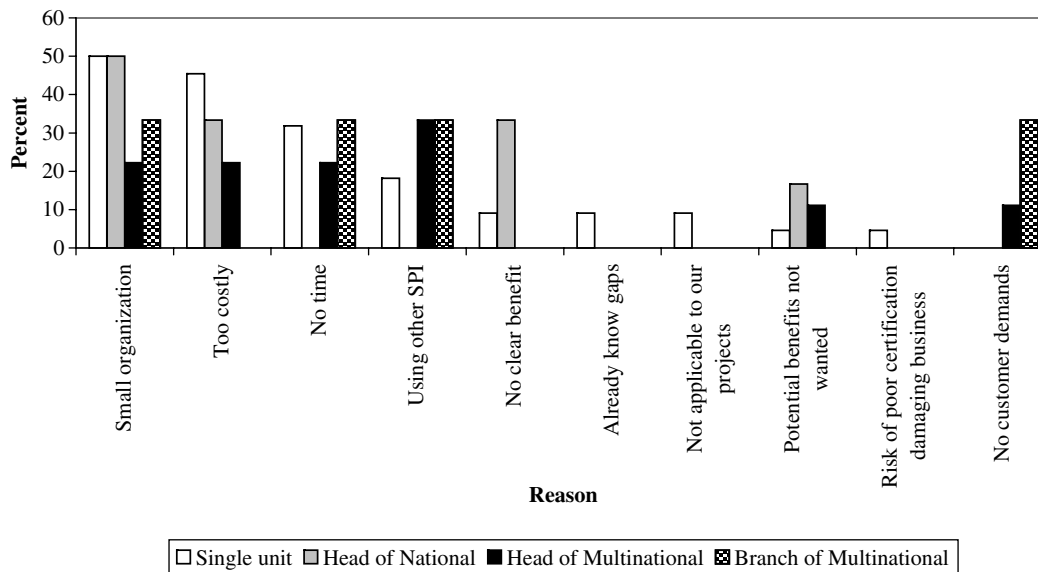

Fig. 5. Percent of each organizational type for organizations under analysis giving each reason.

than MEDIUM and LARGE organizations. Specifically, SMALL organizations tend to give "we could not" reasons, and tend not to give "we should not" reasons, whereas MEDIUM and LARGE organizations tend to give "we should not" reasons. Certainly, to the extent that CMMI adds cost overheads, these would be more readily absorbed by larger organizations with greater economies of scale. However, the "no time", "small organization", and "not applicable to our projects" reasons within the "could not" reason group may indicate that small organizations face other additional barriers than simply being unable to absorb additional cost overheads.

An earlier survey (Brodman and Johnson, 1994) found that small organizations (especially those with fewer than

20 employees) do not have the resources or funds to implement CMM, and that many of the practices of CMM are not applicable in a small context. Our high-level findings are in broad agreement with this, despite our study being over a decade later, about a successor to CMM, with surveyed organizations in a different country, servicing different industries (not Defense), and with different levels of awareness about capability maturity models. This agreement provides support for our findings.

How are reasons given related to organizational type (i.e. head or branch office, national or multinational)? We hypothesized that there would be a significant relationship between reasons and type. Moreover, as we found relationships between organizational type and size, and between

size and reason, we might have expected the transitive relations between type and reason to also be significant. However, we could not conclude that these relationships were significant. Perhaps the factors influencing the relationships between size and reason are different to those influencing the relationships between type and reason.

Some previous studies (Jung et al., 2001) have used a definition of "small enterprise" as those with 50 or fewer employees. Under that definition, we retain significant relationships between size and giving reasons in one or both of the "should not" or "could not" groups ($p = 0.020$, Fisher's exact test), but not between size and giving the "small organization" reason ($p = 0.185$, Fisher's exact test). In light of this, the distribution of number of employees in organizations giving the "small organization" reason noted in Section 3.2, and our agreement with an earlier survey (Brodman and Johnson, 1994), we consider our definition of SMALL organizations as those with less than 20 employees to be appropriate.

4.2. Improving SPI for small organizations

How can SPI researchers improve SPI, and in particular make SPI more accessible and beneficial for small organizations? The preliminary evidence from our study on CMMI may be able to inform CMMI researchers and perhaps SPI researchers more generally. To benefit from improved practices, organizations must first adopt them. Our study shows that many software-developing organizations will never gain any benefit from process capability maturity SPI because they consider it infeasible to adopt. We found that this was especially true of small organizations, which constitute the majority of software-developing organizations. To what extent are these negative perceptions of CMMI justified? Certainly some small organizations (Batista and Dias de Figueiredo, 2000) have successfully adopted CMM-based SPI. However, the perceptions revealed in our study are broadly similar to the perceptions revealed in an early study (Brodman and Johnson, 1994) of organizations most of which had tried to adopt CMM. Moreover, many authors have previously raised concerns about the applicability of CMMI-based SPI for small organizations. So although there is a lack of studies of failed adoption of CMMI-based SPI, this is no doubt due to inevitable publication bias in the field. (Organizations heavily invested with CMMI will be less commercially motivated to publicly expose failures associated with CMMI than successes.) Further justification for the negative perceptions about CMMI held by small organizations comes from the lack of widespread adoption of process capability maturity SPI among small organizations, despite the decades of existence of CMM.

While CMMI admits tailoring, the emphasis is on tailoring "down" by demanding justification for excluding defined aspects of the model, rather than on tailoring "up" by starting with minimal practices and adding to them as necessary to address specific issues. Wilkie et al. (2005) have found that different CMMI process areas have different perceived value. CMMI tends to suit the requirements of large organizations – small organizations need CMMI to be more tailored to suit, which creates additional burdens for them. Ironically, small organizations are likely to have fewer resources available to do tailoring.

How do the needs of small organizations differ from larger ones? Wilkie et al. (2005) have found that small organizations have a focus on product quality assurance rather than on process quality assurance. However, we speculate that small organizations also have more specific interest in product quality ("the product working well") than product quality assurance ("evidence of the product working well"). Many of the CMMI process areas and goals are concerned with process quality and quality assurance, which we believe small organizations perceive as costly and time-consuming lower-value activities. We hypothesize that small organizations are more variable than larger ones in their business context, needs, and constraints. The extent and nature of variability in small software development organizations is not yet well understood, and neither is the issue of how SPI frameworks can be created to best help small organizations.

Much SPI research gathers evidence about the relative costs and benefits of adopting SPI. However if the cost and time to adopt an approach are perceived as low enough, it may not be important to establish convincing prior evidence of benefit. Organizations will be able to inexpensively trial such approaches directly and observe any benefit in their own context for themselves. We speculate that recent interest in Agile methodologies is driven because they are perceived to target issues of greatest concern to small organizations (product quality instead of process quality or quality assurance) and are perceived to required low cost and time to trial and adopt.

We speculate that radically recasting CMMI for the needs of small organizations may make it more widely successful in the industry, especially if there is a focus on issues of immediate concern to small organizations. We suggest that in order to broaden the impact of SPI research, researchers must ensure that these common barriers that we have found are addressed: i.e. small organizations must be the target audience, and the adoption of SPI approaches must require radically less cost and time.

4.3. Sales data for empirical research

The data for this study was collected as part of a methodical sales and market research exercise, and was not collected using a survey instrument designed to support scientific study. This has resulted in more threats to internal validity than would be desirable (see Section 4.4). We have persevered because this is very rare data which sheds light on important but previously unaddressed questions.

Sales data and market research data may be a significant and largely untapped source of data for the software

engineering research community, especially to help understand issues related to the adoption of new technologies, technology transfer, and industry perceptions. Sales and market research data has distinct benefits for use in empirical academic research: it should generate large volumes of data; and it is supported by commercial outcomes and so does not require massive academic investment. Sales data and market research data has its own kind of validity in that it is collected in a "real world" context; its collection is driven by simple commercial interests that will make apparent the likely direction of any bias; and it is often gathered with an explicit sales or market research strategy defined to satisfy an explicit commercial goal.

Sales and market research have different goals and methodologies than academic research, but they may not necessarily be mutually inconsistent. The research community should try to find methodologically "safe" incentives to get access to and use the findings of market research. Companies may find it commercially or politically risky to report negative findings, while positive findings may reflect a competitive advantage that companies want to retain privately for themselves. So, companies may understandably feel reluctant to let aspects of their sales or market research data be analysed for public review. Anonymous reporting and a limited-term embargo on the publication of findings may ameliorate this concern. Companies may see many benefits working with academic researchers in this way: they may receive a deeper and more rigorous analysis of their data than they would otherwise be able to undertake; they will inevitably be provided with early access to the research results; and academic research may contribute to "growing the pie", rather than simply the company's individual "slice of the pie" in the market.

Even if companies are prepared to let their internal data be used for academic research, it is not yet clear how the academic community will accept the results of research based on such data. Companies and their employees are driven by commercial and private incentives that can bias the collection and use of sales and marketing data. The use of sales and market research data within the software engineering research community is not yet well understood. There are neither guidelines for improving the design of commercially realistic sales and market research instruments to reduce threats to internal validity for academic research, nor for understanding the generalization of empirical findings based on sales and market research data. Our study points out some validity issues that should be considered in the development of such guidelines. We believe that such guidelines may share similarities with guidelines for survey research, e.g. such as (Fowler, 2002). For example, as with surveys, data validity could be checked and improved by conducting follow-up interviews.

4.4. Threats to internal validity

The major internal validity threats and their partial controls are listed here.

We have used "buying" CMMI appraisal and improvement services as a proxy for "adopting" CMMI. These two notions are not in exact correspondence, but do overlap and are affected by similar issues. Our findings about responses to proposals for "out-sourced" CMMI services may not apply to organizations that want to perform in-house CMMI. This threat is partly controlled by our exclusion of organizations that said they were already using process capability maturity SPI. Related to this, the specific services and the nature of the company offering the services may have influenced the sort of responses that were given. For example, it is possible that another company offering similar services might have converted some of the sales opportunities into sales by being able to effectively address a specific sort of excuse, and thus the associated "reason" might not appear so frequently.

Organizations might be tempted to give inaccurate reasons that reflected better on them that their real reasons. This threat is partly controlled by allowing organizations to give the "Insubstantive" responses – organizations were not forced to invent meaningful reasons. Related to this, it is possible that the reasons recorded by the salespeople may not have been recorded accurately if there had been a potential for the various different reasons to be used as a measure of the salespeople's performance. However, there was no such incentive provided by the company for the salespeople/consultants.

Size data for some organizations in our study may be up to a year out-of-date, and it is also possible that in some cases the size of entire multi-site organizations may have incorrectly been given for individual sites. We also have no good data on the number of employees directly involved with software development within these organizations. These threats are partly countered by our analyses being only in terms of the very coarse categories of SMALL, MEDIUM, and LARGE. Also, our categorization errors (if any) arising from either mistaking the number of employees at an individual site for those of their entire multi-site organization, or from not distinguishing software development teams from other employees are both likely to lead us to mistake smaller software development organizations for larger ones. Such errors would only incorrectly weaken the significance of relationships we have discovered concerning SMALL organizations – our findings of significant relationships are conservative in this sense.

In an idealized setting, organizational decisions are cost vs. benefit judgments. From this perspective the reasons "too costly" and "no time" should be moved to the "should not" group of reasons. In principle, if the likely return is high enough, organizations could in a perfect market simply acquire new debt or investment to acquire new resources to overcome these "barriers". However, we reject this simplistic theoretical perspective as being invalid in practical terms for any individual organization make a decision about adopting SPI within a short-to-medium timeframe, and especially for the resource-constrained small organizations which were a large part of our study.

There is a maturation threat that could plausibly affect the kind of reasons given by respondents: i.e. the age of an organization may be a better predictor of reasons than its size. There is a history threat that could plausibly affect perceptions of CMMI and thus the reasons given by respondents: we have not controlled for previous exposure or awareness of CMM, CMMI, or related process capability maturity models. This is only partly treated by our exclusion of organizations that said they were already using process capability maturity SPI.

4.5. Threats to external validity

It is possible that our findings do not generalize to other contexts. Our study uses neither a random sample of organizations nor an exhaustive survey. However, as discussed in Section 2.1 the number of contacts was fairly large, and the contacts were drawn from several sources, including large industry databases. Nonetheless there may be hidden systematic bias affecting the valid generalization of our results.

For example, we could not meaningfully identify the distribution of industries serviced by organizations in our study. (However, we believe few if any organizations serviced the Defense industry.) Also, organizations in our study are all Australian – our findings may not generalize to organizations in other countries. However as noted in Section 4.1, the agreement of our high-level findings with an earlier survey (Brodman and Johnson, 1994) conducted in another country (USA) with organizations largely in another industry (Defence) provides support to control for this external validity threat. Independent replication studies could address this threat.

5. Conclusion

5.1. Summary and conclusions

Our study has proposed new research questions and new ways to investigate these questions. We investigated why organizations do not adopt CMMI. Our study is the first to have investigated the population of organizations that have chosen to not adopt CMMI, and their reasons for not adopting CMMI. Our sample population consisted of 40 software-developing organizations neither using process capability maturity model SPI nor wanting a CMMI Class A appraisal, that within a two month recording period decided not to buy CMMI Level 2 Class B or C appraisal and process improvement services, and gave at least one substantive reason for their decision. Their most common reasons were: being a small organization, it being too costly, having no time, and already using another form of SPI. We found a significant relationship between size and the reason of being a "small organization". In particular, SMALL organizations (less than 20 employees) were associated with this reason. This partly validates our definition of SMALL organization, but more importantly shows that SMALL organizations considered themselves too small for CMMI. We believe that this was based on concerns either about the perceived applicability of CMMI or about constrained resources (available budget or time). The size of organizations was related to the sorts of reasons they give. Small organizations were (positively) associated with saying that adopting CMMI would be infeasible ("could not" reasons), and (negatively) associated with saying that adopting CMMI would be undesirable ("should not" reasons). We found no significant relationship between an organization's type (i.e. head or branch office, national or multinational) and the reasons they gave, despite there being a significant relationship between size and organizational type. The factors influencing the relationship between size and reason may be different to those for type and reason.

To investigate our research question, we analysed sales and market research data collected by a company selling CMMI services. Where this data can be obtained for use in software engineering research, the use of such data is attractive for many reasons, but nonetheless carries with it threats to validity that we have described. The valid collection, analysis and interpretation of such data is not yet well understood within the software engineering research community. Because of potential validity threats we have noted, partly due to the novelty of our research methodology, we regard our study as an exploratory one. Although our results must be read with some caution, we believe that we have established a new level of evidence about research questions and populations of software-developing organizations that have not been previously studied.

5.2. Future work

We encourage independent studies on this topic. This could increase confidence in our preliminary findings and also track changes in attitudes to SPI over time. We encourage further research into the extent and nature of variation of business context, needs, and constraints of smaller software-developing organizations, and their differences to larger organizations. We believe that a good understanding of these issues is vital in improving and creating effective SPI approaches.

To increase the impact of software engineering research into SPI, it is not enough to increase the benefits of SPI. Smaller organizations are the most numerous software-developing organizations, and many of them will not gain any benefit from process capability maturity SPI because they consider it infeasible to adopt. Small organizations should not be seen as being "at fault" for not adopting CMMI – instead the SPI approaches, its transition mechanisms, and sales and marketing should be improved. We believe that to achieve broader impact, SPI approaches must target small software-developing organizations and require very little cost and time to adopt. Radically different approaches to SPI may be required. One size may not fit all, and it is possible that research into "tailoring"

existing process capability maturity models may not be sufficient.

For any future studies that may use sales data to investigate these research questions, we encourage researchers to collect additional data to address some of the validity threats that we have noted in our study: record the number of employees in the software development team, their local office, and also the entire organization; record the age of the organization; and record the industries serviced by the software development team. Of course collecting any additional data should not compromise the commercial effectiveness of sales activities! We believe commercial sales and market research data is an under-utilised source of empirical data for software engineering research. We would like to see the academic community develop guidelines and norms for the use of such data, suggestions for the design of sales and market research instruments to better support empirical research, and the creation and recognition of methodologically "safe" ways to use the data and to provide incentives for companies to contribute their sales data to empirical researchers.

Acknowledgements

References

Baddoo, N., Hall, T., 2002. Motivators of software process improvement: an analysis of practioners' views. The Journal of Systems and Software 62, 85–96.

Baddoo, N., Hall, T., 2003. De-motivators of software process improvement: an analysis of practioners' views. The Journal of Systems and Software 66, 23–33.

Batista, J., Dias de Figueiredo, A., 2000. SPI in a very small team: a case with CMM. Software Process Improvement and Practice 5, 243–250.

Brodman, J.G., Johnson, D.L., 1994. What small businesses and small organizations say about the CMM. In: Proceedings of 16th International Conference on Software Engineering (ICSE 1994). IEEE Computer Society Press, New York, pp. 331–340.

Chrissis, M.B., Konrad, M., Shrum, S., 2003. CMMI: Guidelines for Process Integration and Product Improvement. Addison-Wesley, Boston, MA.

CMMI Product Team, 2001. Appraisal Requirements for CMMI, version 1.1, Technical Report CMU/SEI-2001-TR-034, Carnegie Mellon University Software Engineering Institute.

Cohen, J., 1960. A coefficient of agreement for nominal scales. Educational and Psychological Measurement 20, 37–46.

Conradi, R., Fuggetta, A., 2002. Improving software process improvement. IEEE Software (July/August), 92–99.

Fowler Jr., F.J., 2002. Survey Research Methods, third ed. Sage Publications.

Goldenson, D.R., Herbsleb, J.D., 1995. After the Appraisal: A Systematic Survey of Process Improvement, its Benefits, and Factors that Influence Success. Technical Report CMU/SEI-95-TR-009, Carnegie Mellon University Software Engineering Institute.

Herbsleb, J., Carleton, A., Rozum, J., Siegel, J., Zubrow, D., 1994. Benefits of CMM-Based Software Process Improvement: Initial Results. Technical Report CMU/SEI-94-TR-013, Carnegie Mellon University Software Engineering Institute.

Hyde, K., Wilson, D., 2004. Intangible benefits of CMM-based software process improvement. Software Process Improvement and Practice 9, 217–228.

Jung, H.W., Hunter, R., Goldenson, D.R., El-Emam, K., 2001. Findings from phase 2 of the SPICE trials. Software Process Improvement and Practice 6, 205–242.

Mehta, C.R., Patel, N.R., 1986. Algorithm 643 FEXACT: A FORTRAN subroutine for Fisher's exact test on unordered $r \times c$ contingency tables. ACM Transactions on Mathematical Software 12 (2), 154–161.

Paulk, M.C., 1998. Using the Software CMM in Small Organizations. In: Joint Proceedings of the Pacific Northwest Software Quality Conference and the Eighth International Conference on Software Quality, pp. 250–361.

R Development Core Team, 2005. R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria.

Shaw, M., 2003. Writing good software engineering research papers: minitutorial. In: Proceedings of the 25th International Conference of Software Engineering (ICSE 2003). IEEE Computer Society, Washington, DC, pp. 726–736.

Staples, M., Niazi, M., 2006. Systematic Review of Organizational Motivations for Adopting CMM-based SPI. Technical Report PA005957, National ICT Australia.

Trewin, D., 2002. Small Business in Australia: 2001. Australian Bureau of Statistics report 1321.0.

Wilkie, F.G., McFall, D., McCaffery, F., 2005. An evaluation of CMMI process areas for small-to-medium-sized software development organizations. Software Process Improvement and Practice 10, 189–201.

Mark Staples is a Senior Researcher in the Empirical Software Engineering program at National ICT Australia. Before joining NICTA, he worked for medium-sized software-developing companies in roles covering verification, configuration management, product line development, and process improvement. His current research interests cover these areas of experience. He holds a Ph.D. from the Computer Laboratory, University of Cambridge.

Mahmood Niazi is a Lecturer at School of Computing and Mathematics at Keele University. He has spent more than a decade with leading technology firms and universities as a process analyst, senior systems analyst, project manager, and lecturer. He has participated in and managed several software development projects. These have included the development of management information systems, software process improvement initiatives design and implementation, and several business application projects. He holds a Ph.D. from the Faculty of IT, University of Technology Sydney.

Ross Jeffery is Program Leader of the Empirical Software Engineering program at National ICT Australia, and is a Professor at the Department of Computer Science and Engineering at the University of New South Wales. He holds a Ph.D. from the University of New South Wales.

Alan Abrahams is a consultant with Object Consulting Pty Ltd. He provides consulting support for capability improvement activities including developing and implementing SPI programs. He has had experience in general business management and consulting, with a strong background in measurement and improvement within several different industries and companies.

Paul Byatt is a consultant with experience in quality management, software process improvement, project management, and consulting. Paul has held management positions in a variety of companies and industries and has implemented and supported application systems in Australia, Europe and South America.

Russell Murphy is a consultant at Downing Consulting, Melbourne Australia with experience as a Software Engineer, IT Manager, Project Manager and Process Consultant. He has a detailed knowledge of standards and models such as EIA 632 Processes for Engineering a System; IEEE Documentation Standards for Software/Systems Engineering; ISO 9001:2000; CoBIT; CMM/CMMI; TSP and PSP; and ITIL. He has participated in IT infrastructure projects including IT/IS strategic planning, risk management, business continuity planning, disaster recovery planning, acquisition and policies, procedure and processes. He holds a Diploma of Information Technology from the RMIT.