Evaluating the Structure of Research Papers: A Case Study

Roel Wieringa[1] and Hans Heerkens[2]

[1] Faculty of Electrical Engineering, Mathematics, and Computer Science, P.O. Box 217, 7500AE Enschede, The Netherlands,
roelw@cs.utwente.nl,
http://www.cs.utwente.nl~roelw

[2] Faculty of Business, Public Administration and Technology, P.O. Box 217, 7500AE Enschede, The Netherlands,
j.m.g.heerkens@utwente.nl

Abstract. This paper is triggered by a concern for the methodological soundness of research papers in RE. We propose a number of criteria for methodological soundness, and apply these to a random sample of 37 submissions to the RE'03 conference. From this application, we draw a number of conclusions that we claim are valid for a larger sample than just these 37 submissions. Our major observation is that most submissions in our sample are solution-oriented: they present a solution and illustrate it with a problem, rather than search for a solution to a given problem class; and most papers do not analyze why and when a solution works or does not work. We end with discussion of the need to improve the methodological soundness of research papers in RE.

1 Introduction

This paper is triggered by a concern for the methodological soundness of submissions to RE conferences, and also by an attempt to improve the methodological structure of our own papers. Although ten years of RE research has produced many results, uptake by industry is slower than can be expected by the speed at which research results are produced. Also, the RE knowledge base has a peculiar structure that resembles a garden with a thousand flowers, with each gardener tending his or her own plot, rather than it resembles a single building where workers build upon each others results. There are many different schools of thought, each revolving around one technique, proposing additions, variations and improvements. Occassionally a completely novel technique is proposed, leading to a new stream of papers. Why has this fecundity not lead to a massive improvement of RE practice? Our hypothesis is that this is because the research methodology by which these results have been produced, is not sound. If we would use a sound research methodology, we would build upon each other's results, validating what has been proposed before, relating our solutions to real problem classes, and analyzing problems that have not yet been solved. Our aim in this paper is not to validate this hypothesis but to answer a more modest question, namely

– What is the methodological structure of submissions to RE conferences?

To answer this question, we set out a checklist of criteria for methodological soundness (section 2), applied these to 37 randomly chosen submissions to the RE'03 conference (section 3), tabulated our observations (section 4) and analyzed them (section 5). This gave us more insight in the nature of the problem, or at least of what we consider to be a problem, and indicated some lessons learned by which to improve our research methodology (section 6).

2 Criteria for Methodological Soundness

2.1 Knowledge problems and action problems

To set up a list of criteria for methodological soundness, we start from the observation that each research paper proposes a solution to a problem. We distinguish two kinds of problems:

– **Action problems** consist of a difference between the way we perceive the world to be and the way we think it should be. We normally solve an action problem by changing the state of the world. As a side effect, this produces knowledge, and in special cases, this knowledge may even be sufficient to make the problem go away. By trying to implement a change we may learn that the world is quite different from what we thought it to be, and that it in fact already agrees with our desires. However, the general approach to solving an action problem is to change the world, not to change our perception of the world.
– **Knowledge problems** consist of a lack of knowledge about the world. To solve a knowledge problem, we need to change the state of our knowledge, and when we do that, we try not to change the world. As a by-product of gathering knowledge we usually do change the world, e.g. by doing an experiment or by observing people (who thereby may change their behavior). However, the intention still is to change our knowledge state without altering the state of the world.

Given this distinction, we can distinguish papers presenting a solution to an action problem and papers presenting a solution to a knowledge problem.

– An example of a knowledge problem in RE is the lack of knowledge of the structure of RE negotiations. A paper contributing to a solution of this problem would study RE negotiations and report about the structures found.
– Another example of a knowledge problem is the lack of in-depth knowledge of RE processes in health care. A paper treating this problem could present the results of a case study of an RE process for a health care application.
– An example of an action problem in RE is the need for better integration of formal and informal techniques for RE. A paper treating this problem could propose an integrated technique and explicitly show that it solves the problem to some extent. If its contribution to knowledge has been explicitly related to the scientific body of knowledge, this paper would contribute to our knowledge of possible solutions to this action problem.

– Another example of an RE action problem is the need for repeatability in RE processes over different project groups. A paper analysing this problem could investigate a large number of RE processes and search for causes of variability. Such a paper would contribute to our knowledge of the action problem, and so would solve a knowledge problem to some extent. Another kind of paper could take this analysis as a starting point, propose a solution to the action problem and show that this improves the situation to some extent. This paper would contribute to our knowledge about solutions to this action problem.

We now distinguish papers into those that treat an action problem and those that treat knowledge problems.

2.2 Criteria for the presentation of a solution to an action problem

When a paper presents a solution to an action problem, it should answer the following questions:

– Which action problem is solved?
– What is the relevance of this solution?
– How was the problem solved?
– Is the solution valid?

Appendix A gives the complete checklist based on these questions. Here we discuss each of the criteria in turn. First, *Which action problem is solved?* Action problems consist of problematic phenomena in the world, and so they can be structured by identifying the relevant *phenomena* and the *norms* with respect to which these phenomena are problematic. Norms may have to be justified, e.g. whose norms are they, or are they part of a wider system of norms? In addition, the *relationship* between norms and phenomena may have to be explained, e.g. to which phenomena do the norms apply? Where there are norms, there are *stakeholders* and these may have to be named explicitly. Finally, every action problem is usually part of a bundle of problems. Are these problems distinguished? Are are *priorities* given to the problems in the bundle? Did the author make a choice which of the problems in the bundle to discuss in the paper?

Second, once it is clear which action problem is solved, the *relevance* of the proposed solution for theory and practice should be indicated. Does the solution contain a significant addition to our knowledge base? Does it contribute to an advance in practice?

Third, *How was the problem solved?* To solve a problem, it must be *analyzed* first. Note that the analysis of a problem bundle is a knowledge problem itself that may deserve a paper in itself. To analyze a problem in the real world, one must identify (classes of) phenomena and find *causal relationships* between them. Based on a problem analysis, *criteria* for a solution must be identified. Next, solutions should be generated. This may be by stroke of genius, or alternatively the author may have had a systematic way of finding solutions. In any case, is the *method* for generating solutions indicated? Usually, there is more than one

solution. Were alternatives considered and is there an argument for the chosen solution in terms of solution *properties* and an *evaluation* of the solution against the criteria? Note that this too is a knowledge problem that may deserve a paper in its own right.

Solutions to action problems may be implemented. Is the implementation *described* and is it *evaluated?* Implementation evaluation may involve setting up operationalized *criteria* for success, *justifying* those criteria, performing *observations* and *relating* the observations to the original action problem. Again, this is a knowledge problem that may deserve a separate paper to report about.

Finally, *Is the solution valid?* Why should the reader believe that the solution worked for this particular problem? And if it worked for this particular problem, for which other problems is the solution claimed to work as well? These questions concern *internal* and *external validity,* respectively.

Discussion. We consider these criteria to evaluate the structure of papers that treat solutions to action problems self-evident. As far as we are concerned, they follow from the definition and analysis of action problems. They are based on what we think is a consensus in the literature [1–6]. However, what we consider self-evident and consensual may be false or misguided. After we discuss our application of these criteria to 37 RE submissions, we will discuss the appropriateness of our criteria for papers about action problems.

At this point we must make clear that we do not mean that each paper about action problems should satisfy all these criteria. The criteria can be used to position and analyze a paper to see what contribution is proposes. However, if the paper falls short on *all* of these criteria, we submit that it is not about an action problem. In that case it should be made clear what the paper *is* about.

A second remark is that each action problem contains three knowledge problems: Problem analysis, solution analysis, and implementation analysis. We will return to this point after we discussed the structure of papers solving a knowledge problem.

2.3 Criteria for the presentation of a solution to a knowledge problem

We only consider empirical knowledge problems, not mathematical knowledge problems. So our knowledge problems are about phenomena in the real world, not about logical relationships or mathematical constructs. We claim that all knowledge problems have the structure proposed below. Again, this is based on our analysis of the structure of knowledge problems and our reading of the literature [7–11].

Analogous to action problems, we can ask the following questions about a paper presenting a solution to a knowledge problem.

- Which knowledge problem is solved?
- What is the relevance of this solution?
- How was the problem solved?

– Is the solution valid?

The corresponding checklist is given in appendix B. First, *Which knowledge problem is solved?* Empirical knowledge problems are identified by the *phenomena* of interest and the *properties* of those phenomena that we are interested in. (Sometimes, these properties are called *variables* but we will not do that here.) The knowledge problem may consist of finding causal relationships among the phenomena. It may also consist of simply finding the value of the properties of interest in the observed situation. If there are several knowledge problems, are they distinguished and *prioritized*? Is it clear which one the paper is proposing to solve?

Second, just as for action problems, the next question is *What is the relevance of this solution?* Again, the question is what the contribution to theory and to practive is.

Third, *How was the problem solved?* This involves an explicit list of one or more *research questions* to be asked, and a *design* of the research by which the answer to these questions will be found. Note incidentally how a knowledge problem contains an action problem (research design). This makes the relationship between action problems and knowledge problems recursive, for we saw above that each action problem contains knowledge problems.

A description of research design involves a description of the *units of observation* and a description of the way *data* is collected. In a fully-blown research design, we also need to know how the properties to be observed are operationalized, i.e. what the relationship is between what is actually observed and these properties. This may involve introducing indicators (properties that can be observed) that have a causal relationship with the properties of interest.

Having designed the research, we perform it, i.e. we perform *observations.* Here we can ask according to which *procedure* the observations were done and what *happened* when this procedure was followed. Observations made must be *analyzed*. Are there *theoretical explanations* of the observations? Are *fallacies* in the interpretation of the data avoided?

Finally, *Is the solution valid?* Just as for action problems, we can ask for *internal validity* and *external validity*. Internal validity is here the soundness of the research design, i.e. the research does indeed provide information about the properties of interest. External validity is the validity of the answers to the research questions for a larger population than the sample studied.

Discussion. We claim that the proposed structure of reports about solutions to knowledge problems is applicable to all empirical knowledge problems. In particular, it is applicable to the three knowledge problems that appear in action problems: problem analysis, solution analysis and implementation analysis. This checklist can thus be viewed as a refinement of the appropriate parts of the checklist for action problems. If a paper discusses a solution to a knowledge problem in the context of an action problem, both checklists should be used.

3 Case study

We analyzed 37 randomly chosen submissions the RE'03 conference by scoring them on our checklists. Seven of these were accepted papers, so the acceptance rate in our sample is 19%, which is identical to the acceptance rate of the conference. One of us scored all papers on the checklists and the scores were verified by the other.[3] We believe this gives a reasonable internal validity of our scores. That is, we believe that even if a third person who would apply the checklists on these 37 papers might occassionaly give a different score on one attribute to a paper, our results below would not change significantly.

We cannot make a similar claim for external validity: Our sample is too small to generalize over all submissions to all RE conferences. However, the reader may nevertheless gather knowledge from our results if he or she recognizes these patterns in RE papers he or she is familiar with.

4 Observations

Problem definition. About one third of the papers we reviewed are about knowledge problems. Three of these discussed both an action problem and a knowledge problem. Our sample thus contains a bias towards action problem.

Action problems.

- **Action problem definition.** Almost all papers about action problems make clear what the relevant phenomena and relevant norms are, and who the stakeholders are. However, the norms used in the paper are usually not motivated. About one 10 papers do give a motivation but in 7 of these the motivation is so weak that we are not sure it is present.
  Almost all papers discuss action problems in isolation: Only three papers consider problem bundles and give priorities to the problems in the bundle. Related to this, most papers about action problems contained no problem analysis, or one so weak that we are not sure it is present.
- **Action problem solution generation.** Only three papers in our sample indicated their method of solution generation. The others did not do so, or did it so very weakly and we are not sure a method is described. There is little attention in giving the criteria for evaluating proposed solutions: Two papers gacve these criteria, the others did not, or did it so weakly that we are not sure criteria were given.
- **Solution choice.** About half (18) papers gave an argument for the selected solution, but half of these (10) did so very weakly. Alternatives are usually not considered. Only three papers gave clear attention to alternative solutions and gave an argument for the selected option.
- **Implementation.** There is little attention to implementations in our sample. Only four papers described an implementation of the solution.

[3] This biases the results because the verifier has not formed an independent judgement.

Knowledge problems.

- **Knowledge problem definition.** There is hardly any attention to this in our sample: Only one paper listed research questions.
- **Research design.** Most knowledge papers did describe this. Eight papers in our sample, i.e. most of the papers describing a knowledge problem, described a research design. Those papers also described observations done.
  There was less attention to operationalizations of the properties to be observed. Only two papers in our sample spend attention to this.
  Only one paper analyzed the results of observation, albeit very weakly. None of the papers provided explicit explanations of the observations in terms of a theory.

Validity. This is not a point of attention for the papers in our sample. Only four papers spend attention on the validity of the results. Four other papers discussed validity so weakly that we are not sure they did so, and the others did not discuss validity.

Relevance. Most papers claim relevance: About 20 papers claim a contribution to practice. However, half of these do this in a very weak way, i.e. we are not sure they really do make such a claim. Only 5 papers make a claim of contributing something to the state of knowledge. Three of these actually set out to solve a knowledge problem. So most papers that are about knowledge problems do not make clear what their contribution to knowledge is.

5 Analysis

The chosen sample of papers has a slant towards action problems above knowledge problems. Action problems are usually considered in isolation, and the relevant norms that make phenomena problematic are not motivated. Problem anlysis is most papers about action problems is weak or absent.

Solutions are usually simply proposed and criteria by which solutions are to be evaluated are usually not given. Often, no alternatives are given and the only solution considered is the solution argued for. One explanation of these observations is that the author already knew the solution to be described, and added a single problem that illustrates this solution and also gives an occassion to give arguments favoring this solution.

The emphasis of most papers is on presentation of solutions to action problems rather than considering different kinds of solutions for a problem class. There is no effort at evaluating or generalizing solutions (no external validity).

The papers in the sample are slanted towards solution description, not towards implementation description and analysis.

The papers describing knowledge problems do usually not list their research questions. Research designs are usually described, but observation results are not usually analyzed nor explained. If the knowledge problem is investigated in the context of an action problem (i.e. it is a problem analysis, a solution analysis,

or an implementation analysis), then paper do not usually ask why something works, where and when it works,and when it does not work; and it is not usually asked whether it works better than what we already have.

6 Discussion and Conclusions

The picture of methodological soundness emerging from this is pretty bleak. Let us assume that our observations and analysis is internally valid for this sample. As we already indicated, the sample size is too small to conclude anything about all RE submissions, nor even about all RE'03 submissions. However, the reader may recognize phenomena that he or she observed in papers read by him- or herself.

Applying our own checklist to our own paper, we set out a problem context, asked a research question, set out the framework by which we answered the question, described the case study, tabulated the obnservations, analyzed them and discussed the validity of the results. But how can we explain the results?

One response to the results is that there is nothing to be explained, because we do not know whether these results are valid for a significantly large population of RE papers. This may be so, but here we will proceed on the assumption that the results are valid for a significantly large population of RE papers.

Proceeding on this assumption, a second response is that researchers in requirements engineering are more interested in proposing new solutions than in analyzing problems, investigating known solutions, or investigating existing implementations. This in turn may be caused by the desire of researchers to create a paradigm shift, or a revolutionary new technique and establish a name this way. However, these explanations are very speculative and it will be very hard to verify them empirically.

A third possible response is to forget about possible explanations and look at our original action problem: How can we improve the methodological structure of RE papers? This takes us back to our checklists. Can we use these as guidelines for researchers to design their research? We think so, and our argument is that these checklists are based on accepted theories in research methodology [8, 12, 9, 13, 14, 3, 4, 15, 16].

However, a fourth possible response is that these checklists are not applicable to the kind of research done in RE. For example, it may be said that our checklist for knowledge problems are oriented to standard empirical research, whereas an analysis of our sample shows that RE researchers do design research, oriented to proposing novel solutions. We do not agree with this argument: We do not see why doing a problem analysis, analyzing a proposed solution, or analyzing an implemented solution are different kinds of research questions than those asked in other empirical sciences.

It may also be said that our checklist for action problems is not applicable to design research. Again, we do not agree with this: Proposing a novel design solution without problem analysis or solution evaluation is like proposing a

physical experiment without listing a research question to be answered by the experiment, nor performing measurements with the experiment.

The safest conclusion is that there is a need for a debate about research methodology in RE that shows either that our results are not valid for most RE papers, or shows that there is a need for better education in research methodology in RE.

Acknowledgements. This paper benefited from discussions with Wiebe Hordijk, Dulce Pumareja, and Zlatko Zlatev.

References

1. Checkland, P., Scholes, J.: Soft Systems Methodology in Action. Wiley (1990)
2. Evans, J.: Creative thinking in the decision and management sciences. South-Western Publishing Co. (1991)
3. Heerkens, J.: The Managerial Problem Solving Method. TSM Business School (1998)
4. Hicks, M.: Problem solving in business and management; hard, soft and creative approaches. International Thomson Business Press (1999)
5. Leeuw, A.d.: Organisaties; management, analyse, ontwerp en verandering (In Dutch). Van Gorcum (1997)
6. Visscher, K.: Design methodology in management consulting. PhD thesis, University of Twente (2001)
7. Baarda, D., Goede, M.d.: Basisboek methoden en technieken (In Dutch). Stenfert Kroese (2001)
8. Cooper, D., Schindler, P.: Business research methods, 8th edition. Irwin/McGraw-Hill (2003)
9. Easterby-Smith, M., Thorpe, R., Lowe, A.: Management Research. An Introduction. Sage Publications (1993)
10. Hart, H.t., Dijk, J.v., Goede, M.d., Janssen, W., Teunissen, J.: Onderzoeksmethoden (In Dutch). Boom (1998)
11. Swanborn, P.: Methoden van sociaal-wetenschappelijk onderzoek (In Dutch). Boom (1987)
12. Cronbach, L.: Essentials of Psychological Testing. Stanford University (1990)
13. Eisenhardt, K.: Building theories from case study research. Academy of Management Review **14** (1989) 532–550
14. Ericsson, K., Hastie, R.: Contemporary approaches to the study of thinking and problem solving. In Sternberg, R., ed.: Thinking and problem solving. Academic Press (1994)
15. Miles, M., Haberman, A.: Qualitative Data Analysis. An Expanded Sourcebook. Sage Publications (1994)
16. Strauss, A., Corbin, J.: Basics of Qualitative Research. Grounded Theory, Procedures and Techniques. Sage Publications (1998)

A Checklist for Papers about Solutions to Action Problems

1. Which action problem is solved?

   (a) Phenomena
   (b) Norms
   (c) Relationship between norms and reality
   (d) Stakeholders
   (e) Priorities
2. What is the relevance of this solution?
   (a) For theory
   (b) For practice
3. How was the problem solved?
   (a) Problem analysis
      i. Causal relationships between phenomena
   (b) Criteria for solutions
   (c) Method for solution generation
   (d) Choice of a solution
      i. Alternatives
      ii. Solution properties
      iii. Solution evaluation
   (e) Implementation description
   (f) Implementation evaluation
      i. Criteria
      ii. Justification of criteria
      iii. Observations
      iv. Relating observations to action problem
4. Is the solution valid?
   (a) Internal validity
   (b) External validity

B Checklist for Papers about Solutions to Knowledge Problems

1. Which knowledge problem is solved?
   (a) Phenomena
   (b) Properties
   (c) Causal relationships to be investigated
   (d) Priorities
2. What is the relevance of this solution?
   (a) For theory
   (b) For practice
3. How was the problem solved?
   (a) Research questions
   (b) Research design
      i. Units of observation
      ii. Data collection method
      iii. Operationalization
   (c) Observation
      i. Observation procedure
      ii. Description of what happened
   (d) Analysis
      i. Theoretical explanations
      ii. Fallacies avoided?
4. Is the solution valid?
   (a) Internal validity
   (b) External validity