Requirements Eng (2006) 11: 102–107
DOI 10.1007/s00766-005-0021-6

ORIGINAL ARTICLE

Roel Wieringa · Neil Maiden · Nancy Mead
Colette Rolland

Requirements engineering paper classification and evaluation criteria: a proposal and a discussion

Received: 29 July 2005 / Accepted: 18 October 2005 / Published online: 24 November 2005

Keywords Requirements engineering research · Research methods · Paper classification · Paper evaluation criteria

1 Introduction

In recent years, members of the steering committee of the IEEE Requirements Engineering (RE) Conference have discussed paper classification and evaluation criteria for RE papers. The immediate trigger for this discussion was our concern about differences in opinion that sometimes arise in program committees about the criteria to be used in evaluating papers. If program committee members do not all use the same criteria, or if they use criteria different from those used by authors, then papers might be rejected or accepted for the wrong reasons. Surely not all papers should be evaluated according to the same criteria. Some papers describe new techniques but do not report on empirical research; others describe new conceptual frameworks for investigating certain RE problems; others report on industrial experience with existing RE techniques. Other kinds of papers can also be easily recognized. All of these types of papers should be evaluated according to different criteria. But we are far from a consensus about what classes of paper we should distinguish, and what the criteria are for each of these classes.

We see a variety of evaluation criteria in journals too. At one extreme is the set of nine genres used by *IEEE Software* [15], all of which have different evaluation criteria. At the other extreme is the single paper class recognized by the *Requirements Engineering Journal*, which has the following evaluation criteria: originality, utility, technical contribution, and relation to previous work. Apparently, the only paper class recognized by the *Requirements Engineering Journal* is a paper describing an original and useful solution technique. This corresponds to the "how to" genre of *IEEE Software*. This leaves authors and reviewers for the *Requirements Engineering Journal* in the dark about how other classes of papers should be judged, such as experience reports, empirical studies, or tutorials, none of which describe an original technique. This might lead to the use, by reviewers, of evaluation criteria unknown to authors, or even to the use of mutually inconsistent evaluation criteria by different reviewers of the same submission.

The calls for papers of successive RE conferences, in which some of us acted as program chair, show an evolution of paper classification and evaluation schemes. Each scheme was based on the experience of the previous chair, and we tried to pass on our experience to the next chair. We also discussed our ideas with other members of the Steering Committee of the RE conferences, and with RE researchers outside the Steering Committee. This short note presents the outcome of those discussions in the form of a proposal for paper classification and a set of evaluation criteria for different paper classes. We hope to include more people in the discussion and thereby further improve the classification and evaluation scheme.

In Section 2, we sketch the rationale for our classification. Section 3 presents the classification, and Section 4 concludes with a discussion of background ideas and related work.

R. Wieringa (✉)
University of Twente, Twente, The Netherlands
E-mail: roelw@cs.utwente.nl

N. Maiden
City University, London, UK
E-mail: n.a.m.maiden@city.ac.uk

N. Mead
Software Engineering Institute, Carnegie Mellon university, Pittsburg, PA 15213-3810, USA
E-mail: nrm@sei.cmu.edu

C. Rolland
Université Paris1, Panthéon Sorbonne, Paris, France
E-mail: rolland@univ-paris1.fr

2 Rationale for the classification

The starting point of the classification scheme is the observation that much of what is called research in our community is really design. We propose new techniques for doing RE. Designing is, by definition, the activity of proposing a technique for a purpose [19, 26]. To do research, on the other hand, is to investigate something in a systematic manner. The outcome of designing is an artifact, such as a new technique, notation, device, or algorithm. The outcome of research is new knowledge.

Researchers in such diverse fields as product development and systems engineering have identified the engineering cycle as the logical structure of engineering activity [1, 2, 6, 13, 16, 20, 23]. This is because the engineering cycle is basically the structure of rational decision-making [14, 18].

2.1 The engineering cycle

Design and research are distinct but closely related, as can be seen in the following list of activities, which we call the *engineering cycle*. We demonstrate this cycle with papers that underpin, report, and apply the University of Toronto's *i** approach

(a) **Problem investigation**: Investigation of the current situation. For example, we may investigate the current way of doing RE, such as El Emam and Madhavji's [10] field survey of requirements practices for information systems development.
(b) **Solution design**: Propose an improvement to the current situation. For example, we may propose a new RE technique, such as the *i** approach described in Yu and Mylopoulos [27].
(c) **Solution validation**: Investigation of proposed solution properties. For example, we may investigate the properties of a new RE technique such as *i**, and predict whether this will improve the current way of doing RE in a certain aspect. Easterbrook et al. [9] recently investigated whether *i** models, applied with a viewpoints approach, would lead to improved requirements discovery.
(d) **Solution selection**: The literature has shown a succession of *i** proposals each an improvement over the previous one. The process by which improvements were selected has however never been reported about, so we cannot give a reference here.
(e) **Solution implementation**: Realizing the selected solution, for example, introducing a new RE technique in an organization. Maiden et al. [17] report how *i** was implemented as part of the RESCUE process to specify requirements for air traffic management systems.
(f) **Implementation evaluation**: Investigation of the new situation. For example, we may investigate the practice of RE in an organization, where this organization has recently introduced a new way of doing RE. Bush [5] reports how the *i** approach introduced previously in the UK's National Air Traffic Services through the RESCUE process later informed analysis of safety goals.

The engineering cycle is a list of activities, not a sequential program. People may start with problem analysis, or with solution design, or do both at the same time. We may even start with validation. Cross showed that experienced designers develop their understanding of the problem in parallel with designing a solution and validating the solution properties [7]. And Witte [25] showed that in major management decision processes, all tasks in the cycle are performed in parallel. Figure 1 shows the activities in the engineering cycle, plus the use activity, and some of their non-sequential impact relationships.

The engineering cycle is a classification of engineering tasks and their logical relationships. To justify that a design solves a problem, the designer should refer to an investigation of the problem. To justify the selection of one solution rather than another, the designer should

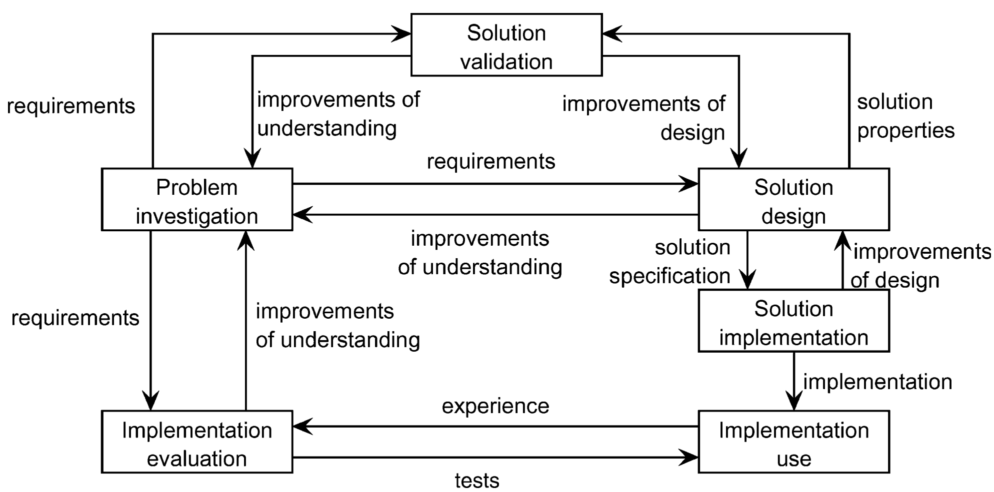

Fig. 1 Activities in the engineering cycle, plus the use activity. *Boxes* represent activities, *arrows* represent impacts. There is no preferred sequential relationship among the activities

refer to the different properties of the solutions as uncovered by solution validations. To justify an implementation, the designer should refer to the solution design that had been chosen.

We now classify RE papers into those that describe research activities (Section 2.2), a design activity (Section 2.3), or other relevant activities related to the engineering cycle (Section 2.4).

2.2 Research activities

The engineering cycle contains three research activities, namely

problem investigation (which problems exist in RE practice?),
solution validation (what are the properties of a proposed solution?), and
implementation evaluation (what are the experiences with this implemented solution?).

The first and third of these tasks are very hard to distinguish in practice, because in both kinds of research, we investigate the use of RE techniques in practice. In problem investigation, we do this to understand problems with RE techniques; in implementation evaluation, we do this to understand the use of a particular technique in practice. We propose to group these two activities together under the heading of *evaluation research*.

The second research task, solution validation, has a quite different nature, because here we investigate the properties of a technique not yet implemented. As illustrated by a recent historical study of the engineering sciences, this is the core business of engineering research [3]. Engineering researchers are in the business of proposing new techniques and investigating their properties. Civil engineers propose new techniques for building roads and investigate their properties; aeronautic engineers propose new techniques for flying aircraft and investigate their properties, etc. The important difference with evaluating existing situations is that in validation research, the techniques investigated are novel and have not yet been implemented in practice. One would not expect field research to be a useful research method in validation research. Mathematical analysis or laboratory experimentation, on the other hand, would be useful research methods. We propose to give this research task a class of its own, called *validation research*.

The difference between validation research and evaluation research is that in the first, techniques not yet implemented in practice are investigated, whereas in the second, techniques-in-practice are investigated. As with all research, hypotheses (expected outcomes) may be falsified. Techniques may not work as expected, and both kinds of research, validation, and evaluation, can send the researcher back to the drawing board to improve the technique.

2.3 Design

The second task in the engineering cycle, solution design, is a creative task in which some new technique is proposed. To describe a new technique, we must explain its ingredients, show how these fit together, explain the technique's intended use, illustrate it with an example, and state how we think it works. Solution design is not a research activity, but engineering journals and conferences contain many contributions that describe a design. This is useful for other engineering researchers even if the design is not validated, because they could replicate the technique and validate its properties, or use it to solve their own problems, which might be problems the designer of the technique had not considered. We call contributions that fall under this class as *proposal of a solution*.

2.4 Other: new conceptual frameworks, opinions, and experiences

Both in research and in design, we use conceptual frameworks to structure the world we are investigating or designing. For example, we view RE as a goal-analysis activity, a specification activity, a negotiation activity, or a problem framing activity. Each of these views comes with a framework consisting of concepts such as goals, objects, stakeholders, frames, etc. Researchers and designers will usually adopt an existing conceptual framework, but occasionally they may develop a totally new one. This is different from research as well as from design. Research presupposes a conceptual framework that structures the world to be investigated; in order to make observations at all, one must have a language in which to describe the observations. Developing such a language is a philosophical activity. Similarly, developing the conceptual framework by which we describe a solution technique is not the same thing as inventing a design in terms of a given view of the world. We call papers that describe a new conceptual framework, implying a new way of viewing the world, *philosophical papers*. These have their own evaluation criteria, which are different from the criteria by which research papers or design papers are evaluated.

Yet another class of papers is those that express an opinion. These do not describe new research results, designs, or conceptual frameworks, but rather are the author's opinions of what we should do. Authors may express opinions about the desirable direction of RE research, what is good or bad about something, what as a community we should do or not do, or anything else about values and preferences. Examples are the viewpoints in *Requirements Engineering Journal* and columns in *IEEE Software*. These papers too have their own evaluation criteria, and we call these papers *opinion papers*.

A final paper class that is valued at RE conferences is experience papers written by practitioners in the field.

These describe a personal experience using a particular technique. They do not propose a new technique—that would be a solution proposal. They are not scientific experiments—that would be an evaluation or validation paper. They need not even contain lessons learned: that would make it an evaluation research paper, evaluating experiences with a technique in practice by means of action research. The criterion here is that the paper describes the personal experiences of the author and that these are interesting enough to communicate to other practitioners. We propose to call this class of papers *experience papers*.

We have not been able to find paper classes other than the ones above, but we remain open for motivated proposals for other paper classes.

3 The classification

We now summarize the classification, and extend it with a proposal for evaluation criteria in each paper class.

Evaluation research: This is the investigation of a problem in RE practice or an implementation of an RE technique in practice. If it reports on the use of an RE technique in practice, then the novelty of the technique is not a criterion by which the paper should be evaluated. Rather, novelty of the *knowledge claim* made by the paper is a relevant criterion, as is the soundness of the research method used. In general, research results in new knowledge of causal relationships among phenomena, or in new knowledge of logical relationships among propositions. Causal properties are studied empirically, such as by case study, field study, field experiment, survey, etc. Logical properties are studied by conceptual means, such as by mathematics or logic. Whatever the method of study, it should support the conclusions stated in the paper.

Evaluation criteria for this kind of paper are

Is the problem clearly stated?
Are the causal or logical properties of the problem clearly stated?
Is the research method sound?
Is the knowledge claim validated? In other words, is the conclusion supported by the paper?
Is this a significant increase of knowledge of these situations? In other words, are the lessons learned interesting?
Is there sufficient discussion of related work?

Proposal of solution: This paper proposes a solution technique and argues for its relevance, without a full-blown validation. The technique must be novel, or at least a significant improvement of an existing technique. A proof-of-concept may be offered by means of a small example, a sound argument, or by some other means.

Evaluation criteria are

Is the problem to be solved by the technique clearly explained?
Is the technique novel, or is the application of the techniques to this kind of problem novel?
Is the technique sufficiently well described so that the author or others can validate it in later research?
Is the technique sound?
Is the broader relevance of this novel technique argued?
Is there sufficient discussion of related work? In other words, are competing techniques discussed and compared with this one?

Validation research: This paper investigates the properties of a solution proposal that has not yet been implemented in RE practice. The solution may have been proposed elsewhere, by the author or by someone else. The investigation uses a thorough, methodologically sound research setup. Possible research methods are experiments, simulation, prototyping, mathematical analysis, mathematical proof of properties, etc.

Evaluation criteria are similar to those for evaluation research

Is the technique to be validated clearly described?
Are the causal or logical properties of the technique clearly stated?
Is the research method sound?
Is the knowledge claim validated (i.e., is the conclusion supported by the paper)?
Is it clear under which circumstances the technique has the stated properties?
Is this a significant increase in knowledge about this technique?
Is there sufficient discussion of related work?

Philosophical papers: These papers sketch a new way of looking at things, a new conceptual framework, etc.

Evaluation criteria are

Is the conceptual framework original?
Is it sound?
Is the framework insightful?

Opinion papers: These papers contain the author's opinion about what is wrong or good about something, how we should do something, etc.

Evaluation criteria are

Is the stated position sound?
Is the opinion surprising?
Is it likely to provoke discussion?

Personal experience papers: In these papers, the emphasis is on *what* and not on *why*. The experience may concern one project or more, but it must be the author's personal experience. The paper should contain a list of lessons learned by the author from his or her experience. Papers in this category will often come from industry practitioners or from researchers who have used their tools in practice, and the experience will be reported without a discussion of research methods. The evidence presented in the paper can be anecdotal.

Evaluation criteria are

Is the experience original?
Is the report about it sound?
Is the report revealing?
Is the report relevant for practitioners?

Papers can span more than one category, although some combinations are unlikely. It is quite possible to write a paper proposing a new technique and presenting a sound validation of the technique, ending with a discussion in which the author airs his or her opinion about what other researchers should do. The point of this classification is not to force authors to write papers that fit within one class. The purpose is to avoid papers that belong to one class to be evaluated by criteria that apply to another class.

Evaluation research and validation research papers should use a sound research method; this is not a criterion for the other paper classes. We deliberately refrain from listing all possible sound research methods. The set of possible research methods is only bounded by the creativity of the researcher. Well-known empirical research methods include laboratory experiments, simulations, field experiments, case studies, and action research, and each has many variants. The criterion to be used in evaluating papers covering evaluation research or validation research is not whether a known research method is used. The criterion is whether the knowledge claims made by the paper are interesting, and are justified by the research method followed. See the list of criteria given before.

Another criterion not to be used in evaluating research papers (evaluation research or validation research) is whether the techniques investigated are sufficiently new. That is a criterion for Solution Proposal papers, not for research papers. Conversely, for Solution Proposal papers, criteria relating to the use of sound research methods and interesting knowledge claims are not appropriate.

4 Discussion

Compared to Zave's earlier classification of RE research efforts [28], our goal is to identify paper evaluation criteria and not to define a classification of topics that belong to RE. We are concerned about the use of a correct research method by authors and of proper evaluation criteria by reviewers, and not about the RE topics that can be researched.

The role of scientific research methods has been debated by a number of researchers. Brooks [4] suggests that engineers aim to producing useful things and therefore do not have to follow scientific methods. Auyang [3] surveys the history of a few branches of engineering and observes no difference between research into the properties of artifacts and research into the properties of natural objects. Engineering researchers usually have a larger obligation than natural scientists do to motivate their research by the expected utility of their results. But this is a matter of degree, and engineering researchers are just as curiosity-driven about their subjects as natural scientists are. One important observation Auyang makes is that engineering is not simply the application of knowledge taken from the natural sciences to practical problems. Rather, engineering is the application of the *scientific method* to practical problems [3, p 134]. We should make clear here that we accept both quantitative and qualitative research methods, ranging from controlled experiments to case study and action research. The essential element of any research method is that the scientists bends over backwards to check every possible way in which the knowledge claim made by the scientist could be wrong [11]. Vincenti [22, p 229 and further] too makes clear that the use of critical, scientific methods plays a central role in the growth of engineering knowledge—not to be confused with transfer from science, which plays a secondary role.

One science from which we could learn how to propose and validate techniques is medicine. Davis and Hickey [8] describe how the model of laboratory tests, clinical tests, and pilot applications of new medicine could be used to define a hierarchy of validation levels for a new RE technique, leading up to actual implementation of the technique. Currently, validation is weak in software engineering papers. Tichy et al. [21] and Zelkowitz and Wallace [29] show that validation was absent in roughly 30–50% of the software engineering papers that require validation. A case study by Wieringa and Heerkens [24] indicates that the situation in RE may be just as bad. At the other extreme, Glass et al. [12] show that papers in information systems tend to be empirical and propose no solutions for the problems they cover. The classification scheme proposed in this paper shows that we need to do all of these things: investigate problems empirically, as information systems researchers do; propose novel designs as software engineers do; and validate these solutions, as researchers in the engineering sciences do.

Acknowledgments This paper benefited from discussions with Al Davis, Sol Greenspan, Hans Heerkens, Ann Hickey, and Pamela Zave and from comments by the anonymous reviewers.

References

1. Archer LB (1984) Systematic method for designers. In: Cross N (eds) Developments in design methodology. Wiley, Chichester, pp 57–82

2. Asimov M (1962) Introduction to design. Prentice-Hall, Englewood Cliffs
3. Auyang S (2004) Engineering: an endless frontier. Harvard University Press, Cambridge
4. Brooks F (1996) The computer scientist as toolsmith II. Commun ACM 39(3):61–68
5. Bush D (2005) Modelling support for early identification of safety requirements: a preliminary investigation. In: Proceedings RHAS'2005 workshop, 13th international IEEE conference on requirements engineering, Paris, August 2005
6. Cross N (1994) Engineering design methods: strategies for product design, 2nd edn. Wiley, Chichester
7. Cross N (2001) Design cognition: results from protocol and other empirical studies of design activity. In: Eastman C, McCracken M, Newstetter W (eds) Design knowing and learning: cognition in design education. Elsevier, Amsterdam, pp 79–103
8. Davis AM, Hickey AM (2004) A new paradigm for planning and evaluating requirements engineering research. In: Second international workshop on comparative evaluation in requirements engineering, Kyoto, Japan, pp 7–16
9. Easterbrook S, Yu E, Aranda J, Fan Y, Horkoff J, Leica M, Qadir RA (2005) Do viewpoints lead to better conceptual models? An exploratory case study. In: Proceedings of 13th IEEE international conference on requirements engineering, IEEE Computer Society Press, pp 199–208
10. El Emam K, Madhavji NH (1995) A field study of requirements engineering practices in information systems development. In: Proceedings of 2nd IEEE international symposium on requirements engineering, IEEE Computer Society Press, pp 68–80
11. Feynman RP (1994) Six easy pieces. Perseus Books
12. Glass R, Ramesh V, Vessey I (2004) An analysis of research in the computing disciplines. Commun ACM 47(6):89–94
13. Hall AD (1962) A methodology for systems engineering. Van Nostrand, Princeton
14. Hicks MJ (1991) Problem solving in business and management. Hard, soft and creative approaches. International Thomson Business Press
15. IEEE (2004) Genre submission guide. http://www.computer.org/software/genres.htm. Cited 22 May 2005
16. Jones JC (1984) A method for systematic design. In: Cross N (ed), Developments in design methodology. Wiley, pp 9–31
17. Maiden NAM, Jones SV, Manning S, Greenwood J, Renou L (2004) Model-driven requirements engineering: synchronising models in an air traffic management case study. In: Proceedings CaiSE'2004, Springer-Verlag LNCS 3084, pp 368–383
18. March JG (1994) A primer on decision making. How decisions happen. The Free Press
19. Merriam-Webster. http://www.m-w.com/
20. Roozenburg NFM, Eekels J (1995) Product design: fundamentals and methods. Wiley
21. Tichy W, Lukowicz P, Prechelt L, Heinz E (1997) Experimental evaluation in computer science: a quantitative study. J Syst Softw 28:9–18
22. Vincenti WG (1990) What engineers know and how they know it. Johns Hopkins University Press, Baltimore
23. Wieringa RJ (1996) Requirements engineering: frameworks for understanding. Wiley, New York
24. Wieringa RJ, Heerkens H (2004) Evaluating the structure of research papers: a case study. In: Gervasi V, Zowghi D, Sim SE (eds) Second international workshop in comparative evaluation of requirements engineering, pp 29–38
25. Witte E (1972) Field research on complex decision-making processes—the phase theorem. Int Stud Manage Organ 2:156–182
26. WordNet. http://wordnet.princeton.edu/
27. Yu E, Mylopoulos JM (1994) Understanding "why" in software process modelling, analysis and design. In: Proceedings of 16th international conference on software engineering, IEEE Computer Society Press, pp 159–168
28. Zave P (1997) Classification of research efforts in requirements engineering. ACM Comput Surv 29(4):315–321
29. Zelkowitz M, Wallace D (1997) Experimental validation in software engineering. Inf Softw Technol 39:735–743