

# A practical strategy for the evaluation of software tools

*Antony Powell, Andrew Vickers*  
*Department of Computer Science, University of York*  
*Heslington, York, UK.*  
*Phone: +44 1904 432722, fax: +44 1904 432708*  
*Email: [alp|andyv]@minster.york.ac.uk*

*Eddie Williams, Brian Cooke*  
*Rolls-Royce plc*  
*PO Box 31, Derby, UK.*  
*Phone: +44 1332 771700, fax: +44 1332 770921*

## Abstract

This paper describes a working strategy for software tool evaluations that has resulted from work within Rolls-Royce plc in response to the difficulty, and mixed successes, we have experienced in the selection of software tools. The lack of an acceptable methodology has meant that industrial evaluations are commonly time-consuming, fail to capture both tool and problem *knowledge* in a form suitable to aid future evaluations, and frequently give inconclusive results. Even where rigorous selection methods are used we raise the concern that tool evaluators are failing to address perhaps the most important factors in determining final success namely the non-technical or '*soft*' factors.

In an attempt to overcome some of these problems the proposed strategy provides a qualitative list of important *issues* distilled from many years experience of making tool selection decisions. This generic issue checklist is used to form *domain specific criteria* against which tools can be compared in a more quantitative manner. This process ensures traceability between issues, tool requirements criteria and supporting evidence in order to document decisions and provide assurance that all issues have been addressed. It also helps us to capture valuable corporate knowledge for future evaluations in order to become more

efficient at evaluating tools, provide more consistent criteria and to limit the risk of expensive mistakes.

This industrial perspective on tool selection will be of interest to managers and evaluators of organisations who purchase software tools. To a lesser degree the issue guidelines cover method evaluation and tool *emplacement* but further refinement and practical application is recognised. The strategy may also form the basis of a process for tool evaluations as required by higher levels of the SEI Capability Maturity Model (Humphrey, 1988; Humphrey, 1990). Finally, we hope that tool vendors will use it to provide better support for eliciting and meeting customer requirements during the evaluation process.

### Keywords

Software tools, industrial practice, evaluation, experience

## 1 INTRODUCTION

For many years the software engineering community has argued (and vendors have claimed) that significant increases in software productivity can be achieved through the effective use of software tools. Unfortunately, the process of tool selection and emplacement is inherently difficult and expected tool benefits have often failed to materialise. Most organisations who employ software tools can testify to bad experiences of tools that did not meet expectations and legends of more 'spectacular' tool disasters are commonplace. Yet software tools are a necessary and integral part of the modern software development process and are key drivers of productivity and ultimately profitability. It is therefore quite surprising that few companies appear to have a rigorous process in place for the evaluation of software tools.

This paper describes an attempt to address some of the difficulties we have experienced in tool evaluation. This experience comes from many years of selecting and introducing software tools and methods for the development of aeronautical control and monitoring systems. Despite long and very expensive evaluations for major tool purchases we still have, with the luxury of hindsight, made regrettable tool selection decisions. Here we have attempted to distill what distinguishes good from bad tool evaluations and form a more structured method to learn from these experiences and avoid costly mistakes.

We start by explaining our experience of the reasons for the difficulty of the evaluation process, give practical advice in overcoming these problems and describe limitations in the 'state of the art' in the evaluation domain (Section 2). Firstly, we address the problems of the tool selection *decision* itself including the problem of multiple and often conflicting criteria, the environment and nature of the decision, and the flawed decision making process. Secondly, we highlight the problem that evaluators tend to focus on technical capabilities of the tool at the expense of non-technical or '*soft*' factors such as human and business issues. We also acknowledge the problems of successfully introducing the tool into the project domain - a process that we term '*tool emplacement*' (as opposed to 'implementation' in order

to distinguish from actual tool usage). Tool emplacement can fail because of poor documentation, lack of compatibility with existing tools, and lack of flexibility. The interesting conclusion is that, despite basic reasons for mistakes in tool selections, decisions are still made with irrational neglect for even elementary issues.

In response to these problems, a joint working party set out to form a strategy for tool evaluation based on the strong belief that we can capture and exploit past experience to improve the quality of decision making in tool selection and emplacement. Section 3 describes the basic strategy that arose from this collaboration. It consists of a method to place 'sanity' checks on decisions and a structured process to aid organisational learning. The result is a strategy that is simple to apply and generic across different problem domains but robust enough to help improve the quality of decision making. We illustrate our approach with a simple example and also contrast our approach with those of others with whom we are aware.

We conclude by describing the more unexpected benefits of the strategy (such as the ability to develop repositories of criteria for different tool domains) and further work required (to address in more detail the problems of method evaluation and tool emplacement). The strategy is currently in use as a practical and evolving set of evaluation guidelines. Whilst the strategy may not necessarily reduce the cost and duration of evaluations (although benefits have been experienced) we believe it will help to avoid costly mistakes.

The paper expresses our concerns about current practice in tool evaluation in particular that bad decisions are made on basic but recurrent mistakes, and hence a need for a more rigorous process. We hope it will act as a prompt for evaluators and decision makers to be introspective and as a result demand higher standards from, and add learning to, a critical corporate process.

## 2. BACKGROUND

### 2.1 The problem with tool evaluations

From our practical experience we summarise the key problems of tool evaluation as follows:

**Decision criteria** - The complexity of evaluating tools appears to rise exponentially with their functionality. Evaluators face the difficulty of weighting many nebulous product criteria (such as 'features', 'usability', 'robustness' and 'quality') often without reference points or benchmarks. This is made worse by limited understanding of the problem which the tool has to solve, imperfect tool information and also, due to cost, the need to evaluate tools outside of their operational environment.

**Decision risk and uncertainty** - To add to this problem managers need quantification of the benefits of a new system against the risks of change. However, the difficulty of quantifying and comparing criteria leaves a high degree of uncertainty on the likelihood of success.

Consequently an inherently complex multi-criteria decision (even with sound understanding of the problem environment and tool capability) can degrade into a best guess supported by a weak set of figures manipulated to meet the required decision outcome.

**Decision pressures** - The search for a tool can be triggered by both internal and external stimuli. *Internal* stimuli occur as current systems (gracefully) degrade under the greater demands of a changing operational environment. Tools outlive their cost-effective lifespan by surviving on patches, unofficial procedures or 'skunk-works' with users pushed to, and beyond, their patience threshold. Alternatively a decision can be forced by *external* stimuli such as advances in technology, contractual pressure, changing standards or an irrational need to keep up with competitors tool decisions. The cost and risk of introducing the new tools means the decision to search for a tool is typically made when the existing application is on the verge of collapse. The resulting urgency means that evaluation decision gets pressured and a difficult decision becomes even more risky.

**Decision horizon** - Decision pressures can also influence the horizon on which a decision is based. Immediate budgetary demands may mean an unwillingness to pay the *prima facie* high price of a tool with disregard to the high through-life costs of apparently 'cheaper' alternatives. The result is a short term fix rather than a long term 'solution'.

**Decision making process and responsibility** - A key problem is that due to the significant risk of tool failure responsibility for the change is naturally avoided (why risk a promising career?). As a result decisions go underground, are left undocumented or are dissipated by committee dissonance. Lack of documentation leads to loss of both accountability and valuable corporate knowledge.

**Post-decision myopia** - There is a tendency for the process to end once the purchase decision is made. The tool is thrown 'over the wall' into the user area for them to cope the best they can; by which time the decision makers, evaluators (precisely those people who should understand the tool best) and any potential product 'champions' have long gone. Many tools which are technically sound fail because of inadequate consideration of this emplacement process. Inadequate management of the change process (such as lack of management support, cultural problems, poor training) can destroy any remaining chance of success.

**Inevitable or premature tool decline** - The volatile environment of tools, tool markets, vendors, standards, upgrades etc. means there is a high likelihood of a premature and somewhat abrupt end to the tool's useful lifespan. Even in the best case scenario; despite the most thorough review of a tool's capabilities, a good decision, and well managed emplacement cannot stop the eventual changing environment that renders the current tool obsolete.

These problems are considerable even when evaluating '*small-scale*' tools that apply to well defined functional problems (such as Statemate for the behavioural specification of systems) but they are exacerbated when considering 'womb to tomb' tools such as Cradle-SEE, Teamwork, RDD-100, etc. Under these circumstances the beaten and downtrodden evaluators

are left in a lose-lose situation. Little progress is made, much money is wasted and little is learnt.

## 2.2 Conventional wisdom and ‘state of the art’

The deliberately cynical account described above is intended to highlight the inherent difficulties of the evaluation process. Few are surprising, all are basic, and certainly none are new but readers may nevertheless recognise some or all of the symptoms within their own organisations. Of course, it is not difficult to proclaim sound advice based on the above observations, both our experience and conventional wisdom (Heller, 1991; Kitchenham et al., 1995; Mosley, 1992) reminds us of the importance of:

- **Managing the decision process** - It is critical to document the decision and make people involved responsible and accountable for the decision outcome. Add checks on the decision process to recognise and reduce pressures on decision makers to help avoid foolish decisions.
- **Managing tool emplacement** - Preferably the tool should be evaluated by the people who are going to use it or at the very least ensure their active participation. Careful change management should foster champions, fund appropriate training, and respond to problems as they arise.
- **Managing tool decline** - Try to recognise earlier the degradation of current tools due to internal and external triggers. Planned tool lifecycle management will allow more time for successful tool replacement.
- **Trying before buying** - Experience suggests - and we maintain - that case studies are of greater benefit than stand-alone evaluations particularly for large and important evaluations.
- **Identifying when and what to change** - At some point we need to recognise the hard truth that the tool is good and it is our process that needs change.

How come bad decisions are still made? They are all simple common-sense points but all too frequently missed. What we need to know is *how* to do them effectively, or more importantly from our point of view, what issues should be considered when devising procedures to implement these policies. Hence our reasoning for a more formal and consistent approach to evaluation of both technical and non-technical issues in tool selection.

There are polished accounts (Kitchenham et al., 1995; Shin and Lee, 1996) of mechanisms for the evaluation of tools which address many of the concerns outlined in Section 2.1 and espouse much of the wisdom listed in Section 2.2; however, from a ‘functional’ or technical point of view. It is our experience that the software issues concerned with *how* the tool may/will be used are the issues that are crucial to a successful evaluation, emplacement and use. It is these issues which we have concentrated on.

### 3. AN INDUSTRIAL STRATEGY

#### 3.1 Strategy overview

Clearly we need to pay more respect to the intricacies of both the selection decision and the *emplacement* process. To aid this we need more structured evaluation methods to give insight into the factors of tool assessment. Here we describe the approach adopted to improve the evaluation process based on the problems highlighted in Section 2. The strategy has two main characteristics:

**Defined process** - We view evaluation process as consisting of two main activities: understanding the problem environment (which the tool is to address) and understanding the tool. It is important that we capture knowledge of both these activities for use in future evaluations. The mechanism we use is a defined basic process model with three phases: pre-evaluation (Section 3.2) evaluation (Section 3.3) and post-evaluation (Section 3.4). The phases help ensure consideration of decision impacts and allows decision checkpoints to be monitored. The chosen high level of abstraction allows us to introduce element of standardisation without loss of local flexibility. Throughout we stress the importance of having a managed and documented approach to the process.

**Issue checklist** - Within the process model we introduce the technique of *criteria generation* from an *issue checklist* (Section 3.3.2). The benefits are two-fold; firstly it prompts consideration of non-technical factors and secondly it helps us to deduce our requirements more appropriately.

The focus of this paper will be the evaluation process itself but we hope the reader will recognise how, by the use of a structured evaluation process the quality of, both pre- and post-evaluation phases can be strengthened. In Section 3.4 we conclude our description of the process with the current status of this strategy and results to date of its operation.

#### 3.2 The pre-evaluation process

The pre-evaluation process is based on a first-pass through of the evaluation process described in Section 3.3. It should consist of the following events:

1. **Trigger recognition** - Recognition of the need for the tool and documentation of the trigger stimulus (i.e. the perceived problems with old tools).
2. **Problem statement** - Formation and documentation of a preliminary statement of the problem which the new tool is intended to address.
3. **Criteria assessment**. Generation of a list of the most important criteria against which tool(s) will be evaluated (based on the problem statement and issue checklist or past evaluation template).

4. **Tool search and shortlist** - Conducting tool search (including finding if anyone within the organisation has prior experience of the tool) and shortlisting of tools against the initial criteria list. Of course this selection may be based on the 'lesser of evils'.
5. **Sourcing decisions** - At some point a decision is made whether to buy a tool or develop an in-house solution. This may be a high-level policy decision.
6. **Evaluation proposal** - A proposal is produced to describe and justify the selected evaluation strategy against a first pass at the evaluation process.
7. **Decision to evaluate** - A formal decision to proceed with (or cease) evaluation of the tool(s) and trigger for full evaluation process.

The relative importance of the tool decision will determine the need for, and extent of, each of these activities.

### 3.3 The tool evaluation process

#### 3.3.1 Process and Mechanisms

Figure 1 shows the basic process by which an evaluation takes place. The pre-evaluation process is likely to have already made one high-level pass at the evaluation phase in order to shortlist tools and produce the evaluation plan. The full evaluation process is based on the concept of phased evaluations to address the need for an evaluation process which is efficient and economical for the nature and risk of the tool selection decision. We describe the stages of our process as follows:

**Demonstration** - Used as part of the initial search for tools but should only be used as a basis for the most minor tool purchases (where cost of mistakes is minimal in comparison to evaluation overhead) or where adequate information exists by which to make an informed purchase (evaluators must be confident that this meets all criteria).

**Assessment against criteria** - Evaluation of tool against the list of issues described in Section 3.3.2 (in full in Appendices A1-A8) and preferably in comparison with another tool for benchmarking purposes. This should be the minimum evaluation level against which tool selection decision made.

**Pilot study** - Significant tool selection decisions should not be made without recourse to some form of pilot study - either shadowing part of a project or applying it to a low risk part of a live project.

On completion of each technique a decision is made whether to proceed to the next level. Each evaluation process decision is recorded with its reasoning and a final report is produced to record reasons for the tool selection or rejection decision. This process guides evaluators, aids resource allocation and can be used by other parties (managers, quality departments, other evaluators) to confirm reasoning and issue coverage.

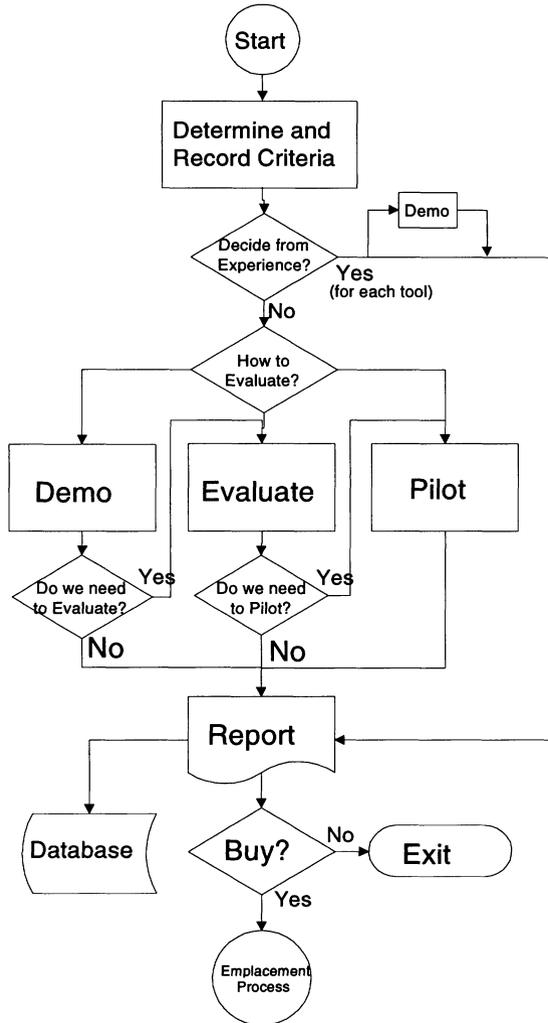


Figure 1: Basic evaluation process.

### 3.3.2 Issues in the selection of software tools

At the core of our strategy is the issue checklist shown in Figure 2 (presented in full in Appendix A1-A8) is used to form an initial set of more quantitative criteria against which a tool is evaluated. It is suggested that each tool should be assessed against a particular requirement prepared for that tool linked to a particular issue category. The requirement itself should ensure that issues are addressed and indeed provide traceability to them. When forming a business case for tool purchase, evaluators should provide evidence that all relevant issues from the list have been addressed.

The issues are listed in no particular order and may be addressed in sequence of perceived importance.

**Business issues** reflect the top level business constraints against which the business case for the tool will be made. They represent the environment for the decision and may constrain the tool evaluation process such as limiting the cost of new tools, policies of sourcing and risk to current projects. It is important to accept that sometimes it is impossible to calculate quantitative indices such as return on investment with confidence, in which case qualitative alternatives should be made explicit.

**Evaluation issues** are those which impact upon the evaluation process itself and represent the method by which evaluation must be justified. These issues form a self-check on evaluators including preconceptions, structuring to capture evaluation knowledge and selection of benchmarks.

**External reference issues** concern how the tool evaluation and emplacement relates beyond the boundaries of the organisation. The intention is to ensure that evaluators take advantage of information repositories pertaining to the tool (including past evaluations). These can significantly reduce the amount of time and therefore cost needed to evaluate a tool.

**Vendor support issues** represent concerns about futures in the uncertain tool market. These are frequently overlooked but critical particularly in high integrity and/or defence applications.

**Financial issues** represent the importance that all the costs incurred in selecting a tool be taken into account and assessed as accurately as possible. Where figures are problematic the level of uncertainty should be emphasised and bounded.

**Tool issues** represent the critical fitness for purpose of the tool and its method. These are the focus for traditional tool evaluation methods (including things assessment of how a tool affects the process and product integrity, compatibility, acceptability). It is recognised that the need to act to solve a technical problem may outweigh other problems such as vendor . And, once again we must recognise when the tool is good but our process needs changing.

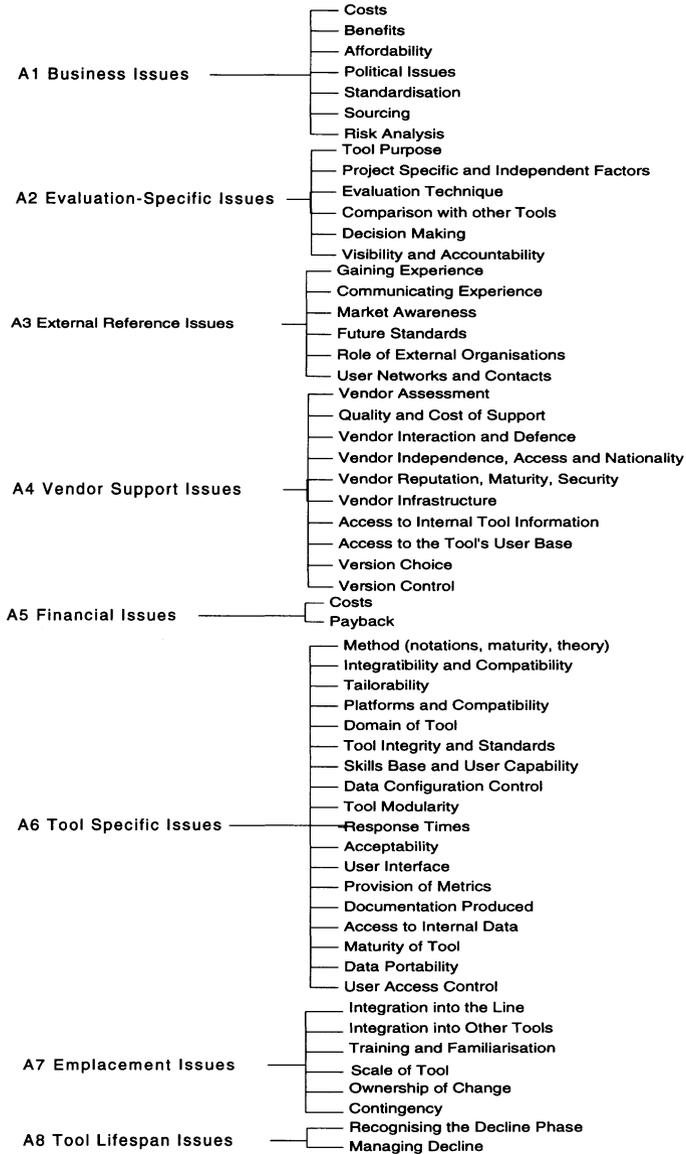


Figure 2: Issues in tool evaluation (Appendices A1-A8).

**Emplacement issues** reflect the difficulty of successfully integrating a tool into the line and achieving business benefits. They should be addressed by evaluators *before* a tool selection decision is made.

**Tool lifespan** issues reflect the need to recognise and manage the inevitable decline of a tool against the introduction of its replacement. This is frequently neglected but can impact highly on the quality of the next evaluation decision.

The list may not be 100 percent complete but we have found it to be a useful starting point for addressing the key issues that affect success in tool selection and emplacement in our organisations. It is not very qualitative, but we are not alone in making use of subjective measures, see (Dick and Hunter, 1994) for example. By structuring and documenting tool selection criteria in relation to these issues we stand more chance of ensuring completeness and sensible decisions. In addition, we can construct templates for evaluation and comparison of certain types of tools (Section 3.3.4).

### *3.3.3 Post-evaluation (emplacement and decline) process*

Development of the process for the emplacement and decline of tools has been identified as an issue for further enhancement of the evaluation process. Some of the issues of emplacement and decline have been discussed elsewhere in this document. We believe that more empirical research is necessary into the factors that affect successful emplacement of software tools. We also believe that there is a requirement for more advanced monitoring changes in the internal and external operating environment of a tool and to make replacement decisions earlier. These issues will be the subject of future work.

### *3.3.4 Operation and results*

The described strategy has been incorporated into local quality guidelines for tool evaluations since mid-1995. Feedback forms have been used to communicate results of strategy application within operational areas. We have already observed considerable benefit in compiling knowledge from tool evaluation activities on two levels:

**Collection of domain criteria checklists** - Different tool domains have varying criteria and weightings, which are capable of being recorded and reused. Such tool domains include general design, requirements management, systems engineering and requirements capture, configuration, testing, document managers, and control specification. There are also academic examples of domain criteria; for instance, Schamp (1995) records the criteria for configuration management tools. These internal (company generated) and external criteria are forming the basis of a tool requirement repository which will be of use in future evaluations.

**Collection and publication of tool evaluation results** - Through use of organisation-wide process improvement teams we have stimulated cross-organisation communication about tools, toolsets, and tool experience. Tool evaluation intentions and results are published to all relevant parties - allowing us to make best use of organisational knowledge, track weaknesses in particular tools and ensure a proactive attitude in tool policy.

Time will allow us to evaluate and refine the strategy whilst improving our knowledge repository. Initial feedback has indicated strong potential for cost savings and considerable support for valuable cooperation.

#### **4. AN EXAMPLE**

A real demonstration of the usefulness for our approach would require a number of experiments, involving the assessment of a number of different tools, using a variety of evaluation processes (e.g. (Misra, 1990)). Such a presentation would be a paper in itself, but it is useful to provide a sketched example of how our technique could be used to evaluate a software tool. We consider a mythical tool called the Requirements Management Tool (hereafter referred to as RMT) and its evaluation by company XYZ. The scenario is not derived from practices within Rolls-Royce.

1. Members of XYZ have a need for some form of requirements management tool.
2. Members of XYZ receive some technical literature on the RMT tool from the tool manufacturer.
3. The tool looks promising and the XYZ members review their existing evaluation library and discover a review of a requirements specification tool (RST) that offers some similar features. The review for RST is retrieved and previous experience shows that it is the integration of such tools that is often the critical success factor. An RMT review check list is drawn up which includes the following issues:
  - Business issues - Is the RMT seat price less than £2000?
  - External reference - Is there an RMT user group?
  - Vendor assessment - Is the vendor a known supplier to XYZ?
  - Tool specific - Does the tool link to the existing DMT tool?
4. A demonstration is arranged and the tool looks sufficiently appropriate that an evaluation copy is ordered.
5. During evaluation the RMT tool scores well technically, but it cannot integrate with the DMT tool. As the DMT tool has been used successfully for many years it is decided not to pilot the RMT tool and the evaluation process is terminated.
6. The RMT tool is rejected and the process and supporting arguments are logged in XYZ's tool evaluation library.
7. XYZ management are pleased that an inappropriate tool has not been purchased and that the evidence for such a decision has been formally recorded.

This is a deliberately artificial example but shows how a simple procedure can act as a means of protecting a company from an isolated evaluation performed purely on performance criteria. The process encouraged an orderly evaluation based on a variety of issues, making

use of previous knowledge where appropriate. Most importantly, the reasons for the rejection are recorded and placed within the corporate memory.

## 5. RELATED WORK

Whilst the process presented in this paper is primarily a description of our current practice, it is still appropriate to take some time to place our work in the context of that of others of which we are aware. The significance of software tools to the quality of an overall engineering capability has been documented (Polvia, 1992) and many have written of the importance of software tool evaluation (Heller, 1991; Kitchenham et al., 1995; Mosley, 1992). Our ideas on creating a process for this evaluation (and subsequent emplacement) of tools is not uncommon to that of Rowley's (Rowley, 1993) goal of treating tool evaluation as a project in its own right. Such a view aids repeatability and encourages the recording of decisions and rationale, as well as introducing group accountability. This is a good thing and is not dissimilar to the work of Mosley for instance (Mosley, 1992).

The inclusion of evaluation issues other than just functionality and performance is recognised elsewhere, in particular with regard to the role of different types of tool (Anderson, 1989; Cheng and Pane, 1991; DeSantis, 1994; Miller and Jeffries, 1992). Other organisations have recognised the wider scope of tool evaluations than just performance and functionality and have instigated their own corporate advice centres, see (Scheffler and Marshall, 1991).

Other tool evaluation checklists do exist (Jeanrenaud and Romanazzi, 1994; Klopping and Bolgiano, 1994) but can be criticised (as indeed can ours) for being overly procedural and bureaucratic. McDougal (McDougal and Squires, 1995) discusses the problems of checklists, but we can only really draw on our own experiences. Checklists are laborious to fill in, but they do provide a focus or thought, and do draw the evaluators away from a technocentric approach to evaluation.

## 6. FINAL COMMENTS

This paper has presented a brief overview of a more detailed tool evaluation process currently in use within Rolls-Royce plc. It is intended to be simple, generic and a working strategy that will evolve with use in order to capture and exploit valuable organisational knowledge generated during tool evaluations. The strategy addresses the weaknesses of the existing evaluation process by attempting to ensure that all issues are structured and addressed particularly *soft* factors overlooked by many formal evaluation techniques. The result is a quality evaluation process that we believe is more effective, more efficient, and less risky.

We believe that the strategy improves our chances of ensuring that conventional wisdom about tool evaluation (Section 2.2) is addressed and against these goals it performs well (and considerably better than no process whatsoever!) However we emphasise the need for exposure of the strategy to generations of tool evaluations and its necessary evolution to

capture more detailed evaluator knowledge. Further work on the strategy is required to capture experiences of tool emplacement (which is critical in determining the success of a tool) and enhance it to provide greater support for the more difficult area of method evaluation. These experiences will help us to define precisely how business cases for tool acquisition should be constructed, though we will need more qualitative measures, see Williams (Williams, 1992) for example. We hope to define a business case template which would use the results of our process to format a standard business case for tool selection. We see this as a significant step forward. Finally, we are in the process of assembling a knowledge repository on tool capabilities and application domain criteria. Taking this to the logical conclusion we could lay a foundation for a repository of this kind on the World Wide Web (WWW) thereby making a global tool information knowledge-base (clearly this has problems of regulation, evidence, objectivity, copyright etc. despite the attraction for tool purchasers).

We have made (and perhaps laboured) the point that there are many issues in the use of tools of which only one is the technical capability. Experience has taught us that the subtleties of tool evaluation, selection and emplacement must not be underestimated. A more successful process will come from greater attention to the organisational and behavioural issues. Tools can provide business benefits but all too easily they can be eroded by unsuccessful introduction. In reality it is even harder than this as the introduction of a tool to an organisation will itself change the nature of the organisation. A tool may therefore contribute to its own lack of success because of the changes it introduces! Selecting tools is difficult (good from bad) and whilst we have not addressed all the issues, we believe the techniques presented in this paper are a step towards ameliorating the problems of tool selection.

Finally, we believe that our experience is of benefit both to the users of tools, the developers, and vendors as a means of bringing the parties together for mutual benefit. We encourage use of checklist in Appendix A1-A8 to appraise the reader's organisation's process (if they have one) and we would welcome feedback. Whilst strategy may not necessarily prevent disasters occurring we at least have an opportunity to learn from them.

## 7. ACKNOWLEDGMENTS

Peter Jeffery and John Anderson for contributions to the ideas presented in earlier drafts of this paper. Peter Summers of Rolls-Royce who contributed to early versions of the evaluation strategy. Mike Burke at the British Aerospace Dependable Systems Computing Centre at the University of York. Rolls-Royce plc for permission to publish. The anonymous reviewers for providing sharply focused observations on improving the paper.

*Please note: This document represents the opinions of the authors and in no capacity does it indicate official policy for Rolls-Royce plc.*

## 8. REFERENCES

- Anderson, E. E. (1989) A Heuristic for Software Evaluation and Selection. *Software: Practice and Experience*, 19(8), 707-717.
- Cheng, D. Y., and Pane, D. M. "An Evaluation of Automatic and Interactive Parallel Programming Tools." *Supercomputing '91*, Albuquerque, USA, 412-423.
- DeSantis, J. (1994) Evaluating Multi-Platform Development Tools. *Object Magazine*, 4(4), 41-44.
- Dick, R., and Hunter, R. "Subjective Software Evaluation." *Software Quality Management II: Building Quality into Software*, Edinburgh, UK, 321-334.
- Heller, R. S. (1991) Evaluating Software: A Review of the Options. *Computers and Education*, 17(4), 285-291.
- Humphrey, W. A. (1988) Characterising the Software Process: A Maturity Framework. *IEEE Software*.
- Humphrey, W. S. (1990). *Managing the Software Process*, Addison-Wesley.
- Jeanrenaud, J., and Romanazzi, P. "Software Product Evaluation: A Methodological Approach." *Software Quality Management II: Building Quality into Software*, Edinburgh, UK, 59-69.
- Kitchenham, B., Pickard, L., and Pfleeger, S. L. (1995) Case Studies for Method and Tool Evaluation. *IEEE Software*, 12(4), 55-62.
- Klopping, I. M., and Bolgiano, C. F. (1994) Effective Evaluation of off-the-shelf Microcomputer Software. *Office Systems Research Journal*, 9(1), 46-40.
- McDougal, A., and Squires, D. (1995) A Critical Examination of the Checklist Approach to Software Selection. *Journal of Educational Computing Research*, 12(3), 263-274.
- Miller, J. R., and Jeffries, R. (1992) Interface-Usability Evaluation: Science of Trade-Offs. *IEEE Software*, 9(5), 97-102.
- Misra, S. K. (1990) Analysing CASE System Characteristics: Evaluative Framework. *Information and Software Technology*, 32(6), 415-422.
- Mosley, V. (1992) How to Assess Tools Efficiently and Quantitatively. *IEEE Software*, 9(3), 29-32.
- Polvia, P. (1992). "A Comprehensive Model and Evaluation of the Software Engineering Environment." *Information Resources Management Association International Conference*, Harrisburg, USA, 302-307.
- Rowley, J. E. (1993) Selection and Evaluation of Software. *ASLIB Proceedings*, 45(3), 77-81.
- Schamp, A. (1995) CM-Tool Evaluation and Selection. *IEEE Software*, 12(4), 114-118.
- Scheffler, F. L., and Marshall, R. R. "The Software Technology Support Centre: Help for Acquiring Software Tools." *National Aerospace and Electronics Conference*, Dayton, OH, USA, 647-653.
- Shin, H., and Lee, J. (1996) A Process Model of Application Software Package Acquisition and Implementation. *Journal of Systems and Software*, 32, 57-64.
- Williams, F. (1992) Appraisal and Evaluation of Software Products. *Journal of Information Science, Principles and Practice*, 18(2), 121-125.

## APPENDICES A1-A8

### A1 - Business issues

These strategic issues reflect the top level business constraints against which the business case for the tool will be made. It is important to accept that sometimes it is impossible to calculate quantitative indices such as return on investment with confidence, in which case qualitative alternatives should be made explicit.

- **Costs** - The total cost, from licences to attaining necessary competencies, should be calculated and depending on the level of confidence an upper bound be determined.
- **Benefits** - The benefits should be estimated in as quantitative way as possible but where qualitative arguments from considerations such as those in the rest of this section are used they must be made explicit.
- **Affordability** - There should be adequate resources - human, financial and material to perform the assessment. Adequate evaluation time should be budgeted.
- **Political issues** - The constraints arising from contractual commitment or the need to visibly conform to an industrial consensus should be addressed.
- **Standardisation** - Evaluators should recognise policy decisions on preferred methods, notations, or platforms to which tools should conform. They should also try to be aware of any likely future directions for relevant standards.
- **Sourcing** - The policy as to the selection of third party, in-house or commissioned tools should be identified and agreed.
- **Risk analysis** - Risk analysis should be performed before any significant purchase decision is made.

### A2 Evaluation-specific issues

Evaluation issues are those which impact upon the evaluation process itself. Evaluators should address these issues when planning an evaluation. These include:

- **Analysis of the purpose of the tool** - The evaluator should have a clear understanding of the process and technical activity which the tool is intended to support. The evaluator should also understand the role which the tool is intended to fulfil and ensure that the means to meet this are addressed by the evaluation process.
- **Project specific and project independent factors** - Evaluators should address both project specific (e.g. specific platform and network requirement) and non-project specific factors (e.g. quality of the method), which may require evaluation by different personnel.
- **Evaluation technique** - Evaluators should consider the comparative value and cost of different evaluation mechanisms:

- demonstrations (by the supplier);
  - structured assessments using criteria checklists;
  - pilot studies.
- **Comparison with other tools** - The evaluator should form a strategy for comparison of one tool against another including the baselining and use of scoring systems for prioritisation.
  - **Decision making** - Evaluators should clarify the basis upon which decisions will be made including who is responsible for the tool selection decision.
  - **Visibility and accountability** - Evaluators should provide visibility of the evaluation to all interested parties and ensure accountability and documentation of all decisions.

### A3 External reference issues

External organisation issues concern how the tool evaluation and emplacement relates beyond the boundaries of the organisation. These include:

- **Gaining experience** - The experience of other users, projects and/or companies should be identified and exploited where possible.
- **Communicating experience** - The sharing of experience with other users, projects and/or companies should be an integral part of the evaluation process.
- **Market awareness** - Line departments should keep abreast of the tool market to ensure that all potential solutions are considered and that the timing of tool evaluations and purchase is appropriate.
- **Future standards** - Line departments should keep abreast of potential changes to methods and standards which may affect the lifespan of the tool.
- **Role of external organisations** - Line departments should make active use of external organisations who can provide consultancy, tool information and contacts.
- **User networks and contacts** - Line departments should establish and maintain links with other users to ensure that experience continues to shape the use of the tool once it is in operation in the line.

### A4 Vendor support issues

Vendor support issues represent concerns about futures in the uncertain tool market. These include:

- **Vendor assessment** - Evaluators should consider undertaking a formal vendor assessment exercise.

- **Quality and cost of support** - Evaluators should balance the level and quality of technical/product support provided and the cost of alternative support options against emplacement needs.
- **Vendor interaction and defence** - Evaluators should appraise and define the nature of the relationship between the vendor and the organisation including quality and style of interaction (sales driven, technically capable and responsive) and form barriers as a defence from unnecessary interruption.
- **Vendor independence, access and nationality** - Evaluators should establish the acceptability of the vendor in terms of their independence, ease of access and problems of nationality.
- **Vendor reputation, maturity and security** - As far as possible evaluators should ensure the vendor's long-term commitment to the product, the standards to which they operate and their likely survival as a supplier.
- **Vendor infrastructure** - Evaluators should ensure that the infrastructure of the vendor is suitable to effectively support such a product.
- **Access to internal tool information** - Evaluators should test vendor openness about the tool's construction, quality of supporting documentation and tool certification evidence to ensure sufficient contingency.
- **Access to the tool's user base** - Evaluators should seek to consult reference sites and attend user groups for mutual benefit and co-operation and to gain a feel for the vendor.
- **Version choice** - Evaluators should ensure that they evaluate the right version of tool to used. Avoid evaluating beta versions.
- **Version control** - Evaluators should establish the consequential risks of changes in the version of the tool on support software and interfaces with other tools.

## **A5 Financial issues**

It is important that all the costs incurred in selecting a tool be taken into account and assessed as accurately as possible. Where figures are problematic the level of uncertainty should be emphasised and bounded.

- **Costs** - Evaluators should calculate the costs arising from (at least) tool licences and maintenance, hardware support, achieving and holding the necessary competence level of personnel, migrating from or integrating with the current toolset and administration.
- **Payback** - Evaluators should estimate the amount potentially saved by automating process and/or lower cost of quality.

## A6 Tool specific issues

Tool issues represent the critical fitness for purpose of the tool and its method. These include:

- **Method (notations, maturity, theory)** - The underlying method should preferably be proven and accepted in the field and within the intended user community. The risk of method obsolescence should be minimised.
- **Integrability and compatibility** - The tool should be demonstrably compatible with existing and prospective tools and techniques.
- **Tailorability** - Evaluators should consider the level of need and provision for tailorability and customisation, but this should be contrasted with potential lack of standardisation.
- **Platforms and compatibility** - Evaluators should address the risks of platform dependence and version incompatibility.
- **Domain of tool** - Evaluators must understand the purpose, functions and applicability of the tool against the problem domain.
- **Tool integrity and standards** - Evaluators should consider tool integrity and standards and should ensure that the integrity of the tool meets its intended use.
- **Skills base and user capability** - Evaluators should consider the level of skills required against currently available skills.
- **Data configuration control** - The tools should allow the user to manage the input and output data in a configurable way.
- **Tool modularity** - Evaluators should consider the benefits of buying tools in discrete functional units and should cost unused features.
- **Response times** - Evaluators should ensure that the tool can withstand realistic usage and demand.
- **Acceptability** - Evaluators should consider the tool's acceptability internally and externally, prejudices should be surfaced.
- **User interface** - Evaluators should be sure that the user interface will be acceptable to the users (by understanding the tool and the user) .
- **Provision of metrics** - The ability of the tool to support any required metrics programme should be considered e.g. such that the effectiveness of the data management (and other) facilities can be judged.
- **Documentation produced** - The tool should enable documentation to be produced of sufficient quality to be easily introduced into existing company standards. This includes the ability to produce sections intended to merge with other documentation sources.
- **Access to internal data** - The need for, and provision of, access to internal data structures should be examined.

- **Maturity of tool** - Evaluators should be aware of the maturity of the tool in the market place, e.g. by examining the level of problems being solved by patches, upgrades etc.
- **Data portability** - Evaluators should ensure that data from the new tools should be machine portable to other tools already in use (or planned to be brought into use).
- **User access control** - Evaluators should consider if the tool has appropriate internal security mechanisms to ensure control of user access at appropriate levels.

## **A7**    **Emplacement issues**

Emplacement issues reflect the difficulty of successfully integrating a tool into the line and achieving business benefits. They should be addressed by evaluators before a tool selection decision is made. These include:

- **Integration into the line** - The steps for introducing the tool including organisation of human resources, transfer of work and overcoming problems should be made explicit.
- **Integration into other tools** - The practical aspects and cost of getting tools to communicate and operate effectively should be reviewed.
- **Training and familiarisation** - Evaluators should explain and cost a strategy for training and familiarisation to support tool introduction including level and type required, internal or externally sourced, and provision of resources. The assumed pre-requisite skill of the trainee should be established.
- **Scale of tool** - The introduction strategy should reflect the scalability of the tool in terms of full-scale introduction on a project or selected (gradual) introduction.
- **Ownership of change** - Change owners and tool owners must be identified, consulted and given responsibility to ensure success of the change process.
- **Contingency** - Change owners should minimise the risk of introducing a tool via the use of contingency plans.

## **A8**    **Tool lifespan issues**

Tool lifespan issues reflect the need to recognise and manage the inevitable decline of a tool against the introduction of its replacement.

- **Recognising the decline phase** - Tool owners and users should monitor tool usage to identify and make contingency for tool decline.
- **Managing decline** - Phase-out and replacement of tools should be managed as part of the tool evaluation and emplacement process.

## **BIOGRAPHY**

**Antony Powell** is a researcher at the Rolls-Royce University Technology Centre at the University of York. Before this Antony worked for Rolls-Royce and Associates helping to introduce a number of measurement and process improvement activities across the Rolls-Royce group. His research interests include software measurement, process improvement, tool evaluation and software change.

**Dr Andrew Vickers** is a Lecturer in the Department of Computer Science at the University of York. His research interests include Requirements Engineering, Software Architecture, Technology Transfer, and Safety-Critical Systems. Dr Vickers is a member of the High-Integrity Systems Engineering group at York. As part of this group, he is Assistant Director of the Rolls-Royce Systems and Software University Technology Centre, and Cotechnical Leader of the associated ASSET process improvement programme. He is a member of INCOSE, and the Industrial Liaison Officer for the British Computer Society's Requirements Engineering Specialist Group.

**Dr Eddie Williams** works for Rolls-Royce and Associates in Derby leading the High Integrity Systems and Software Centre within Rolls-Royce plc.

**Dr Brian Cooke** works for Rolls-Royce and Associates in Derby leading a Systems and Software Metrics initiative throughout Rolls-Royce plc.