ORIGINAL ARTICLE

Quantifying the impact of requirements definition and management process maturity on project outcome in large business application development

Keith Ellis · Daniel M. Berry

Abstract Using data from two surveys of people knowledgeable about requirements for, and the success of the development of, large commercial applications (CAs) in hundreds of large organizations from around the world, this paper reports a high positive correlation between an organization's requirements definition and management (RDM) maturity and that organization's successful performance on CA development projects. Among the organizations that responded with a filled survey, an organization that is assessed at a high RDM maturity is significantly more successful in its CA development projects than is an organization that is assessed at a low RDM maturity, when success in CA development projects is measured as (1) delivering CAs on-time, on-budget, and on-function, (2) meeting the business objectives of these projects, and (3) the perceived success of these projects. This paper presents a comprehensive framework for RDM, describes a quality RDM process, and describes RDM maturity and how to measure it. It describes the two surveys, the first of which ended up being a pilot for the second, which was designed taking into account what was learned from the first survey. The paper concludes with advice to practitioners on the application of the RDM maturity framework in any

organization that wishes to improve its RDM and its performance in the development of large CAs.

1 Introduction

From the very beginning, the requirements engineering (RE) field has been seeking industrial-strength empirical evidence that the improved RE processes that it has been advocating for use in developing computer-based systems (CBSs) have a positive effect on these developments, making the process of development faster, less time challenged, and more predictable, while at the same time producing less faulty, more reliable, more correct, more complete, and more maintainable CBSs. Among these CBSs are commercial applications (CAs)[1] that drive modern businesses and are the kinds of CBSs considered in the research reported in this paper. The field has been seeking this evidence partially to address the complaints by practitioners that the methods really work [15]. This empirical evidence has been slow in coming at least partially because good empirical evidence about RE processes, particularly in the industrial context, is hard to come by.

One reason that good empirical evidence is hard to obtain is the very difficulty of conducting valid experiments that

Written while Keith Ellis was at IAG Consulting, Oakville, ON L6M 3E6, Canada.

K. Ellis
Department of Computing, Lancaster University,
LA1 4WA Lancaster, UK
e-mail: Ellis.in.toronto@gmail.com

D. M. Berry (✉)
David R. Cheriton School of Computer Science,
University of Waterloo, Waterloo, ON N2L 3G1, Canada
e-mail: dberry@uwaterloo.ca

---

[1] Commercial applications are known also as business applications. This paper uses "commercial application" and its acronym "CA" to reserve the "BA" acronym for its common usage to mean "business analyst". No acronym is given to "business analysis".

show what must be shown. Controlled experiments necessarily work with toy-sized artifacts, and it is not clear that results can be generalized to developments of real-life CBSs whose artifacts are orders of magnitude bigger than the artifacts that are used in affordable, controllable, and completable experiments. Conclusions derived from a case study involving the development of a real-life CBS are based on a single data point and are not statistically significant. These difficulties are real regardless of whether the data are gathered while the developments are happening or the data are mined from data saved about past developments.

Secondly, it is difficult to arrange for experiments and case studies in the industrial context. A company that has saved some evidence might be reluctant to share data about its CBS developments because of the possibility of revealing company secrets or information about challenged or failed developments, while other companies are not able to generate data of sufficient quality needed to support an experiment which spans many projects or companies. It is hard to find a company that will agree to subject one of its CBS developments to an experiment in which the CBS is developed multiple times, with different methods. A commercial entity has little interest in or benefit from spending, for example, the $5 Million a large CBS costs and then spending this money again to see whether a different method would produce a better CBS in less time, in the interest of advancing scientific knowledge about development methods. Moreover, data must be gathered over entire development lifecycles, lasting years, from initiation through maintenance. On the other hand, the results of a controlled experiment involving students using CBS artifacts small enough to be acted upon in a reasonable time span, for example, a few hours, are not generalizable to industrial situations.

Nevertheless, there have been a few studies to empirically measure the effect on the development of CBSs the use of RE, its methods, and its tools, mostly since 1994. Each of these studies has been successful in answering one or more questions about the effectiveness of RE, its methods, or its tools. Slowly, the combined results are beginning to provide convincing evidence that RE, the tested methods, and tested tools do help significantly improve the development of CBSs. Section 2 reviews the empirical evidence about the effectiveness of RE, its methods, and its tools. Nevertheless, even this existing empirical evidence has a gap: there are no quantifications specifically of RE maturity, and there are no correlations of measures of RE maturity to project outcomes.

The purposes of this paper are to add to this empirical evidence about RE and to address this gap. The paper describes two surveys conducted among senior organization and project management personnel at hundreds of organizations around the world that yielded data about industrial CA development projects and their requirements definition and management (RDM) processes. It presents a model for RDM maturity that was tested and improved over the course of the two studies, as a result of addressing methodological weaknesses in the first study. The data from each survey show a high correlation between the maturity of a CA development organization's RDM and improved outcomes of its CA developments arising from the specified requirements. Furthermore, Sect. 5.9 shows that even though the two surveys used different questions and RDM maturity measures, their data *can* be compared. Under this comparison, the two surveys' data are found to correlate well; thus, the credibility of each survey is strengthened.

Section 2 of this paper describes past empirical work attempting to validate the positive impact of RE. Section 3 asks the research question this paper attempts to answer, whether the quality of an organization's RDM predicts the success of its strategic CA development projects. It describes the design of the first study, using a survey, that attempted to answer this research question and reports the findings of the first study. While these findings uncovered clear relationships between RDM quality and project outcome, the findings also pointed to both the need for a better measure of RDM quality and the need to define this measure of RDM quality in a framework amenable to implementing in companies. Then, Sect. 4 introduces a framework for describing an organization's RDM process and offers RDM maturity as an improved measure of RDM quality. Section 5 reformulates the research question in terms of RDM quality and describes the design of a new, improved survey for the second study. It presents the findings from the second study, namely that the higher an organization's RDM maturity, the better its CA development performance, and it addresses the limitations and threats to validity of the findings of the second study. Section 6 offers anecdotal evidence that an organization that raises its RDM maturity will improve its CA development performance. Section 7 suggests future research based on questions that were raised but not answered by the second study results. Finally, Sect. 8 concludes the paper.

2 Past related work

This section describes past work that attempts in some way to empirically determine the impact of RE on CBS development.

2.1 Empirical studies in requirements engineering

The recent empirical work to validate RE, methods, and tools is summarized very nicely in a paper titled "An

Analysis of Empirical Requirements Engineering Survey Data", by Paech et al. [34]. Beginning with interview-based Field studies by Curtis et al. [8], studying the design process for large CBSs in the mid 1980s, and by and Lubars et al. [33], studying the state of RE practice in the early 1990s, there have been some interview-, survey- (i.e., questionnaire), workshop-, and focus-group-based studies of

- the state of RE practice,
- industrial uptake or avoidance of RE,
- problems faced in RE for real-life CBS development,
- RE in small-to-medium enterprises,
- RE for CBSs under time-to-market constraints,
- the impact of RE on CBS construction project performance,
- success criteria for evaluating RE,
- best RE practices, and
- user participation in RE.

In addition, Paech et al. report on several studies of specific RE phenomena, artifacts, methods, tools, and activities, for example,

- requirements volatility,
- requirements specifications,
- traces from requirements,
- requirements elicitation,
- requirements analysis,
- scenarios and use cases,
- Quality–function deployment (QFD), and
- reviews and inspections of requirements artifacts.

The overall conclusion reached by Paech et al. is that "It has been established that RE makes a difference for project success." [34, p. 439].

One of the studies cited by Paech et al. on the impact of RE on CBS construction project performance was that done by El Emam and Madhavji [12]. El Emam and Madhavji conducted systematic interviews of personnel at a consulting company whose business was doing RE for clients using the company's own method. They interviewed requirements engineers about the RE processes that they performed in their client engagements in order to learn what is important to do well in an RE process. Among the things that must be done well are

- managing the level of detail in functional models of the CBS,
- learning what is possible from the current system being replaced by the CBS,
- user participation in the RE process, and
- managing uncertainty.

Also, they discovered that their subjects were concerned a lot about the benefits of CASE tools.

The other of the studies cited by Paech et al. on the impact of RE on CBS construction project performance, which is closest in methodology to the present paper, is that by Hofmann and Lehner [23]. With the help of surveys and interviews of 76 stakeholders, Hofmann and Lehner studied 15 RE teams, of which nine developed customized CBSs and six developed COTS-based CBSs, in nine software companies and development organizations in telecommunications and banking. The 76 stakeholders included project managers, requirements analysts, customers, and quality assurers. The studied projects lasted an average of 16.5 months, expending 120 person months of effort. The main positive conclusions of this study were that

1. RE teams that built prototypes or models of their CBSs did better than others,
2. the most successful teams expended an effort on requirements specifications that was twice that expended by the least successful teams, and
3. the most successful teams did RE for a greater portion of the lifecycles of their CBSs than did the least successful teams.

Surprise conclusions, some negative, of this study were that

1. in all but one project, commercially available RE tools interfered with the performance of RE activities,
2. stakeholders found that focusing on functions and data resulted in a "lack of total system requirements attention" and in "incomplete performance, capacity, and external interface requirements",
3. stakeholders felt that lack of traceability of the implications of requirements hurt their projects, and
4. ranking the priority of requirements caused the most difficulty for RE teams.

From these empirical conclusions, Hofmann and Lehner are able to make specific recommendations of best practices.

In yet other work, Lauesen and Vinter [31] conducted a study of the defects in one product of one software development organization to determine the defects' causes. As expected, most of the defects were requirements related. They tried to identify which of 44 requirements engineering techniques described in the literature would be best for avoiding the company's defects. In the end, it came down to about ten well-known techniques that the organization had not used in the past. They helped the organization apply the techniques to a new product. The organization thoroughly studied the user tasks for the new product, built early prototypes for the user interface, and tested these prototypes for usability. The surprising result was that the organization released the product on-time with much less

stress than usual at the organization. Moreover, because the new product's user interface addressed the typical user's needs better than that of any of any competing product, and the new product could be offered at half the price of any competing product, the new product sold twice as well as the competing products.

Hall et al. [19] used 45 focus groups involving about 200 people to conduct an in-depth study of 12 software development organizations' requirements problems. They sought to determine among other things,

- what kinds of requirements process problems the organizations were experiencing and
- whether increased process maturity reduces requirements process problems.

They concluded in the end that *most* requirements process problems are organizational rather than technical. For example, employees' lack of skills and poor employee retention negatively impact the production of initial requirements sets. They concluded also that there *is* a positive relationship between an organization's process maturity and patterns of requirements process problems. High maturity organizations do have fewer problems in their requirements processes.

The year 2005 saw a full-day workshop on the theme that upfront RE pays off in an improved development, CBS, or both. Each presentation described a case study of the development of a real-life or substantial research prototype CBS in which thorough RE was done before development began. The slides of these presentations are available at the workshop Web site [36]. A 1.5-h summary of this workshop was presented at a panel titled "To do or not to do: If the RE payoff is so good, why aren't more companies doing it?" at the 2005 International Requirements Engineering Conference [4].

Damian and three different sets of co-authors report, in three separate papers, the results of an extensive 30-month, three-stage, explanatory case study of the RE process at the Australian Center for Unisys Software (ACUS) [9–11]. The study used mainly questionnaires, interviews, and document inspection to gather data. During the case study the Center was undergoing a concerted RE process improvement following a Capability Maturity Model (CMM) [35] assessment, and as reported in the papers, the RE process improvement did indeed improve the Center's CBS development. The first paper reports the early benefits that were *perceived* by the development teams during RE and the early stages of development as a result of the RE process changes. The second paper reports the benefits *observed* by the development teams during the downstream stages of development as a result of the RE process changes. The third paper explores how the entire RE process interacts with other development processes and how

this interaction affects the development. It concludes that an effective RE process from the beginning of a development improves the entire lifecycle by improving the effectiveness of other processes in the lifecycle, including project negotiation and planning, management of feature creep, testing, and defect finding and rework, and ultimately, it improves the quality of the product developed during the lifecycle.

Sadraei et al. [37] did a field study of RE practice in 28 software projects in 16 Australian software companies using semi-structured interviews and a detailed questionnaire. They were able to examine in detail the characteristics of the RE processes in the projects. From this examination, they were able to model each project's RE process and then to compare the models to each other to draw conclusions. The main finding is that more upfront RE leads to less rework. The subsidiary findings include that

- skimping on RE activities results in spending more time in later activities,
- organizations with more experience with mission critical systems tend to follow a more structured approach to their RE processes and documentation,
- using a standardized or formalized RE process in a project is not essential for the project's success; instead there must be timely distribution of accurate requirements information, and having a customer or users continuously on site and involved works as well, and
- there is no universally applicable RE process; each project must find its its own that fits its context.

2.2 Empirical studies in software development process improvement and maturity

There have been several models developed of general software development process maturity in an attempt to measure and encourage software development process improvement, a.k.a. software process improvement (SPI). Examples include the Software Engineering Institute (SEI)'s CMM [35] and CMMI (CMM Integration) [39]; the International Standards Organization (ISO)'s ISO/IEC 15504-1:2004 [27], which is known also as the Software Process Improvement and Capability dEtermination (SPICE) suite; the European Strategic Program on Research in Information Technology (ESPRIT)'s Bootstrap [18]; and the business-oriented process improvement method, Six Sigma [20]. Each of these models suggests that at higher levels of process maturity, developers perform better with less process volatility. Each suggests also that systematic gains are mode by moving from a low maturity level to a higher level.

These models have been accepted by some as valid without empirical validation and have been rejected by others for the very lack of empirical validation. There have

been attempts to empirically validate some of these maturity models. For example, there have been several studies of the effect on the CBS developments carried out by an organization of the organization's achieving a high CMM level or improving its CMM level.

A notable one of these is that by Herbsleb and Goldenson from the SEI that developed the CMM. They conducted a systematic survey of 155 organizations that had undergone CMM-based SPI [21]. They received completed questionnaires and data from 138 of the 167 approached individuals representing 61 assessments from over 400 CMM assessments. Herbsleb and Goldenson found that CMM-based SPI pays off in improved

- ability to meet schedules,
- ability to meet budget,
- product quality,
- developer productivity,
- customer satisfaction, and
- staff morale.

A later SEI study by Gibson, the same Goldenson, and Kost [17] of 35 organizations, some large, showed specific quantified improvements in several performance measures as a result of CMMI-based SPI.

Galin and Avrahami [16] summarized and aggregated 19 studies to reach a quantitative conclusion that was stronger than the sum of the individual conclusions, that investing in CMM level improvement definitely leads to improved software development and maintenance, assuming that the averages used as input to their study represented failure adequately. They showed that the impact from a level improvement from a low level is higher than from a level improvement from a high level.

Wang et al. [43] report a quantitative comparison of SPICE, CMM, ISO 9000 [26] and Bootstrap for the purpose of allowing any organization to choose the most appropriate model for its SPI attempt.

None of these models is focused specifically on RDM maturity or process improvement. Sommerville and Sawyer [40, p. 79] observe that "[software productivity improvement] has improved the process of developing products from requirements, but it hasn't helped the development of requirements from customers … many SPI programs do not adequately address the requirements problems that underlie poor product quality. Timeliness, cost control, and quality may improve, but the level of achievable improvement is capped by flaws in the requirements process."

Existing models of RDM maturity are far more limited and are, as yet, unquantified. The structure of the RMM used in this paper is similar in many respects to the REGPG work of Sommerville and Sawyer [40] in a number of respects, but most significantly in that the RMM uses a

Table 1 Cost driver and capability ratings versus multipliers

Cost driver	Ratings				
	Very low	Low	Nominal	High	Very high
Analysts' capability	1.46	1.19	1.00	0.86	0.71

hybrid of CMM and SPICE concepts to create a model that is both implementable and measurable. However, Sawyer et al. did not quantify the impact of requirements improvement on project performance beyond concluding that [38, p. 84]. "It is reasonable to aim for a 20 percent reduction in the number of reworked requirements in each improvement cycle [through implementing REGPG]."

2.3 Empirical studies of RE effectiveness based on frameworks for estimation

Empirical evidence of the importance of RDM within the overall SDLC can be found in the software cost estimation literature. Barry Boehm [5] accounts for the analyst's capability as one of the fifteen effort adjustment factors applied to scale up or down an initial cost estimate based on the size of the software to be developed. Boehm's model suggests that an organization with very low capability analysts would estimate its cost at 205.6% times that of an organization with very high capability analysts, if all other variables, including code size, were held equal. Table 1 shows all the multipliers for the analysts' capability effort adjustment factor. Further, Boehm's model has a higher multiplier for very low capability analysts than for any other very lowly rated personnel or project effort adjustment factor, suggesting that a very poor analyst would increase costs more than any cost increasing factor would.

2.4 Bold statements in the popular press

By 2008, the date of the first study, many industry researchers, practitioners, and government oversight organizations had issued to the popular press bold statements about the impact of flawed requirements:

- "Flawed Requirements Trigger 70% of Project Failures" [29].
- "finding and fixing requirements errors consumes between 70% - 85% of total project rework costs" [32].
- "between 40 percent and 60 percent of software defects and failures can be attributed to bad requirements" [1].
- "we estimate that about 48 percent of the federal government's major IT projects have been rebaselined…. The most commonly cited reason for rebaselining was changes in project requirements, objectives, or scope—55 percent." [42].

- "Deficient requirements are the single biggest cause of software project failure. From studying several hundred organizations, Capers Jones discovered that RE is deficient in more than 75 percent of all enterprises [28]. In other words, getting requirements right might be the single most important and difficult part of a software project." [23].

These bold statements made for wonderful, scary headlines and certainly brought the need for better RDM into the spotlight in a helpful way, but the conclusions and data are rather limited in quantifying benefit. The tacit underlying assumption of most of these studies seems to be that RDM is akin to taking out an insurance policy, something that one does to avoid or at least recover from an unexpected catastrophe, rather than as something proactively done to ensure project success. Moreover, these types of statements promote the thinking that if a project did not fail, it must have had good RDM. We take the opposite view and maintain that the benefit of good RDM is not simply disaster avoidance or recovery but actual project improvement. Rather, RDM has a significant value that increases with increasing RDM maturity. The first study of this paper was designed to measure of the strategic benefit of RDM by assuming that most requirements lie on a spectrum of goodness somewhere between excellent and terrible. By comparing a measure of RDM quality to project outcomes, a more forward looking and benefit-centric viewpoint of RDM is tested.

2.5 Summary of related empirical work

While there have been some empirical studies to evaluate RDM and its methods and tools, the number of organizations covered by each study has been quite small, reporting on only 1 through 28 projects in 16 organizations. The only study cited above with more than one hundred organizations is the one by Herbsleb and Goldenson, with 155 organizations about CMMI—general software engineering maturity—levels, not specifically RDM maturity levels. The studies described in this paper concern the RDM maturity levels of 110 and 437 large CA development projects in those many organizations.

3 The first study: quantifying the impact of RDM process maturity on project outcome in CAs

The first study was conducted largely by Ellis in 2007 and 2008.

3.1 Research question and studied projects

The research question driving the first study [13] is:

Does the quality of an organization's RDM process predict the success of the organization's strategic CA development projects?

In this question, the term "strategic project" describes a CA development project with three properties:

1. The project's budget must be in excess of $250,000 for development, needed software, and external services.
2. The project must involve software development or application implementation.
3. The project must deliver business capability or software functionality that is significantly different from that which existed prior to the project.

Thus, a strategic project does not include (1) any simple or routine project of only moderate complexity, (2) any project to develop infrastructure or to roll out new technology, and (3) any maintenance, bug-fixing, or platform migration project that does not change the organization's business. Restricting the study to strategic projects is intentional, even though doing so means that the study's conclusions may not apply to the large number (See below for an estimate of just how big that number is.) of small projects, in which different laws may hold.

Strategic projects can be contrasted with simplistic projects, those that are equipment centric or technology centric. Restricting the study to strategic projects excludes about 90% of all projects, but includes projects that consume more than 50% of project spending.[2] Strategic projects receive the majority of capital spending on CAs. The study author believed that a simplistic project does not have the same need for high levels of business interaction in defining requirements. Thus, a simplistic CA project should less impacted than a strategic CA project by low quality RDM. The belief is that the relatively high failure rate among strategic projects might be attributable mainly to low RDM quality. Fortunately, because of the high visibility of strategic projects, the failure rate of strategic projects is measurable. Therefore, there is the possibility to empirically relate failure rate and RDM quality for strategic projects.

The research question was to be answered by conducting a survey of business or IT executives, managers, and professionals from around the world using their experience-based opinions to establish both (1) the quality of the RDM processes in their organizations and (2) the outcomes of strategic CA development projects in their organizations. Of course, since the respondents would be reporting their perceptions, what is really measured are the *perceived*

---

[2] These percentages are guestimated by the first author, Ellis, based on his more than 19 years of full-time experience as a software development methods consultant, and the second author, Berry, agrees from his more limited but longer experience.

quality and the *perceived* outcomes. This paper, as do others reporting survey results, continues to describe what is measured as simply "quality" and "outcomes".

3.2 Goal definition

The goal definition [45] of the survey is as follows:

Object of study The object studied is the entire RDM process as a step of a CA development process.

Purpose The purpose is to measure the impact of business requirements quality on the outcome of strategic projects.

Quality focus The quality foci are (1) the completeness of the RDM process versus (2) the completeness of the CA and the timeliness and on-budget-ness and the perception of success in the delivery of the CA.

Perspective The perspectives are from the viewpoints of business or IT executives, managers, and professionals.

Context A survey is conducted among business or IT executives, managers, and professionals from around the world, getting each's perceptions about the *most recent strategic* CA development project at his or her organization with questions that ask about (1) the quality of the organization's RDM process in the project, (2) the outcomes of the project, and (3) the quality of the project's delivered CA.

Thus, the first study examined only the strategic projects that were selected for reporting by the survey respondents, creating the possible threat of non-representativeness of the selected projects. However, it was hoped that requesting the same specific strategic project, that is, the most recent, from each respondent would introduce enough systematic pseudo-randomness to ensure representativeness of the selected projects.

The survey's questions had to be designed to elicit the data necessary to answer the research question. The set of potential respondents had to be representative of the population of business or IT executives, managers, and professionals concerned with developing CAs and the response had to be high enough to allow generalizing from the results to the population. See Sect. 3.4 for details about the population, the responders, and the questions.

3.3 Threats

Therefore, the four main threats to the validity of the results of the first study are

1. that the set of potential respondents is not representative of the population of business or IT executives,

managers, and professionals concerned with developing CAs,
2. that the response level does not allow generalizing to the population,
3. that the respondents' perceptions do not match their organizations' realities, and
4. that the strategic CA development projects selected by the respondents as the subjects of their answers are not representative of all CA development projects. Indeed, there is a small chance that each respondent selected his or her best, recent project, and thus, failed projects are underrepresented. However, let us compare the results of the first study to those of another study on overall project success rates for a similar period, that is, the Standish Group's 2009 CHAOS report [41]. If the first study were revised to use the Standish definition of success, that a project is successful when it is delivered on-time, on-budget, and on-function, then the first study results, that 30% of the projects were successful are virtually identical to the Standish results, that 32% of the projects were successful. Unfortunately, the Standish group defined project failure as cancelation of the project, and the first study did not include this concept. Nevertheless, the first study's overrun statistics shows a somewhat higher average functionality delivered per project than did the CHAOS report, but it shows also comparatively higher overruns on-time and cost. These comparisons lead us to believe that the typical respondent reporting a failing project in the first study had spent greater time and money to achieve a higher proportion of delivered functionality, due to the strategic nature of the projects studied. This tradeoff may be what distinguishes the projects of first study from the project population as a whole.

The first study survey was developed and evaluated and its findings were reported [13] by Ellis, then of IAG Consulting, with the help of Michael O'Neil of the Info-Tech Research Group. The Infotech Research Group took care of field testing the questions. The final survey was fielded using the InfoTech Research Group's Internet survey infrastructure.

3.4 Conduct of first survey

The first study's survey was sent to over 10,000 people, mostly IAG clients, targeting people likely to be in leadership roles in strategic projects. The first study's survey triggered over 400 responses from around the World, each describing its respondent's chosen project. Excluding responses about non-strategic projects winnowed the data to those describing 110 strategic projects. Over 75% of the projects reviewed in the included responses were critically

important or very important to the respondents' organizations. The typical CA developed in these projects was fundamental to the performance to the developing organization, large and cross-functional with lots of interdependencies with other projects, costing at least $250,000 for its development, costing, on average, $3 million for its development, and representative of the kind of large-scale developments initiated at companies today. It was therefore expected that the results would be good, generalizable, and therefore useful.

Among the 110 respondents there were

8, about 6.4%, business managers and executives,
47, about 37.6%, business or IT project managers,

47, about 37.6%, business analysts,
10, about 8.0%, principal technology executives, and
13, about 10.4%, technology professionals.

3.5 Survey contents and evaluation of answers

Tables 2 and 3 capture the essence of the first survey, showing the actual text of the factors the participants were asked if they were present in their projects. The first table is titled "Factors Related to RDM Quality" and the second table is titled "Factors Related to Organization". Each line of either table shows one of 21 questions from the survey, which represented a factor whose correlation to project

Table 2 RDM quality factors assessed in first survey

Factors related to RDM quality	Correlation to project success
Describing project goals and objectives in concise, clear, and unambiguous terms	0.2320
Facilitating cross-functional group sessions where requirements were discovered	0.3457
Conducting efficient meetings, and making effective use of stakeholder time	0.4108
Accurately documenting the business process behind the application	0.1935
Describing the information flow, and key data needed by users at any given time in the business process	0.2427
Assuring that the scope of the project neither significantly changed nor had major in-scope functions moved to follow-on phases of the project	0.3834
Uncovering project interdependencies or issues that need investigation	0.4579
Clearly describing project risks and assumptions	0.3384
Presenting the results of analysis in clear, well-organized, and readable documentation	0.3061
Assessing change requests: determining the impact of these on scope and cost, and the impact of systems changes on business processes	0.2368
Achieving consensus on requirements amonge stakeholders	0.4449
Getting requirements documented in a short, concentrated period of time	0.4264

Table 3 Organizational factors assessed in first survey

Factors related to organization	Correlation to project success
Our organization has a formalized approach to doing business requirements which is consistently followed by business analysts on projects	0.2434
Our organization has defined standards for business requirements documentation quality, and assess the work of analysts against these standards on projects	0.3755
Our organization treats business analysis as a profession and has trained staff dedicated to this function	0.2509
Our organization can predict how much stakeholder time will be needed, and which stakeholders will be involved in the requirements phase of a project	0.2349
Business requirements are traceable, and well integrated into testing at our organization	0.1717
Stakeholders feel that the process of extracting and documenting requirements is efficient at our organization	0.3249
Our organization is excellent at transitioning requirements from business departments into the information technology department	0.2106
I believe the automated business analysis tools that we current have in place are excellent at helping us elicit requirements	0.2022
I believe the automated business analysis tools that we currently have in place are excellent at helping us manage requirements and requirements change	0.2265

success was being measured. Each of the twelve questions in the first, "RDM-quality" table represents one task in a good RDM process, and each of the nine questions in the second, "Organization" table represents one way that an *organization* can be good about its RDM process. Each respondent answered each question with one value in a five-point ordinal scale. For each first table question, the ordinal scale rated the skill, from very low through very high, of the respondent's organization in performing the question's task. For each second table question, the ordinal scale rated the respondent's agreement, from strongly disagree through strongly agree, that his or her organization is good in the way stated in the question. After answering these 21 questions, each respondent was asked to rate the success of his or her selected, most recent strategic project on a five-point ordinal scale from failure through unqualified success. It was expected that a project's timeliness and cost was a factor in determining its success.

For the purpose of correlating RDM process quality with project outcomes, RDM quality for a respondent's organization was taken as the numerical sum of the respondent's answers to the twelve questions of the "RDM-quality" table. This sum has a maximum of 60 and was scaled by 1.67 to get values in the range of 1 through 100.

3.6 Findings

As shown by the numerical data in the second columns in Tables 2 and 3, the study found a stronger correlation between high skill in RDM tasks and project success than between organizational RDM process goodness and project success. However, no one factor of the 21 appeared to be strongly correlated to project success—whether the factor focused on a RDM task or on organizational RDM process goodness. The study did conclude that an organization could not, for example, "define standards for business requirements documentation quality" and in the absence of other changes, expect predictable and measurable improvement in project outcomes.

3.7 Retrospective on the study

During the analysis of the data from the first survey, Ellis realized that instead of seeking individual factors about an organization's RDM process or documentation standards that affected project success, a more process-centric viewpoint was needed, especially if the intents were to demonstrate (1) the impact of sustained improvement in RDM and (2) to identify strategies for making this improvement. To prototype a more process-centric study, he re-evaluated the data of the first study with the new process-centric viewpoint. He took the sum of the nine factors related to organization as a proxy for maturity in RDM and divided the existing data into maturity quartiles. From this different viewpoint, a clear and *significant*, with $\chi^2 = 0.000298473$ with 6 degrees of freedom, pattern emerged. The graph in Fig. 1 shows the result of this re-analysis. The graph suggests a strong positive correlation between an organization's RDM maturity and the ultimate success of its chosen strategic CA development project. While these may be interesting results, from the reanalyzed first survey findings, Ellis could neither conclude that making incremental improvement yields positive results nor identify the changes that an organization had to make so that its overall performance on projects will improve predictably. In other words, the first survey did not yield an *implementable* approach to improving CA project outcome. As a result, Ellis decided that for the next run of the survey, he would use RDM maturity as the measure of the quality of an organization's RDM process. He decided also that he would correct in the new survey at least the threat caused by the first survey's instruction to answer the questions about one selected project from the organization; the

Fig. 1 Correlation between project outcomes and project maturity
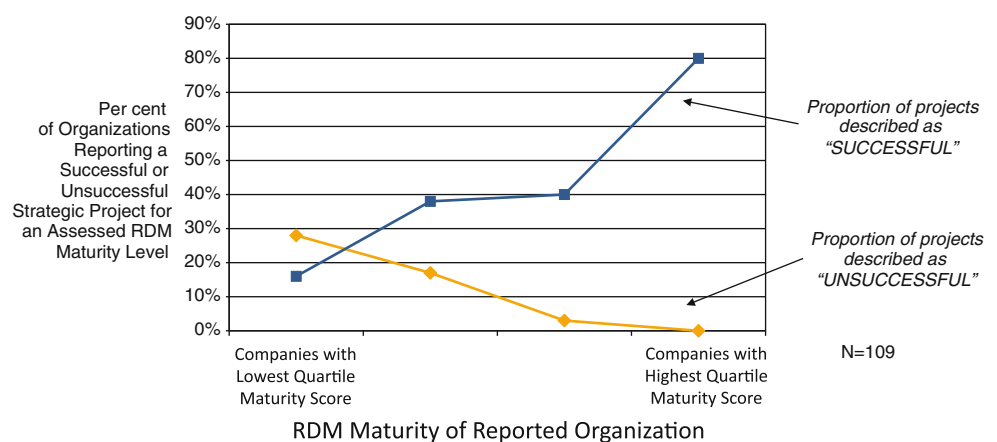
corresponding instruction in the new survey is to answer the questions about the organization's complete portfolio of projects.

3.8 Onward to the second study

While the first study was arguably loose in its definition of an organization's RDM process quality, it did allow Ellis to discover a measurable notion of RDM maturity and to define a model of RDM processes in which RDM maturity makes sense.

The second study [14], described in detail in the remaining sections of this paper, extends the first study with new data into an implementable model by

1. giving precise definitions of "RDM processes" and "RDM maturity", including the incremental levels of maturity
2. changing screening criteria, allowing examination of a larger range of projects sizes and surveyed companies, and dramatically increasing the number of respondents to the survey, and
3. examining and controlling variables such as

   - "adopted system development lifecycle (SDLC)", such as "agile", "waterfall", "iterative", and "visualization centric", and
   - the "skill" of a business analyst, using an emerging industry-accepted standard definition [25] which defines skills distinct from just those required for RDM,

more precisely.

These elements were integrated into the overall study design and are examined in more detail in the rest of this paper.

4 RDM framework and maturity model development leading to second study

This section describes a framework for describing and understanding RDM and model of the maturity of an organization's RDM. The model motivates the questions that were asked in the questionnaire of the second study.

4.1 A framework for RDM research

The RDM maturity model that serves as a basis for the questions of the conducted survey assumes a particular framework for describing the RDM process and its documents. The framework, from IAG consulting, recognizes that there is a tendency in industry to treat requirements as a document. That is, there is a high degree of focus on the requirements document itself—templates for it; its features; how to review it; and even, in the case of agile [2] development, questioning the need for it. Requirements are not a document; requirements are a process. This process is what this paper calls "the RDM process". The documents produced by that process, for example, a requirements specification (RS), communicate a specific state of requirements understanding to other stakeholders in the development process that created the document. These documents should serve the needs of other interdependent business and technological processes in the same organization, as illustrated in Fig. 2 from IAG Consulting [14]. One can legitimately argue that a document that does not serve any need should not even be produced [2].

A central tenet of this RDM maturity model is that a specific RDM-produced document has a value which may be negative as well as positive, just as a poorly developed blueprint for a building may have negative value for a builder attempting to build the specified building. The
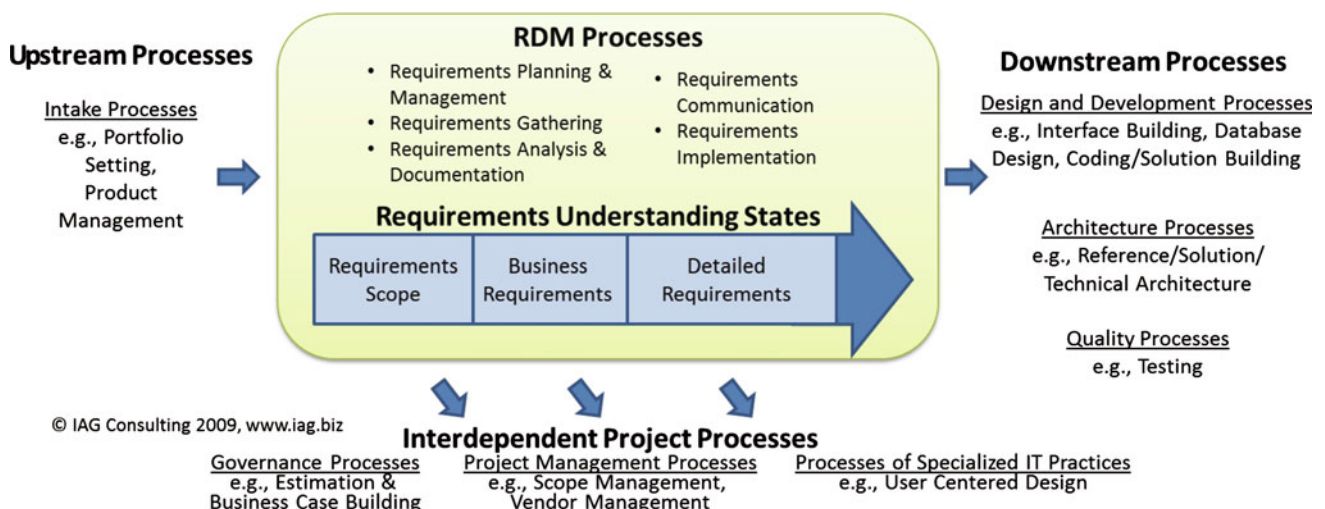

Fig. 2 IAG Consulting's RDM framework

maturity of a RDM process cannot be evaluated from only the RDM-produced documents since a well-constructed document that does not serve the needs of its downstream consuming stakeholders ends up having a negative value to the RDM process.

Many, particularly those in the RE field [4, 36, 38], those in IAG Consulting and other similar companies, and those who have learned some hard knocks lessons, believe that:

- A high quality RDM process has a high probability of yielding high quality artifacts that properly serve the needs of the interdependent processes needed for CA development, and therefore, has a high probability of resulting in a successful development initiative and a high quality CA product. Conversely, a poor RDM process has a high probability of yielding poor results.

- A high quality RDM process evolves with the projects to which it is applied. There are however, specific underlying processes unique to RDM that may be individually assessed and that are important to the functioning of the process as a whole [3, 6, 44].

- Furthermore, in any project of significance, stakeholder understanding of requirements evolves through various states of understanding. With a strong RDM process, these states of understanding build on each other, iteratively adding layers of detail through stages of requirements development activity, while maintaining the understandings of prior stages. Hence, traceability is maintained. With a poor RDM process, more detailed understanding does not naturally build on the understanding of prior work, reducing traceability and creating rework.

- To achieve a high quality RDM process, an organization must have supporting capabilities such as techniques for requirements planning, elicitation, definition, and documentation. It must have also skilled individuals providing services understood and valued by the organization's leadership, as well as a RDM process that engages stakeholders to set expectations and to manage these expectations. A RDM process cannot simply exist without adequate supporting capabilities. Furthermore, these supporting capabilities and the associated processes need to be implemented within the organization and need to be functioning with a degree of consistency. That is, the RDM process must be mature; it cannot simply be on paper.

To summarize, the key elements of the assumed framework are that

1. requirements are a process, the RDM process;

2. a high quality RDM process must not only be defined, but it must also be supported by capabilities that allow the process to be implemented; and

3. a RDM process must be implemented and mature to have value.

By assessing the strength of an organization's capabilities supporting RDM, one can determine the organization's overall RDM maturity.

The assumed framework is critical to this research. This framework accepts that an organization's RDM process can produce requirements documents that have little, no, or negative value to the overall development process in which the RDM process is embedded. It is believed that the probability of an organization's producing low-valued requirements documents is high in the absence of RDM maturity. On the other hand, when quality processes and supporting capabilities are implemented and institutionalized into the fabric of an organization, its RDM process and its artifacts have positive value to its development processes. Quantifying the impact of an organization's requirements documents on its development processes relies on having established measures of RDM maturity and measuring how the organization performs at the model's maturity levels. One purpose of the studies reported in this paper is to validate this belief.

4.2 Components of the IAG requirements maturity model

The concept of capability maturity has been around since Deming first started the quality movement in the 1950s. Since then, hundreds of maturity models, for example, CMM [35], have been developed, with surprisingly few focused specifically on CA requirements; a notable exception is the maturity model of Sawyer et al. [38].

The maturity model developed as part of the research described in this paper is now known as IAG's RDM Maturity Model (RMM). The RMM makes a precise and consistent definition of the observable states of an organization's maturity in RDM that can be both quantified through research, and be implemented into organizations. Therefore, from this point forward, whenever this paper refers to "an organization's RDM maturity", it is referring to the organization's maturity with respect to IAG's RMM, which is hereinafter called simply "the RMM".

The framework presented below for describing RDM and the RMM and the concepts in Figs. 3, 4 and 5 were constructed by Ross Little of IAG Consulting and tested by having successfully evaluated the RDM maturity of many complex organizations, having had effective discussions based on the findings of RDM maturity assessments on

Fig. 3 IAG maturity model stages
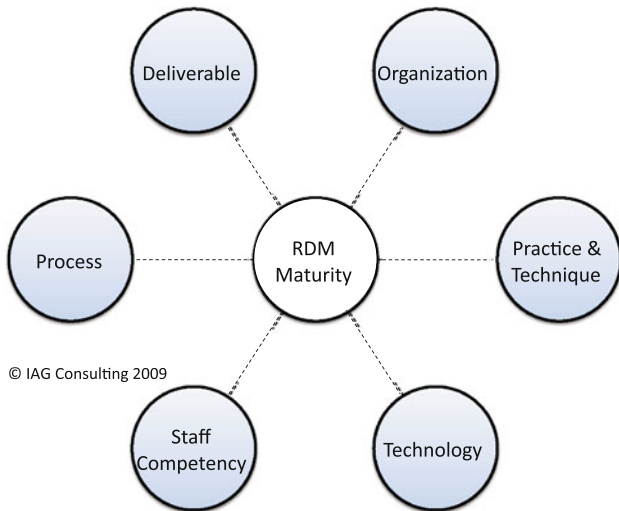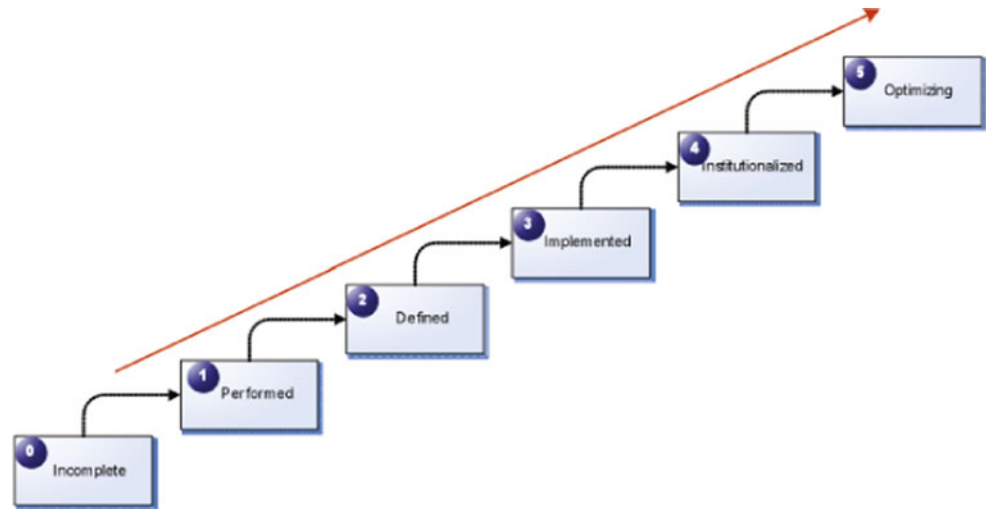
Fig. 4 IAG RDM maturity model capability areas

Maturity Level Capability Area	Level 1	Level 2	Level 3	Level 4	Level 5
Process					
Staff Competency					
Technology					
Practices & Techniques					
Organization					
Deliverables & Results					

Fig. 5 Matrix of RMM maturity levels and capability areas

how to improve RDM maturity in a specific organization, and having implemented this RDM maturity model at many customer organizations of IAG Consulting.

4.2.1 Structure of the requirements maturity model

The RMM has two dimensions,

1. maturity levels and
2. capability areas:[3]

Maturity levels The RMM is a staged and gated maturity model similar to those used by several industrial standards bodies, for example S, SEI's CMM. It is staged because there are five defined levels that an organization can achieve, namely Levels 1 through 5: Performed, Defined, Implemented, Institutionalized, and Optimizing, respectively. There is, for completeness, an additional level, Level 0: Incomplete, for any organization that has no RDM process at all. The RMM is gated because to be regarded as having achieved any specific level, an organization must have achieved measurable level-specific thresholds. Level 0 is an exception to this gatedness; there is no gate through which to get into it. An organization starts at Level 0 and progresses up the staircase shown in Fig. 3 as goals are achieved and thresholds are surpassed. Each level of maturity shifts the organization's emphasis to different requirement practice characteristics. Each level builds a foundation for succeeding levels.

Capability areas Figure 4 shows the six capability areas that are assessed and combined to determine the maturity

---

[3] "CA" continues to mean "commercial application" and not "capability area".

level of an organization's RDM. These six capability areas are the fundamental building blocks for RDM and include the following:

Process definition, usage, and management of RDM procedures.

Staff competency level of knowledge, skills, and ability of the RDM workforce.

Technology provision, usage, and integration of software tools in the context of the RDM practice.

Practices and techniques definitions of (1) how analysts will perform work and (2) the efficiency and effectiveness of these activities.

Organization (1) the RDM organizational model and the RDM services delivered to stakeholders, (2) the provision of resources and resource management in the delivery of these RDM services, and (3) the framework of RDM process and tool governance.

Deliverables and results templates, work products, and artifacts from the RDM process.

4.2.2 Measuring maturity

Overall maturity in this study is a composite of performance across the six capability areas. Each capability is weighted slightly differently based on IAG's experience in what drives effective long-term performance and consistency in RDM. Hence, each organization's maturity is determined first for each of the six capability areas, and then an aggregated RDM maturity score is calculated from the six capability area maturity scores.

Maturity is very step-like in nature. Thus, there are Levels 1 through 5, and, the calculations described later notwithstanding, there is no Level 2.3, for example. An organization progresses up the ladder as tangible goals are achieved and thresholds are surpassed. The gating to levels facilitates practical maturity level assessment of an organization. In practice, the capability and maturity assessor of an organization must be both (1) objective to be defensible and (2) able to describe to the organization what it must achieve to be gated to the next level of maturity. In the absence of specific criteria for each of the 30 cells in Fig. 5, it is very difficult to assess companies objectively or to systematically determine what must be done to improve maturity.

Here is how the matrix of Fig. 5 is used by an assessor during an assessment of an organization: For each cell at row r and column c, there is a set of characteristics, forming a criterion, that tells the assessor things to look for at the organization in order that the organization be considered to be at c's maturity level for r's capability area.

For example, for the capability area "Practices and Techniques", there are eight characteristics. An organization must meet all eight of these characteristics is considered to be at Level 1 for the capability area. In an assessment, an assessor checks off each cell whose criterion is met by the organization. At the end, the organization receives for each capability area the maturity level of the rightmost cell checked off. Note that for any capability area, the checklist for the criterion at Level i includes all items in all the criteria at Levels j for $j < i$. The organization's aggregate maturity score is a weighted average of its six capability area maturity levels. Section 5.3 explains how the second survey was designed so that a respondent's answers to the survey questions would simulate an assessment of his or her organization. Thus, what is true of an assessment's aggregate maturity score should be true also of an aggregate maturity score calculated from a survey response.

Therefore, *even though RDM maturity improvement is very steplike in nature and is best described using ordinal numbers, any way of measuring RMM must make use of rational numbers*, because it is unlikely that the weighted average score of 6 ordinal values will result in an integer. For this study, the concern is to classify each organization into the level containing organizations most like it, and a decision must be made on how best to segment rational values into ordinal categories.

In this study,

1. we preserve the gated concept of the RMM, that is, an organization's weighted average score must be at least 2 for it to be considered at Level 2, and
2. we chose to drop from consideration any organization whose score was too far from any integer value, whenever doing so was possible given the sample size, in order to focus on organizations that are most like Level 1, most like Level 2, most like Level 3, etc. This exclusion reduced the data, but sharpened the results.

The mapping to RMM levels from aggregate maturity scores is as follows:

Level 1 a survey maturity score of less than 2,[4]
Level 2 a survey maturity score of at least 2.0 but not more than 2.499,
Level 3 a survey maturity score of at least 3.0 but not more than 3.499,
Level 4 a survey maturity score of at least 4.0 but not more than 4.499, and
Level 5 a survey maturity score of at least 5.0.[5]

---

[4] While it would have been appropriate to use a score of 1 to 1.499, the sampling of companies with a score of under 1.5 was too small to accommodate this segmentation.

[5] While this level is defined, in fact no respondent had a high enough maturity score to achieve this level.

This approach allows approximating ordinal levels from rational number data. In an assessment of an organization that is trying to improve its RDM maturity, an assessor may present non-integer maturity scores to the client to help it understand the assessed maturity level more precisely and visualize how fast it is moving to the level at which it wants to be.

It is very useful to have a model of requirements maturity that is being measured in terms of what an organization and its members do at each level of maturity. Such a model directs the invention of questions to test maturity level and informs the association of possible answers to these questions to maturity levels. Without such a model, the search for question can be haphazard and the interpretation of answers to these questions can be guess work.

5 The second study

The second study was conducted by Ellis in 2008 and 2009.

5.1 Reformulated research question and studied projects

The research question given in Sect. 3 has to be reformulated in terms of RDM maturity:

> Does the RDM maturity of an organization's RDM process predict the success of the organization's strategic CA development projects?

This research question drove the second study [14].

The definition of "strategic project" is loosened from that of the first study. The description given to respondents of a strategic project was, "The survey is focused on people who have experience on larger scale strategic projects in medium to large sized companies, and can comment knowledgeably on strategy and performance". This instruction enabled any CA development initiative of any significance to be included. As with the first study, the new research question was to be answered through a survey, again of business or IT executives, managers, and professionals from around the world. The second survey asked each respondent to think of the portfolio of all strategic projects undertaken in his or her organization and to answer the questions about the portfolio as a whole.

5.2 Goal definition

The goal definition of the second survey is as follows:

Object of study The object studied is the entire RDM process as a step of a CA development process.

Purpose The purpose is to measure the impact of an organization's RDM maturity on the outcomes of the organization's projects to develop strategy CAs.

Quality focus The quality foci are (1) the maturity of the RDM process versus (2) the completeness of the CA and the timeliness and on-budget-ness of the delivery of the CA.

Perspective The perspectives are from the viewpoints of business or IT executives, managers, and professionals.

Context A survey is conducted among business or IT executives, managers, and professionals from around the world, getting the perceptions of each about the portfolio of all CA development projects at his or her organization with questions that ask about (1) the maturity of the organization's RDM process in the project, (2) the outcomes of the project, and (3) the quality of the project's delivered CA.

Thus, the second study examined all strategic projects in the organizations of the survey respondents, thus avoiding the first study's possible threat of non-representativeness of the respondents' selected projects.

5.3 Second survey design

The second survey was designed by first establishing the business issues that needed to be examined by the research and then mapping these issues to their knowledge domains. The mapping, which was produced in several long consensus-seeking focus sessions at IAG Consulting led by Ellis, is shown in Fig. 6. These domains were then massaged into survey questions that capture this knowledge.

A number of factors were critical in the survey design:

Method independence The RMM does not assume any SDLC, for example, agile, waterfall, iterative, and visualization centric, or even any RDM or development method, to be superior to another. For any capability area, as long as an organization reports being consistent in its use of some set of RDM and development methods, it is considered mature in the capability area. The organization needs neither to consistently use one method nor to use all methods to be considered mature in the practices and techniques capability area.

Consistent participant interpretation The survey question were field tested early to discover and rephrase unclear questions, for consistency of question interpretation, and to determine the minimum time required to answer each question to serve as a reality check on a response to the survey.

Balanced coverage of capability areas The survey questions were checked for adequate and balanced coverage

Organization
- Size
- # Business Analysts
- # Resp for Bus RDM
- Avg Size of Projects
- Industry
- Country
- State

BA Skill/Competency
- Overall Skill Level
- Training
- Skill in each IIBA® defined skill area:
  - BA Planning
  - RM
  - Enterprise Analysis
  - Elicitation
  - Req Analysis
  - Solution Assess & Validate

RMM Level
- Overall Level
- Process Level
- Practice Level
- Technology Level
- Org Level
- Staff Skill Level
- Deliverables Level

BA Infrastructure
- Formal Bus. Analysis/ RM Organization
- RMO/RCoE
- Rpt' g relationship of Business Analysts
- Quality standards
- ETC…

Requirements Results
- % proj on time
- % proj on budget
- % proj on function
- % proj on objectives
- Overrun if overun on any of above
- Avg time to def req.
- % project budget on requirements
- Level of quality of Bus requirements
- 'Success' rate of projects

Technology Usage
- Modeling SW
- RM SW
- Integration of above
- Implement/plan to impl

Methods
- Agile
- Waterfall
- Iterative
- Prototyping/ Visualization
- Usage now and trend

Deliverables
- Clarity, accuracy, completeness
- Documentation standards defined

Practices
- Practice guidelines
- UML/BPMN
- Use Case modeling
- User Stories
- Process modeling
- Data modeling
- Req Mgmt planning
- Requirements discovery practices
- Review practices
- Prioritization
- Change management
- Scorecarding
- Etc.

Processes
- Defined process for elicitation
- Roles are defined
- Change processes defined
- Service levels defined
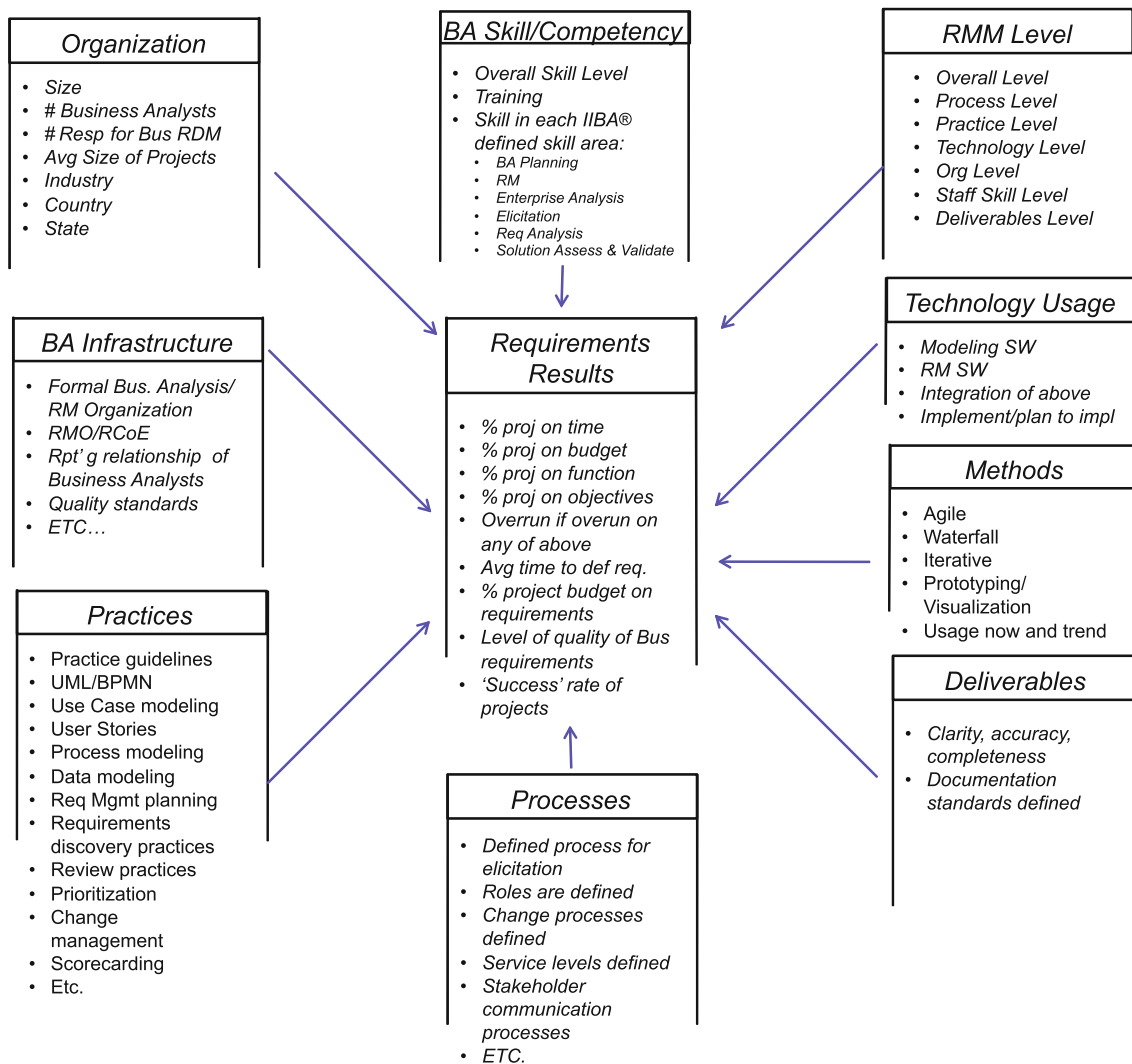- Stakeholder communication processes
- ETC.

Fig. 6 Knowledge domains assessed

of each of the six capability areas of the RMM. A few questions touched on multiple capability areas, and each of these was counted as covering all touched capability areas. *Consistent interpretation of results* For each survey question, each anticipated possible answer was assigned a maturity level of 1 through 5 for its covered capability areas. For example, answering "C" to Question 13 might indicate a maturity of 2 in the staff competency capability area. Moreover, the pattern of maturity levels for each answer was randomly different for each question. During the evaluation of any completed survey, the discovery of an unanticipated answer resulted in cataloguing a quick decision about the maturity level assigned to that answer. It was decided that for any completed survey, the aggregate maturity level for any capability area is the average of the maturity levels achieved in all of the survey questions

covering the area. Finally, a weight was determined for each capability area for its contribution to a total RMM.

5.4 Assessing performance and interpreting development success

The second survey asked each respondent to think of the portfolio of strategic projects undertaken in his or her organization and to report what percentage of these projects were on-time, were on on-budget, were on-function, and had achieved their business objectives.

Secondarily, the survey asked the respondent for the average percentage by which projects that were not on-time, not on-budget, or not on-function missed the time, budget, or function. With these questions, it was possible to separate the probability of achieving project success from the average magnitude of overrun of unsuccessful projects.

In addition, the questions asking about project outcomes appeared in the survey *before* the questions asking about maturity. This ordering, together with the assumption that the typical respondent answered the questions in the order presented helped reduce the chances that a respondent's assessment of his or her organization's maturity would affect his or her assessment of the success of his or her organization's projects. The survey itself can be viewed at http://ba_benchmark.questionpro.com.

For example, to measure an organization's overall success with its projects, the survey asked the respondent:

What proportion of your projects are considered:

_____ outright failures
_____ poor
_____ neither successful nor unsuccessful
_____ successful
_____ unqualified successes

5.5 Conduct of second survey

The second study survey was developed, conducted, and evaluated by Ellis and was field tested with the help of IAG Consulting employees. The final survey was fielded using the QuestonPRO Internet survey infrastructure. The survey results were reported by Ellis and sent to an external panel composed of industry executives, development executives, project managers, managers of business analysts and an industry analyst at InfoTech Research Group for their comments and review. Many of the experts provided personal experiences, other stories, and other data that corroborate the survey results. Some of the inputs provided by these external panel members were reported on in the published report about the second study [14].

In the second quarter of 2009, an online survey portal was used to ask the second study survey questions, to manage skip patterns, and to record responses. A request to participate in the second survey was sent to approximately 40,000 users of IAG's research and to the respondents of the first survey. The request was also posted on the Internet in various discussion forums and was publicized in IAG's newsletters and other communication media.

Just under 550 responses with completed surveys were received, including approximately 50 from respondents of the first survey. The approximately 1.4% response rate is about three times the 0.5% response rate that Ellis expected[6] from a request for participation like this one that was broadcast by e-mail.

These 550 responses yielded 437 qualifying responses. A response qualified only if it

---

[6] Ellis's expectation is based on his 18 years of experience conducting marketing surveys.

Table 4 Demographics of business analysis benchmark survey

	Percentage of respondents
Position	
Executive: head of IT, CIO, head of development, line of business executive	12.16
Head of PMO or project manager	27.06
Head of business analysis competency center or BA	52.52
Other	8.26
Total	100.00
Number of employees in company	
1–99	6.2
100–499	14.3
500–2,499	20.3
2,500–4,999	11.5
5,000–9,999	8.5
10,000 +	39.2
Total	100.0
Industry	
Energy, resources & utilities	3.9
Financial services	17.7
Insurance	9.9
Government & social services	8.7
Health care & pharmaceutical	8.0
Manufacturing	6.0
Media & industry analysts	1.1
Military & defense	1.4
Professional & IT services	14.4
Retail, transportation & distribution	5.0
Software	9.2
Telecommunications	6.6
Education	2.1
Other	6.0
Total	100.0
Region	
United States	53.3
Canada	26.5
Western & Eastern Europe	6.0
India & Pakistan	5.5
Asia & Pacific	5.0
Africa & Middle East	2.6
Central & South America	1.1
Total	100.0

- described an organization that spends over $1 million, net of hardware, annually in development of CAs and other software,
- filled by an individual with experience with RDM and project management for CAs and other software that have added new functionality to the organization's business.
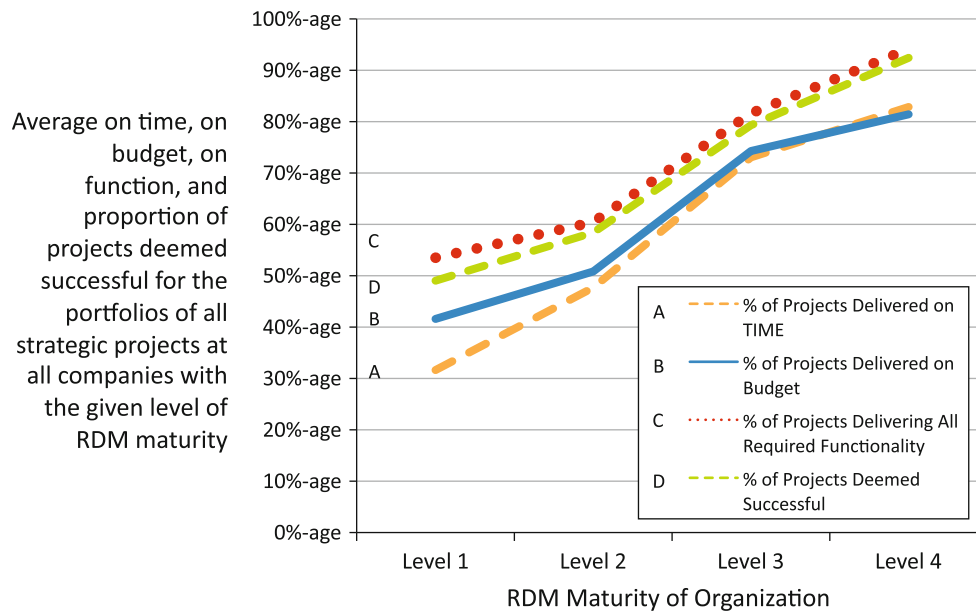
Fig. 7 Percentage of projects on-time, on-budget, on-function, and successful, by RDM maturity level

- described an organization that has run at least four projects each costing at least $250,000 in the last 12 months.

These criteria ensured that the results come from data from only professionals experienced in RDM and project management on CAs.

As with the first study, restricting the study to the responses from organizations dealing with large projects is intentional, even though, again, doing so means that the study's conclusions may not apply to the many CA development projects that are small. Note, however, that agile development projects are not necessarily excluded from this study. Recall that for this study, each projects was classified according to its SDLC. Thus, agile projects are visible, and they show up among the included and excluded data.

The following steps were taken to enhance data integrity:

An incomplete response was dropped from the sample. A response that did not meet the qualifying criteria were dropped from the sample for analysis even if the response was complete. Any response that consumed less than the minimum time needed to complete the survey, as determined by the sum of the minimum times to read every question, was dropped from the sample.

In the rest of this paper, "project" refers to any of the kinds of projects that caused its organization to be qualified for inclusion into the data for the second survey.

The resulting sample included mainly medium and large sized commercial organizations, about 80% from North America. The sampling demographics are summarized in Table 4.

5.6 Summary of the findings

Even though the limitations described in Sect. 5.8 prevent wholesale generalization from the findings, the raw data exhibit correlations relevant to the reformulated research question with high certainties, and they are reported as such. The correlations relate an organization's RDM maturity with indications of the success of its strategic CA development projects.

The graph in Fig. 7 shows that low RDM maturity organizations perform more poorly than high RDM maturity organizations on every measure of development effectiveness. Not only are high RDM maturity organizations noticeably better at servicing the needs of their businesses, they perform nearly twice as well on every measure of development effectiveness assessed by the second survey:

- percentage of projects completed on-time,
- percentage of projects completed on-budget,
- percentage of projects delivering all required functionality, and
- percentage of projects deemed successful.

Accordingly, we say "organization A outperforms organization B" to mean "A is more on-time, on-budget, and on-function in its portfolio of strategic projects than is B and A has a higher success ratio within its portfolio of strategic projects than B has".

A total of 74.1% of the organizations reported on in the second survey, were found to have low levels of RDM

Table 5 Distribution of respondents by maturity level

Assigned level of RDM maturity	No. of surveys	Percentage of used sample
Level 1	57	19.9
Level 2	135	47.2
Level 3	87	30.4
Level 4	7	2.4
Used sample	286	
Total surveys	437	

maturity, that is, maturity Levels 1 or 2. To see the magnitude of impact that RDM maturity has on the performance of an IT organization, compare the typical organizations at Levels 1 and 4. The results of the second survey show that on average, a Level 4 organization

- was on-time on 161% more of its projects,
- was over time on 87% fewer of its projects,
- was on-budget on 95% more of its projects,
- was over budget on 74% fewer of its projects,
- was on-function on 76% more of its projects, and
- was missing function on 78% fewer of its projects

than was a Level 1 organization. That is, in each of the four measures of performance—being on-time, being on-budget, delivering the required functionality, and being considered a success—the confidence level that a Level 4 RDM maturity organization outperforms a Level 1 RDM maturity organization is at least 99.99%, which is calculated using a 1-tail T-test with the assumption of unequal variance in the samples, yielding an average P value of 9.63×10^{-08}.

5.7 Elaborating the findings

Table 5 shows the distribution of survey respondents by the RDM maturity levels of the organizations they described. About half of the respondents described organizations that were at RDM Maturity Level 2, Defined. The typical Level 2 organization has defined its RDM processes and has the required skills, techniques, and deliverables, but fails to implement them consistently, in IAG's experience, usually because the standards are not enforced by the organization or these standards are not sufficiently defined across all capability areas to be operationalized.

Table 6 gives detailed performance statics by RDM maturity level. The table shows that by every measure of developmental effectiveness assessed in the second survey, the average organization at Level N is significantly better than the average organization at Level $N - 1$, for $N = 2$, 3, or 4. Tables 7 and 8 allow assessing the

Table 6 Detailed performance statistics by maturity level

Assigned level of RDM maturity	Avg. calculated maturity score for level (out of 5)	Avg. percentage of projects on-time	Avg. percentage of the time overrun when time overrun occurs	Avg. percentage of projects delivered on-budget	Avg. percentage of budget overrun when budget overrun occurs	Avg. percentage of projects delivering 100% of expected functionality	Avg. percentage of functionality missed when functionality is missing	Avg. percentage of projects that achieve business objectives	Avg. percentage of business objectives missed when business objectives are missed	Avg. percentage of projects considered poor or a failure	Avg. percentage of projects considered neither successful nor unsuccessful	Avg. percentage of projects considered successful or unqualified successes
Level 1	1.8	31.6	51.1	41.6	49.3	53.5	23.0	54.2	25.5	23.8	27.2	49.0
Level 2	2.2	47.7	50.1	50.8	51.9	60.4	23.3	50.9	22.9	16.8	24.8	58.4
Level 3	3.2	73.0	27.6	74.3	26.5	81.6	14.0	70.0	14.7	8.6	12.1	79.3
Level 4	4.0	82.9	11.1	81.4	12.4	94.1	5.1	82.7	13.4	3.0	4.6	92.4
Overall avg.	2.6	56.5	41.4	59.7	42.5	68.0	19.0	59.8	20.3	14.2	19.1	66.7

Table 7 Average and standard deviation by performance measure

Maturity level	No. of surveys	On-time		On-budget		On-function		Success rate	
		Avg.	SD	Avg.	SD	Avg.	SD	Avg.	SD
Level 1	57	31.63	27.49	40.86	31.42	52.56	29.85	49.04	25.94
Level 2	135	47.68	30.25	50.45	30.66	59.10	31.62	58.41	27.49
Level 3	87	72.21	23.12	73.43	27.38	80.68	21.35	79.30	20.15
Level 4	7	92.50	3.54	85.00	7.07	99.50	0.71	96.50	2.12

significance of these differences between organizations at different RDM maturity levels.

Table 7 presents the underlying statistics describing for each RDM maturity level, the sample size, the average, and the standard deviation for each performance measure. Figure 8 shows the plots of the average and of +1 and −1 standard deviations around the average against the maturity level for each of the four performance measures. It is very clear that the general trend in the data is that the higher the maturity level past Level 2, the lower the variance, as measured by standard deviation, for each performance measure. The variance is dramatically smaller for Level 4 than for Level 3.

Table 8 shows the confidence level of the claim that, on average, a Level N RDM maturity organization does better on each of the components of the performance measure, defined earlier, than a Level $N − 1$ RDM maturity organization, for each $N = 2, 3$, or 4. Recall that the components of the performance measure defined earlier are the percentages of strategic projects finishing on-time, on-budget, on-function, and with an evaluation of success. The confidence level that on average, a Level 3 RDM maturity organization does better on each of the components of the performance measure than a Level 2 RDM maturity organization approaches absolute certainty; the potential for error is 4.53 times in 100 million.
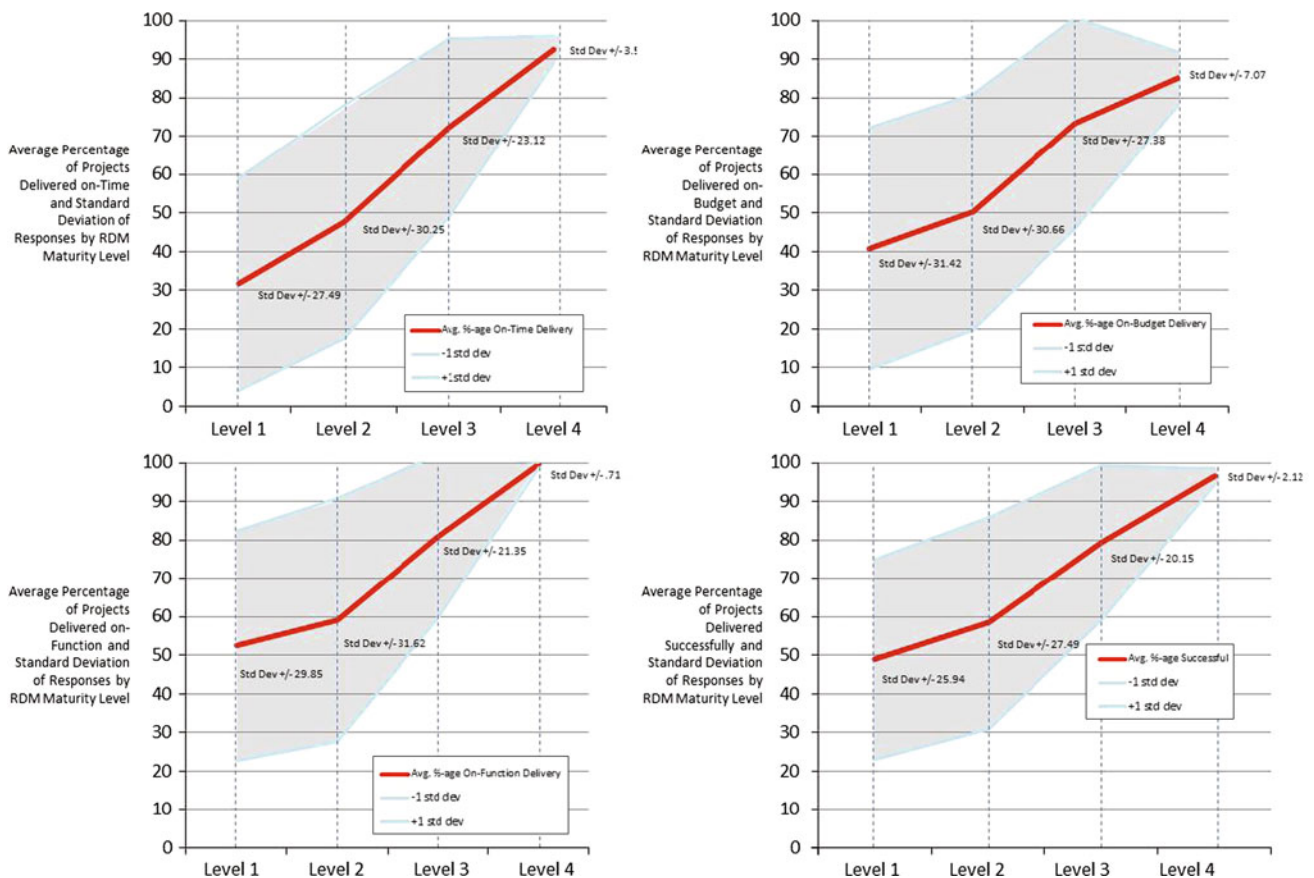

Fig. 8 Plots of average and standard deviation by performance measure

Table 8 Confidence levels that performance will be higher at level N than at level $N - 1$

Maturity level	Number of surveys	On-time	On-budget	On-function	Success
Level 1	57	N/A	N/A	N/A	N/A
Level 2	135	99.98%	97.3%	91.2%	98.7%
Level 3	87	99.99+%	99.99+%	99.99+%	99.99+%
Level 4	7	99.6%	94.8%	99.99+%	99.6%

5.8 Limitations and threats to the validity of the conclusions

The nine main threats to the validity of the results are listed here.

1. The set of potential respondents is not representative of the desired population of business or IT executives, managers, and professionals concerned with developing CAs. As described in Sect. 5.5, the request to participate in the second survey was sent to users of IAG's research and respondents of the first survey. All of these people were in the desired population. The request was also posted on the Internet in discussion forums frequented by people in the desired population. Finally, the request was publicized in IAG's newsletters and other communication media read by people in the desired population. That this request reached more than 40,000 people of the desired population coupled with the fact that desired population is a small fraction of the people working in IT and developing CAs, we believe that the population receiving the request is representative of the desired population.

2. The response level does not allow generalizing to the population. This threat is present in every survey in which responding to the request to participate is voluntary. No survey escapes the problem that a responder may *want* to tell his or her good story and a non-responder may be hiding his or her bad story. Nevertheless, the request yielded 550 responses, of which 437 were usable. As mentioned, the approximately 1.4% response rate is about three times the 0.5% response rate that Ellis expected from a request for participation that was broadcast by e-mail. Therefore, the response to this survey seems to have done a better job of being representative of the population than most.

   For this survey, there was the added limitation that only responses from organizations that spends over $1 million annually for development of CAs were deemed qualified. Only about 10% of software development organizations are qualifying. However, this limitation was intentional, because, as mentioned, this 10% spends about 50% of the total software development money and has the biggest to lose if a CA development effort fails. That the 437 qualifying responses are from only 10% of the software development organizations makes the 437 qualifying responses be a larger percentage of the desired population than it would otherwise be.

3. The respondents' answers do not match their realities. Also, this risk is unavoidable in a survey. The only way to ameliorate this risk is to find other ways to determine the information provided in the survey answers. As mentioned, the answers to the survey were judged as credible by a panel of experts who were sent the survey and answers for evaluation. See also Sect. 5.9 for additional mitigation of this threat.

4. In the first survey, the strategic CA development projects selected by the respondents as the subjects of their answers are not representative of all CA development projects. The first survey looked at only large, strategic CAs that required at least $3 million for development. This limitation was intentional, and it was always understood that the findings may not generalize to all CA development projects. In fact, the cost premium paid on projects with poor RDM were far higher for these strategic CAs than in the more generalized second survey for all projects in the respondents' portfolios.

5. In the second survey, the portfolio of CA development projects selected by the respondents as the subjects of their answers are not representative of all CA development projects. The threat for this item is, by design, less than for Item 3, because each respondent was asked to report average values for all projects in his or her portfolio. To the extent that each respondent's portfolio was representative of his or her organization's CA development projects, the average values are representative of the values for all of his or her organization's CA development projects. See also Sect. 5.9 for additional mitigation of this threat.

6. The results stem not from RDM maturity differences, but from differences in some other unmeasured independent variable. For example, well-managed companies tend to do a lot of things well, and poorly managed companies tend to do a lot of things poorly. As in any research, statistical correlation cannot imply causality. In fact, no amount of correlation data can

imply causality. In either of the two studies, it is possible that some factor not measured caused the project successes and, as a by-product, caused the organization to appear or to become more mature in RDM. The way to establish causality is to make predictions on the basis of assumed causality and to demonstrate empirically that the predictions happen. All of the maturity model work [e.g., 35] predicts that an organization that behaves in a manner measured as mature produces good software. Each experimental validation that good software development behavior is correlated with good software being developed strengthens the case that maturity causes development of good software. The studies reported in this paper contribute to this growing body of empirical evidence.

7. There are variables, other than the six capability areas chosen to define RDM maturity, that may define RDM maturity better and that may be more closely correlated to predictable, successful project outcome. The RMM is only a model and therefore is only an abstraction and simplification of reality. Further research is needed to determine if a simpler model could be found. However, the findings of the first survey make it abundantly clear that no attempt to improve RDM maturity by focusing on few specific activities is likely to be successful. Each tested factor has only limited correlation to project success. Practitioners, working with the RMM in the real world, will be the ultimate judge of what defines RDM maturity in each project in each organization.

8. The survey as developed did not allow accurate assignment of RDM maturity levels to organizations. A survey instrument cannot provide the hours of examination that would happen during a full-fledged maturity audit or during in-depth case study research. Certainly, more questions could have been added to the survey to more accurately assigned RDM levels, particularly to detect the extremes of RDM maturity at Levels 0, 4, and 5. The trade-off, of course, is that a longer survey question set suffers a higher percentage of incompletely answered surveys.

9. This survey, as any empirical research, can be susceptible to researcher bias and other reliability issues. Therefore, the survey's validity and findings must be directly assessed for these issues. Testing this validity is the subject of Sect. 5.9.

5.9 Testing survey validity

A central issue in establishing validity is establishing the repeatability of the findings. In this case, there are two surveys exploring the same issues, but using different survey questions and different definitions of subject CAs.

Yet, the two studies revealed essentially the same results. Moreover, multiple studies improve evidentiary value. In this case, the second survey was designed to be neither a retest of the first survey nor an alternate-form verification of reliability:

- The first survey investigates a single strategic CA development project, namely the most recent one for the respondent, and the second investigates the respondent's entire portfolio of strategic CA development projects.
- The sample sizes are different in the two studies.
- The RDM experience of the respondents in the two surveys is likely different.
- The measures of RDM maturity are different in the two studies.

Despite these differences in the two surveys, if the data of the second survey were recast and analyzed as described below to simulate the approach used to segment the data of the first survey, the correlation between the two survey results that relate project success rates with RDM maturity quartile is 0.63, only slightly below the 0.7 or better expected from a true test–retest investigation of result reliability [30, p. 21].

This recasting of the second survey data involves four steps:

1. The second study's net RDM maturity scores in the range of 1.4–4.2 were scaled to the range of 1–100. This scaling simulates the RDM maturity range used in the first study.
2. The scaled RDM maturity results were segmented by quartile as was done in the prototyping of the future second study at the end of the first study.
3. Since both surveys asked about success rates using the same segmentation, but with the first asking about the success rate of a single project and the second survey asking about the average success rate in a portfolio of projects, we made the simplifying assumption that each organization had the same number of projects in each quartile of success rates.
4. The calculated RDM maturity values were then related to the calculated success rates.

Tables 9 and 10 relate the success rates of projects to RDM maturity for the two studies, using the above analysis.

In other areas, we believe the second survey has reasonable reliability:

1. The questions were developed and field tested to verify (1) that they were understandable, (2) their response rates, (3) their minimum response times, and (4) the reliability and validity of their instrumentation.

Table 9 Second study data, recast to quartiles

Quartile	Percentage of projects that are considered				
	outright failures	poor	neither successful nor unsuccessful	successful	unqualified successes
1st	5.7	15.8	26.4	43.4	8.8
2nd	3.6	10.3	22.6	52.0	11.5
3rd	5.2	8.6	16.7	52.7	16.8
4th	2.5	5.3	10.7	66.0	15.6

Table 10 First study reported results

Quartile	Percentage of projects that are considered				
	outright failures	poor	neither successful nor unsuccessful	successful	unqualified successes
1st	4.0	24.0	56.0	12.0	4.0
2nd	8.3	8.3	45.8	37.5	0.0
3rd	0.0	3.3	56.7	30.0	10.0
4th	0.0	0.0	20.0	55.0	25.0

2. The questions were stated as objectively as possible, asking about specific activities rather than for opinions on subjective issues.

3. For questions about techniques used, whose answers can be subjective, the survey results did not rate one technique over another. Rather, the results rated the consistency and frequency of a particular technique's use. Thus, the study avoids being swayed by bias toward waterfall, iterative, agile, or visualization-centric methods.

4. Content validity was maintained by rigid adherence to the RDM maturity framework and having the same number of questions assessing all six capability areas.

5. The study results were sent to a panel of industry observers to for them to judge whether their personal experiences were consistent with the results.

6. The results are consistent with IAG's experiences in RDM maturity and in other research.

5.10 Conclusion of results

The sum total of the threats argue against generalization to all CA development projects and organizations, particularly because we intentionally limited consideration to large projects and to organizations doing them. Nevertheless, as one reviewer observed, "The challenge of creating a valid experimental setup for RE is very difficult, and in a way what has been done here is as good as can be hoped for". This same reviewer cautioned against over generalizing and insisted that "the claims must reflect the reality of what was (and what can be) done ....".

The research question addressed by the second study was:

Does the RDM maturity of an organization's RDM process predict the success of the organization's strategic CA development projects?

Modulo the threats, the results and conclusions from the data obtained from the second survey strongly support a "yes" answer to this research question.

In reporting the results and conclusions from the data, we stuck to what exactly the data show, that high RDM maturity is correlated very well with high productivity. We refrained from even predicting that if an organization were to *improve* its RDM maturity, it will *improve* its productivity. It's hard to see how that prediction is not valid, but the data themselves do not allow us to infer that prediction.

These results, as all results have limited generalizability. However, these results, combined with the ever growing body of evidence reported in Sect. 2 and studies yet to be reported, contribute to an increasing ability to predict that greater RDM maturity leads to greater CA development productivity.

6 Anecdotal evidence that an organization that raises its RDM maturity will improve its performance

The results of this paper compare the performances of independent organizations at different maturity levels and show that the average organization at one RDM maturity level outperforms the average organization at a lower RDM maturity level. The real concern for the practitioner is whether *one organization* that has moved from one RDM maturity level to a higher one will outperform *itself* at its former RDM maturity level, that is, will improve its performance. We call this real concern the *desired hypothesis*.

The results, strictly speaking, do not allow concluding the desired hypothesis.

Empirically validating the desired hypothesis would require tracking several organizations over several years. No organization that participates would be anonymous from us. While we do not have the data to validate the desired hypothesis, there is little reason to doubt that an organization will indeed improve its performance as it moves upward in RDM maturity level. For an organization not to do so would mean that the organization, which moved upward in RDM maturity level, ended up significantly different from the other organizations at its new RDM maturity level. Therefore, this section tentatively assumes that an organization will improve its performance as it moves upward in RDM maturity level. It first offers some anecdotal evidence supporting the desired hypothesis.

In preparing the second survey, Ellis anticipated the need to verify that *upward movement* in RDM maturity level would yield measurable *improvement* in performance and asked a series of questions in the survey around just this issue. Each of 86.5% of the respondents reported that his or her organization tried to improve its RDM maturity last year. Each of two-thirds of these attempted-improvement reporting respondents reported that his or her organization made substantial improvement in its stakeholder satisfaction and in its on-time, on-budget performance. Making substantial improvement is defined as answering "Yes" to the question, "Did you make significant improvements in the productivity of your analyst teams last year that resulted in improvements in BOTH stakeholder satisfaction with development AND the on-time/on-budget performance of projects?" Thus, each of 57.7% of the respondents of the second survey reported that his or her organization's performance improved as a result of the organization's investment in improving the productivity of its analyst teams. How much of each improvement occurred and how it was measured was not explored; rather, only *that* the improvements occurred was ascertained.

In IAG's experience, the amount of improvement an organization achieves tends to be related to:

1. the extent to which all capability areas are addressed rather than focusing on just one or two,
2. the extent to which the change team engages the whole organization in making the changes in its RDM behavior,
3. the breadth and depth of management support and the approaches taken to build this support, and
4. the extent to which the team members move the creation of a mature RDM process out of the theoretical into practice, applying, testing, and optimizing its RDM process on their current, real projects.

Measuring improvement within an organization is very challenging, since there is usually a long period between when a requirements specification is completed and when the development of the specified CA is completed for any CA of a strategic magnitude. To some extent, the way a typical organization budgets and resources its projects skews thinking to the extent that not every performance improvement is equally important. For example, many an organization has a fixed IT budget and a fixed number of staff, with very limited possibility to use subcontractors. For such an organization, cost savings are irrelevant, since it is not looking for ways to reduce the budget, but meeting its objectives faster is relevant. On the other hand, another IT organization might be heavily outsourced, and the direct cost to develop a single system is more of a driving force. In either case, stakeholder satisfaction and meeting business objectives is a central issue. Meeting these business objectives is very challenging for analysts, because sometimes, a set of superb requirements and an equally superb development effort is applied to a dumb business idea and results in low stakeholder satisfaction. Thus, measuring improvement can be difficult and an outcome that is valuable to one organization is not always valuable to another.

In any case, assessing improvement must have a baseline, either:

- comparing two sets of projects, in which the elements of one are utilizing high maturity RDM and the elements of the other are not,
- assessing the maturity of the organization about to embark on a program to increase its RDM maturity, and then auditing both change compliance and project performance, or
- implementing and complying with a set of key performance indicators (KPIs) that were simply unachievable for the organization previously.

Perhaps the most difficult part of any correct empirical demonstration of improvement is eliminating differences in project personnel and differences in developed CBSs as the source of any observed performance improvement. It would be necessary to find pairs of projects whose developed CBSs are of similar developmental difficulty and whose personnel are of similar competency. Then the data would have to show that one project from each pair significantly improved its performance while increasing its RDM maturity level, but the other project from the pair did not change its performance while staying at its current RDM maturity level.

Nevertheless, between 1993 and 2002, IAG involved 36 projects in six organizations in such an empirical study. For each subject project, IAG conducted assessments at six preset times in its life cycle, one time before project initiation, four times during the project, and one time after

project completion. IAG paired similarly sized projects and tested for between-subject variability using paired *t* tests to compare high-RDM-maturity projects and low-RDM-maturity projects within the same organization. It found that the average high-RDM-maturity project

1. had 75% fewer changes to requirements, as measured by changes per 100 requirements to factor out project size differences,
2. needed 58% less time to complete requirements, as measured in company-wide person weeks,
3. had 60% less total development costs for similarly sized functionality,

than the average low-RDM-maturity project [24].

Thus, it might be possible to set up a good empirical demonstration of performance improvement as a result of increased RDM maturity, but it takes years to get meaningful data. Nevertheless, when IAG did this experiment between 1993 and 2002, the results were fairly promising. Thus, there is some evidence supporting the desired hypothesis.

In IAG's experience, many an organization that embarks on a program of RDM process improvement sees significant and tangible benefits, as suggested by the desired hypothesis. Moreover, an organization does not need to achieve Level 4 or 5 RDM maturity to see this benefit. Using a well-defined RDM Maturity approach is a critical component of making RDM process improvement programs focused and results oriented.

7 Potential areas of future research

The two surveys happened to have examined many other phenomena in lesser detail. Some of these phenomena offer the potential for deeper examination of the effects of RDM on the SDLC.

7.1 People skills versus RDM maturity

The second survey measured the RDM skill of individual analysts using the same Levels-1-through-5 scale that it used to measure an organization's overall RDM maturity. The data show that analysts of lower skill in an organization with higher RDM maturity performed RDM better than analysts of higher skill in an organization with lower RDM maturity. This observation suggests that RDM maturity is not simply a people or skills issue—one must look at the entire RDM process.

7.2 SDLC versus requirements process maturity

We have already established that reported success rates for projects are similar among organizations in any RDM

maturity level. The second survey found also that among organizations at any RDM maturity level, the reported success rates of their projects were independent of the SDLC models, such as agile, waterfall, iterative, or visualization centric, that the organizations used to develop their CAs. Confirming this finding with other studies would help debunk the ideas that there is a best SDLC to use or that one SDLC is significantly better than another. Brooks has already stated that "The hardest single part of building a software system is deciding precisely what to build ...." [7]. RDM comes to play in the hardest part, and the SCLC comes to play in the relatively easier rest of the building.

7.3 Requirements and impact on corporate results

The data from the second survey demonstrate that organizations with low RDM maturity spend more money and expend greater effort to achieve lesser results than organizations with high RDM maturity. Since IT is pervasive in many an organization, and it affects the organization's ability to achieve its business objectives with any degree of agility, the overall financial performance of low RDM maturity organizations should be lower than that of a high RDM maturity organization. The second survey tested a number of variables and found a correlation between an organization's return on assets (ROA), a measure of the organization's efficiency in turning assets into cash, and its RDM maturity. However, it is likely that other variables will show a stronger correlation with RDM maturity.

7.4 Hygiene factors and requirements

RDM may be like Herzberg's famous motivation–hygiene theory [22] of an employee's job satisfaction. Some factors, the motivation factors, affect the employee's job satisfaction, while other factors, the hygiene factors, affect the employee's job dissatisfaction. The first survey found that organizations that focused on specific attributes in requirements documentation had very low failure rates in their IT projects, but low perceived success in these IT projects. On the other hand, organizations with high elicitation capabilities had high perceived success in their IT projects, even if they also focused on specific attributes in requirements documentation. Further research is needed.

7.5 Impact of requirements software tools on project outcome

There are over 100 RDM tools available in the market today, and the number of tools continues to grow. The second survey found that the effectiveness of an organization's use of RDM tools is correlated (1) positively with the organization's RDM maturity level and (2) either

positively or negatively with the organization's RDM effectiveness, depending on its current RDM maturity level.

Specifically, an organization at Level 1 RDM maturity benefited, albeit somewhat weakly, from using a tool, simply because the very use of a tool brings also some improvement in process, documentation, technique, skill, etc., hitting all capability areas. For a Level 2 organization, however, tool use had a negative effect on productivity. However, for a Level 3 organization, which had implemented its RDM and was presumably using tools to stabilize and standardize the RDM implementation within the organization, tool use had a strong positive effect. A typical Level 4 organization showed a positive effect from tool use. Hofmann and Lehner [23] found a negative correlation between RDM tool use and RDM productivity, but at the time, far fewer tools existed. This correlation deserves further study.

7.6 Extremely high maturity companies

The second survey found very few Level 4 RDM maturity organizations and no organizations that could be classified at Level 5. It would be very interesting and useful to be able to add data about Level 5 organizations to the results of the survey. With the current focus on RDM, we believe that that some Level 5 organizations will emerge soon. Data from the second study suggested that when organizations are differentiated largely by IT-based services offered to customers, organizations will begin to achieve sustainable competitive advantage from a high level of RDM maturity. This suggestion arises from the idea that a high RDM maturity organization expends far less RDM effort over far shorter elapsed times to achieve far better results at a far lower cost than a low RDM maturity competitor. Over time, a high RDM maturity organization will out-evolve a low RDM maturity competitor.

7.7 Causality

No degree of correlation data demonstrates causality. In either of the two studies, it is possible that some factor not measured caused the successful organizations to systematically improve their IT performance and, as a by-product, appear or become more mature in RDM. Given the very substantial performance gains documented in these studies, it would be valuable to begin finding and assessing the strengths of these other factors and to understand the relative effects of RDM maturity among all other issues a CIO needs to address.

8 Conclusion

This paper describes two studies, each involving a survey of people knowledgeable about requirements for, and the success of the development of, large commercial applications (CAs) in hundreds of large organizations from around the world.

The research question addressed by the first study was:

Does the quality of an organization's RDM process predict the success of the organization's strategic CA development projects?

The data of the survey for the first study show a high positive correlation between the quality of an organization's RDM process and that organization's successful performance on CA development projects. However, it ended up showing also that "quality of an organization's RDM process" is not correlated with any single variable and that high RDM maturity and project success are not correlated with any single task, any specific skills, or any individual's skills in the RDM process.

A re-analysis of the first study data showed that a notion of RDM maturity is measurable and permits answering the question better than did RDM quality. Consequently, the paper presents a comprehensive framework from IAG for RDM, describes a quality RDM process, and describes RDM maturity and how to measure it. The research question was then reformulated as:

Does the RDM maturity of an organization's RDM process predict the success of the organization's strategic CA development projects?

A second study was carried out using a new survey built to measure the RDM maturity of the respondents' organizations. The data of the survey for the second study show a high positive correlation between a large organization's RDM maturity and that organization's successful performance on strategic CA development projects, when success on a CA development project is measured as (1) delivering the CA on-time, on-budget, and on-function, (2) meeting the business objectives of this project, and (3) the perceived success of this project. The conclusions from these data support a tentative "yes" answer to the reformulated research question when the organization in question is large, and it uses an approach to RDM maturity that is similar to that presented in this paper. A test of survey validity between the two surveys showed a high correlation of their results, thus strengthening the conclusions of both surveys.

The conclusions of these studies add to the growing body of empirical evidence about the effectiveness of RE. They address the gap in existing empirical evidence, that there were no quantifications specifically of RE maturity, and there were no correlations of measures of RE maturity to project outcomes.

The contributions of this paper are

- the RDM maturity model that gave rise to the concept of RDM maturity,
- the concept of RDM maturity as a means of measuring the quality of an organization's RDM process,
- a definition of the six RDM capability areas and a technique for assessing an organization's scores in these areas,
- a definition of an organization's RDM maturity as a weighted average of its scores in the six capability areas,
- the second survey as an instrument for relating a large organization's RDM maturity and the success of its projects to develop strategic CAs, and
- the tentative conclusions of the second survey that a large organization with a high RDM maturity has a high probability of succeeding in its projects to develop strategic CAs.

References

1. Abbott B (2001) Requirements set the mark. InfoWorld 23(10):45–46
2. Agile Alliance Principles (2001) The agile alliance 2001. http://www.agilealliance.org/
3. Alexander L, Davis A (1991) Criteria for the selection of a software process model. In: Knafl G (ed) Proceedings of the fifteenth annual computer software and application conference. pp 521–528
4. Berry DM, Damian D, Finkelstein A, Gause D, Hall R, Wassyng A (2005) To do or not to do: if the requirements engineering payoff is so good, why aren't more companies doing it? In: Proceedings of the IEEE international conference on requirements engineering (RE), p 447, http://www.requirements-engineering.org/REnotDonePanelRE05/
5. Boehm BW, Abts C, Brown AW, Chulani S, Clark BK, Horowitz E, Madachy R, Reifer DJ, Steece B (2000) Software cost estimation with COCOMO II. Prentice Hall, Upper Saddle River, NJ
6. Bourque P, Dupuis R, Abran A, Moore JW, Tripp L (1999) The guide to the software engineering body of knowledge. IEEE Softw 16:35–44
7. Brooks F (1987) No silver bullet—essence and accidents of software engineering. IEEE Comput 20(4):10–19
8. Curtis B, Krasner H, Iscoe N (1988) A field study of the software design process for large systems. Commun ACM 31(11):1268–1287
9. Damian D, Chisan J (2006) An empirical study of the complex relationships between requirements engineering processes and other processes that lead to payoffs in productivity, quality, and risk management. IEEE Trans Softw Eng 32(7):433–453
10. Damian D, Chisan J, Vaidyanathasamy L, Pal Y (2005) Requirements engineering and downstream software development: findings from a case study. Empir Softw Eng 10(3):255–283
11. Damian D, Zowghi D, Vaidyanathasamy L, Pal Y (2004) An industrial case study of immediate benefits of requirements engineering process improvement at the australian center for unisys software. Empir Softw Eng 9(1–2):45–75
12. El Emam K, Madhavji NH (1995) A field study of requirements engineering practices in information systems development. In: Proceedings of the second IEEE international symposium on requirements engineering (RE), p 68
13. Ellis K (2008) Business analysis benchmark. Technical report, IAG Consulting, Oakville, ON, Canada, http://www2.iag.biz/benchmark-2008
14. Ellis K (2009) Business analysis benchmark—2009. Technical report. IAG Consulting, Oakville, ON, Canada http://www.iag.biz/resources/library/download-business-analysis-benchmark-2009.html
15. Fenton N, Pfleeger SL, Glass RL (1994) Science and substance: a challenge to software engineers. IEEE Softw 11:86–95
16. Galin D, Avrahami M (2006) Are CMM program investments beneficial? Analyzing past studies. IEEE Softw 23(6):81–87
17. Gibson DL, Goldenson DR, Kost K (2006) Performance results of CMMI-based process improvement. Technical report. The Software Engineering Institute, Carnegie Mellon University, Pittsburgh, PA, USA. http://www.sei.cmu.edu/reports/06tr004.pdf
18. Haase V, Messnarz R, Koch G, Kugler HJ, Decrinis P (1994) Bootstrap: fine-tuning process assessment. IEEE Softw 11:25–35
19. Hall T, Beecham S, Rainer A (2002) Requirements problems in twelve software companies: an empirical analysis. IEE Proc Softw 149:153–160
20. Harry M, Schroeder R (2000) Six Sigma: the breakthrough management strategy revolutionizing the world's top corporations. Currency, New York, NY
21. Herbsleb JD, Goldenson DR (1996) A systematic survey of CMM experience and results. In: Proceedings of the eighteenth international conference on software engineering (ICSE). pp 323–330
22. Herzberg F (1959) The motivation to work. Wiley, New York, NY
23. Hofmann HF, Lehner F (2001) Requirements engineering as a success factor in software projects. IEEE Softw 18(4):58–66
24. IAG Consulting (2002). IAG performance benchmark. http://www.iag.biz
25. International Institute of Business Analysis (2010) Business analysis body of knowledge (BABOK), Version 2.0. http://www.theiiba.org
26. International Standards Organization (2010) ISO 9000—Quality management. Viewed 10 August 2010. http://www.iso.org/iso/iso_catalogue/management_standards/quality_management.htm
27. International Standards Organization (2010) ISO 9001:2000, Quality management systems—requirements. Viewed 10 August 2010. http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=21823
28. Jones C (1996) Applied software measurement: assuring productivity and quality. McGraw-Hill, New York, NY
29. Kerton B (2006) Flawed requirements trigger 70% of project failures. Infotech Research. http://www.infotech.com/research/flawed-requirements-trigger-70-of-project-failures
30. Kitchenham B, Pfleeger SL (2002) Principles of survey research, part 4: questionnaire evaluation. SIGSOFT Softw Eng Note 27:20–23
31. Lauesen S, Vinter O (2001) Preventing requirement defects: an experiment in process improvement. Requir Eng 6:37–50

32. Leffingwell D (1997) Calculating the return on investment from more effective requirements management. Am Prog 10(4):13–16

33. Lubars M, Potts C, Richter C (1993) A review of the state of the practice in requirements modeling. In: Proceedings of the IEEE international symposium on requirements engineering (RE), pp 2–14

34. Paech B, Koenig T, Borner L, Aurum A (2005) An analysis of empirical requirements engineering survey data. In: Engineering and managing software requirements, Part 3. Springer, Berlin. pp 427–452

35. Paulk MC, Curtis B, Chrissis MB, Weber CV (1993) Capability maturity model, version 1.1. IEEE Softw 10(4):18–27

36. Requirements-engineering.org. RE Day 2005 Website. Viewed 10 August 2010. http://www.requirements-engineering.org/REday05/

37. Sadraei E, Aurum A, Beydoun G, Paech B (2007) A field study of the requirements engineering practice in australian software industry. Requir Eng J 12:145–162

38. Sawyer P, Sommerville I, Viller S (1999) Capturing the benefits of requirements engineering. IEEE Softw 16(2):78–85

39. Software Engineering Institute (2010) CMMI overview. Viewed 10 August 2010. http://www.sei.cmu.edu/cmmi/

40. Sommerville I, Sawyer P (1997) Requirements engineering: a good practice guide. Wiley, Chichester

41. The Standish Group (2009) CHAOS summary 2009. http://www1.standishgroup.com/newsroom/chaos_2009.php

42. United States Government Accountability Office (2008). Report to congressional requesters: information technology: agencies need to establish comprehensive policies to address changes to projects' cost, schedule, and performance goals, http://www.gao.gov/new.items/d08925.pdf

43. Wang Y, Court I, Ross M, Staples G, King G, Dorling A (1997) Quantitative evaluation of the SPICE, CMM, IS0 9000 and BOOTSTRAP. In: Proceedings of the third international software engineering standards symposium (ISESS), pp 57–68

44. Wiegers KE (2003) Software requirements. 2nd edn. Microsoft Press, Redmond, WA

45. Wohlin C, Runeson P, Höst M, Ohlsson MC, Regnell B, Wesslén A (2000) Experimentation in software engineering: an introduction. Kluwer Academic Publishers, Norwell, MA