Pertti Järvinen

On a structure of a research paper

DEPARTMENT OF COMPUTER SCIENCES
UNIVERSITY OF TAMPERE

D-2010-9

TAMPERE 2010
Pertti Järvinen

On a structure of a research paper

Presented in the IRIS33 Conference in Aalborg, Denmark, August 22-24, 2010
On a structure of a research paper

Pertti Järvinen

University of Tampere, Finland

pj@cs.uta.fi

Abstract. Some scientific articles are easy to read, some are not, and it is not based on content. Which kinds of reasons could be behind of those experiences? How could we improve our articles and their writing in order to alleviate problems in reading and understanding? We shall analyze those problems by taking examples from the literature. All the articles have such parts as abstract, introduction, discussion and initial parts of sections. We pay our attention to those parts and how they can be written.

Introduction

Both design-science and other “research must be presented effectively both to technology-oriented as well as management-oriented audiences” (Hevner et al. 2004, Principle 7). The former audience concerns researchers and the latter practitioners. We do not consider the gap between those two interested parties but
we emphasize their both importance as readers of our outcomes, research reports. The importance of writing does not depend on which research method has used nor the type of results (confirmation or support to a certain theory, falsification of the theory, construction of a new information system etc.).

In his editor’s comments Detmar Straub (2009), editor-in-chief of MIS Quarterly, deliberates upon question: “Why top journals accept your paper?” He gives 10 reasons in the priority order. Reasons 1 through 4 (1. exciting ideas, 2. nontrivial research questions, 3. popular themes and 4. theory) are required elements whereas items 5 through 10 are enhancing elements. By ‘enhancing’, Straub means that if they are present, then they heighten the chances of acceptance. He does believe, however, that there is latitude in the extent to which the latter six reasons factor into the decision to accept. Straub continues: “Good writing, reason 7, makes a paper a joy to read, but when the paper is ungrammatical, filled with sloppy errors, badly organized, and awkwardly phrased, it only serves to put readers off.” (Straub 2009, p. vii) We understood Straub’s thinking in such a way that a submitted paper will be rejected if it does not contain exciting ideas, nontrivial research questions, popular themes and theory. But if it contains those four reasons, it can still be rejected, if the submitted paper is ‘ungrammatical, filled with sloppy errors, badly organized, and awkwardly phrased’ Hence, we are interested in how a scientific paper should be structured that it could communicate. We believe that all the researchers are willing to transmit their results to the audience as completely as possible. Hence, we all have a high motivation to write our reports well.

Some parts of our research reports are content-dependent and we cannot give advice to every problematic detail. Hence, we try to concentrate those parts of an article that exist in every paper, i.e. title, abstract, introduction, initial parts of sections and discussion. Our consideration is slightly normative, i.e. we shall give advices how a researcher should write her paper.

Some common components of a paper

In this section we shall pay attention to those common parts mentioned above. Our sincerest wish is that our advice will help a researcher to write her article in such way that a reader could understand it by reading through only once. The core idea of our advice is to lessen a reader’s mental load and hence improve communication and understanding.
The title of the paper describes the report in the shortest way. The title should clearly and succinctly describe the subject of the research. Artificial but uninformative titles are considered inappropriate, but the occasional pun is tolerable.

The abstract is supposed to communicate the results in the paper to the reader who is too busy (or not interested enough) to read the whole paper, and also serve to tempt the curious reader into reading further. It should be short, including a sentence or two of motivation, a sentence or two of problem definition, the selected method and a short description of the major results. Technical details should be kept to a minimum. Since many review publications excerpt only the title and abstract of a paper, the abstract must be completely self-contained, accessible to the non-expert, and must remain appealing when read in isolation.

Introduction can, for example, be divided into five (or six) subsections:
1. Description of problem domain and its importance
2. A gap or conflict in the results and findings achieved this far
3. Exact definition of the problem under study
4. Presentation of (own) approach and its advantages
5. Results
6. Structuring the rest of the paper into sections

Referring to a basic book or the newest review article on the topic you can describe the position of your paper. Some recent statistics can demonstrate the importance of the topic. The transfer from the wider topic to the detailed problem domain directs and focuses attention of readers to the problem under study. Some basic concepts or constructs may also be defined or clarified. Presentation of (often practical) importance of the topic and problem domain will motivate the reader to continue her reading. If the study produces something new and interesting both in theory and in practice, it is then very valuable. The researcher must not be for modesty, because she has used much her energy and time for the study.

In addition to the practical motivation above you must also present motivation from the scientific point of view. There are at least three reasons to publish your results: 1) your results are novel, 2) you succeeded to solve the conflict between earlier results or 3) you are first applying more dense differentiation on your problem domain and hence your results are new.

Reason 1: Your results are novel.
Straub (2009) stresses that the intellectual territory of our study is unexploited, e.g., a new perspective on an old problem. The new idea must not be too radical
and lie too far beyond the reviewers’/evaluators’ experience. Straub also wishes that our research question is neither addressed nor answered satisfactorily before. According to Barley (2006) the papers he considered as interesting differed in some significant and striking way from most of the other papers in academic journals. Some articles interest him simply because they address subjects that depart noticeably from the mainstream. He also recommends that we should bring together ideas or objects from previously unrelated domains.

Our example for a good scientific motivation is as follows: In their introduction Iriberri and Leroy (2009) present that the online community development processes are not yet much studied. To this end they shall do it. “Starting with the Well, the pioneering online community established in 1985 (Rheingold 1993), hundreds of new online communities and social networking sites have emerged (Reid and Gray 2007). Many of those continue to exist and thrive today and show dramatic membership growth. Others draw little participation from their members and some have disappeared completely. In response, researchers from various disciplines are searching for the conditions that make online communities more or less successful. The result of this effort is an extensive body of literature that proposes guidelines and success factors derived from the different perspectives of sociology, psychology, management and economy concepts and theories, a limited number of empirical studies, and a variety of anecdotal stories. Kim (2000) suggested nine strategies for building successful online communities based on her practitioner experience. For example, giving the community a purpose, encouraging etiquette, and integrating rituals increase the chances of success. Preece (2000) articulated participatory design, sociability, and usability concepts and recommended applying these concepts in building communities. From a social psychology perspective, Koh et al. (2007) stressed the need to motivate participation, while Leimester and Krcmar (2004) concluded from their case studies that protecting the privacy of participants is essential. More recently, researchers have started empirically testing independent conditions that can indicate the success of these communities. However, little effort (cursive added by this author) has been made to document the online community development processes and provide guidelines to introduce success factors and design choices in an integrated and orderly way.” (Iriberri and Leroy 2009, p.2)

Reason 2: You succeeded to solve the conflict between earlier results.

On the one hand, Nonaka (1994) claims that “the assumption that knowledge is created through conversion between tacit and explicit knowledge allows us to postulate four different "modes" of knowledge conversion: (1) from tacit knowledge to tacit knowledge, (2) from explicit knowledge to explicit knowledge, (3) from tacit knowledge to explicit knowledge, and (4) from explicit knowledge to tacit knowledge”. (Nonaka 1994, p. 18) On the other hand, Cook and Brown
(1999) claim that organizations are better understood if explicit, tacit, individual and group knowledge are treated as four distinct and coequal forms of knowledge (each doing work the others cannot). Orlikowski (2002) identified two distinct perspectives on organizational knowledge. One proposes that organizations have different types (e.g. tacit and explicit) of knowledge, and identifying and examining these will lead to more effective means to generating, sharing, and managing knowledge in organizations. Another perspective argues that tacit knowledge is the necessary component of all knowledge. Orlikowski solves the conflict above by adopting such a perspective that “tacit knowledge is a form of ‘knowing’, and thus is inseparable from action because it is constituted through such action” (p. 251).

Reason 3: You are first applying more dense differentiation on your problem domain and hence your results are new.

Nielsen P.A. and G. Tjørnehøj (2009) motivate a reader as follows: “There are reports of successful SPI (Software Process Improvement) where CMM (Capability Maturity Model) or other maturity models were not used. Kautz has studied process improvement in three small companies (Kautz 2000, 1999). The success of SPI in these companies is attributed to four factors (Kautz 2000): a tailored approach; an experience network between companies; external assistance; and partial external funding. In another study a medium-sized company’s problems with current software processes were assessed with a technique for problem diagnosis, which was not on the basis of a maturity model and many of the identified problems could later be alleviated (Iversen et al. 1999; Nielsen et al. 2002). In this study the success of the improvement effort was attributed to the particular way experience and knowledge were shared during the problem diagnosis.

Sharing knowledge, also sometimes referred to as sharing experience, is fundamental in all these reports. Hence, we have undertaken research to understand knowledge sharing better and in greater detail (cursive added by this author). We report on an action research effort in the software company SmallSoft on how knowledge sharing can be understood through social network analysis and how software managers can utilize social network analysis to manage SPI efforts.” (p. 1-2)

The replication study is lacking from the list of the three reasons for publishing above. The reason for that is that there are two differing viewpoints. According to Berthon et al. (2002), replications are an important component of scientific method in that they convert tentative belief to accepted knowledge. Colquitt and Zapata-Phelan (2007, p. 1303) wrote that “replications of previously published work and very incremental research rarely offer enough of a contribution to warrant publication”. Although I support the Colquitt and Zapata-Phelan’s view I
can sometimes also see a need for replication studies. My view is based on the self-steering conception on human being and people (cf. Aulin (1989, p. 173). This view assumes that the same state of a human being never returns, i.e. we people always move into a new state. This means that old scientific results might deteriorate, and this motivates me to accept repeating the similar studies as earlier.

When you are presenting gap or conflict in the results and findings achieved this far, Webster and Watson (2002) recommend that you should not criticize too much, because previous work is always vulnerable. It is more important to explain how research builds upon previous findings rather than to claim previous research is inadequate and incompetent. After locating your own study you can present how your study will produce new knowledge about problems not yet studied at all, or how it supplements or complements earlier studies. You can here motivate a reader to read your paper. This motivation often concerns the scientific merits of the study reported.

The problem definition immediately follows from the scientific motivation, i.e., the link between the lack of knowledge and your problem should be as clear as possible. We are performing a scientific study, not practical surveying nor development.

In the most studies some known research approach or research strategy is used. A researcher, however, sometimes develops his/her own approach, i.e. a new scale, a new idea to apply the known theory, method, algorithm or model, a new technique to gather raw data, a new measurement instrument, a new description technique or language, etc. The result may then slightly improve or modify the results achieved earlier. Some ideas about those developments should be already presented in the introduction.

Lee and Hubona (2009) use logic as the description language when they present their view on a scientific rigor as follows: “We are raising three methodological issues in this research essay: a common scientific basis, the fallacy of affirming the consequent, and summative validity. These issues can be controversial and call into question how information systems research is now practiced. To support our argument, we will use a framework that we build from some elementary aspects of formal logic. We call it the MPMT framework, where MPMT refers to a specific way of using modus ponens and modus tollens.

The primary contribution of this essay is to demonstrate that the MPMT framework provides a scientific basis for the rigor of research, where the bulk of our examination focuses on rigor in positivist research and interpretive research. A corollary to this examination will be that the MPMT framework can also provide a scientific basis for the rigor of research which focuses on relevance, such as action
research and design research. (We use the term *design research* as a synonym for *design science.*)” (Lee and Hubona 2009, p. 238)

*Main expected results* can be presented in a form that "in this report we shall show that .... ", in other words, which conclusions will be drawn at the end of the study.

A short description of *structuring the rest of the paper into sections* will give the first view to a reader. It will orient her to the paper. You can tell the thread through the whole report and give some rationale for division of the paper into sections.

**Some general advice on the content**

Straub (2009) says that he give al least one recommendable structure for a scientific paper, namely, “a positivist, quantitative paper to assume the structure of (1) introduction/ motivation/ research questions, (2) literature review, model and hypothesis development, (3) methodology, including instrument validation, (4) data analysis, and (5) discussion, including implications for scholars and managers, limitations, and future research directions. If an author ignores this structure and skimps on the methodology section, as a case in point, the reviewers will be vaguely unsettled and they will then give the authors a long list of methodological issues to deal with in the revision, if they are disposed to even invite a revision. Just having all of these sections at sufficient length does not in any way guarantee acceptance, naturally.” (ibid. p. vii)

Straub (2009) also emphasize a role of theory saying that “theory is King and it is in the evaluation of the theoretical contribution that most reviewers become convinced, or not. What may not be clear is how this attitude relates to blue oceans in that authors very seldom are developing brand new theory; they are most often applying a theory and then showing how variations or refinements of this theory can be applied to a new domain.” (ibid. p. vi) Also Barley (2006) pays attention to theory. He says that some papers have interested me because they propose theories, or at least perspectives, that differ from what has gone before.

Barley (2006) gives rather general advice both for empirical and theoretical papers: “Empirical papers need to flow from introduction to problem statement to methods to data and then to a discussion and conclusions. Theoretical papers need to work though implications of propositions and consider counterarguments. In both cases, readers expect authors to warrant their claims in ways that scholars find legitimate: with logic, mathematical models, data, and counterfactuals, for
example. Without such warrants, a paper too closely resembles opinion, and when it seems to be mere opinion, a paper is unlikely to survive academic skepticism long enough to have a chance to be considered interesting.” (ibid. p. 19) In Barley’s experience, failing to conform to accepted canons of warranted claims is the qualitative researcher’s and the theoretician’s Achilles’ heel.

Rhetoric and metatextual structures

Mauranen (1993) studied rhetorical differences between texts written by academics with different cultural backgrounds. She understood rhetoric as persuasive discourse. Rhetorical strategies consist of the choices that writers make in order to convince readers of their claims. Through rhetorical choices, the writer thus aims at increasing the credibility of certain proposition in the reader’s mind.

According to Mauranen (1993) one text-linguistic variable, metatext, plays an important role in rhetorical strategies. Metatext is “essentially text about the text itself. It comprises those elements in text, which at least in their primary function go beyond the propositional content.” (p. 7) Metatext thus serves to organize the propositional content of the text and to comment on it. The metatext concept can also be seen relative; some part of a text counts as metatext only in relation to another part of it.

Through metatext, the writer steps in explicitly to make her presence felt in the text, to give guidance to the readers with respect to how the text is organized, to what functions different part of it have, and to what the author’s attitudes to the propositions are. Mauranen cited Vande Kopple’s (1985) seven subtypes of metatext (Mauranen 1993, p. 9):

1. Text connectives: *first, next, however, but*
2. Code glosses: *x means y*
3. Illocution markers: *to sum up, to give an example*
4. Narrators: *according to Einstein*
5. Validity markers: *perhaps, might, clearly, obviously*
6. Attitude markers: *surprisingly, it is fortunate that*
7. Commentaries: *you may not agree that, dear reader*

Mauranen limited herself to metatext, which primarily serves the purpose of textual organization. She focused on: connectors, reviews, previews, and action markers “with examples from the material studied:

1. **Connectors.** Conjunctions, adverbial and prepositional phrases, which indicate relationship between propositions in text: *however, for example, as a result*
2. Reviews. Clauses (sometimes abbreviated), which contain an explicit indicator that an earlier stage of the text is being repeated or summarized: So far we have assumed that the corporate tax is a proportional tax on economic income.

3. Previews. Clauses (sometimes abbreviated), which contain an explicit indicator that a later stage of the text is being anticipated: We show below that each of the initial owners will find this policy to be utility maximizing. This preview can be located at the end of the previous section or in the beginning of the section.

4. Action markers. Indicators of discourse acts performed in the text: the explanation is, to express this argument in notation, to illustrate the size of this distortion, …” (Mauranen 1993, p. 10)

Mauranen (1993) explained the functions of a preview and a review as follows: “Even when the primary function of a sentence as either a review or a preview is distinguishable, metatextual sentences are usually strongly both retrospective and prospective. Most preview sentences contain anaphoric references, and most review sentences also prospect clearly ahead. This is not surprising in view of the common interpretation of metatext as helping to orient the reader, to guide the reading process by indicating the organization and progression of the discourse. For successful orientation, it is necessary to build upon what has gone before, as well as anticipate what is to come.” (Mauranen 1993, p. 10) – Our recommendations concerning how to structure Introduction and Discussion sections are in concordance with the functions of a preview and a review, respectively. Especially the practical and scientific motivation sub sections in Introduction have their counterparts (Implications of results to science and Implications of results to practice) in Discussion. In addition to Introduction itself, its last sub section (Structuring the rest of the paper into sections) is another preview to the article proper.

We still want to emphasize that the motivation is important, not only in the introductory section but in the beginning of each section. You can add motivational substance into your preview sub-sections.

The discussion section

As Straub (2009) proposed above we had also recommended that the discussion section should be divided into four subsections:

1. Implications of results to science,
2. Implications of results to practice,
3. Limitations and
4. Further research.

The repetition of results in a concise form helps a reader to speedily recall an overview of the study. By comparing own results with the outcomes achieved by other researchers, the author shows the relational importance of her own achievements, i.e., the implications or consequences of the results to science. It is also reasonable to evaluate how well you succeeded in your problem-solving, e.g., you should describe which hypotheses were supported and which ones not. A gap or conflict in the results and findings achieved this far in the introduction and the implications of results to science in the discussion section are like the Siamese twins; the former motivates a reader to read more and the latter shows how well you succeeded in the scientific sense. Both ground on a carefully performed literature review, e.g. on the systematic one as Kitchenham et al. (2009) and Brereton et al. (2007) recommend. Based on your careful literature review you can decide which ones of your results are: a) novel, b) supporting the earlier literature, and/or c) contradicting with the earlier literature. The novel results are the most significant ones. We in connection with Introduction referred to two conflicting views on the results supporting the earlier literature. The results contradicting with the earlier literature need some comments or speculation on potential reasons behind the contradictory results.

Implications of results to practice contain clear advises how to apply the outcomes achieved. The writer can also give warnings and restrictions about potential misuse of the results. The writing recommendations for practitioners are not always so easy to do, we refer to Ramiller and Pentland (2009): “We take a fresh look at the IS academic community’s enduring concern with the management implications of its research. We examine in particular what we call the ‘variables-centered’ research paradigm, which focuses its attention on covariance among independent and dependent variables. As the predominant research tradition in the field, the variables-centered paradigm ought to constitute a major platform from which our community can speak to issues of managerial interest. Unfortunately, the variables-centered paradigm appears to distance researchers from the organizational actors, such as managers, to whom they would give advice and counsel. Particularly disturbing is the systematic erasure of those very actors from the domain of inquiry. Erased, too, are their actions and means of acting. Thus, when it comes time to offer useful prescriptions for action, our community attempts to do so on the basis of research in which, ironically, neither actors nor action directly appear.” They offer some recommendations that may help to rectify this problem and, thereby, enrich the capacity of variables-centered research to speak in an informative and useful way to issues of practice.
Limitations of the study can contain many things, e.g. a) exclusion of some intervening variables that in the course of this study showed to possibly be important, b) a convenient sample instead of a random sample, c) one or few cases where just this phenomenon existed, d) a new technical idea, which was unknown for a researcher at the beginning of the building process, emerged etc. This sub-section will show to a reader that a researcher knows most of her limitations, and the author has given economic or other reasons why those limitations mentioned are natural or acceptable. Some of the limitations can be converted to research proposals.

The further research sub-section normally contains a long list of new research questions. The researcher can in the course of her study make many observations on what should be studied in the future.

In the recent years the scientific papers seem to end with Conclusion. This can be well understood because otherwise the further research sub-section could stay as the last message from the article to a reader, and it will direct thinking missing the main results. Hence it is reasonable to finish a paper by repeating its main scientific results at the end of the article.

Summary

This article is based on my personal experience and some literature. I have organized our doctoral seminar since 1990 and we have read 3 new good articles per month. During my reading I have recognized that some articles are easier to read than some other ones. One reason is their structure and some guiding and orienting text that alleviate a load of a short-term memory. My own results mainly concern how to structure Introduction and Discussion. Some examples from the current literature are supporting my proposals.

I know that my proposals have some limitations. First, there are different learning styles; according to Rancourt (1986) the empirical, rationalistic and metaphorical learning styles, according to Kolb (1985) the convergent, assimilative, divergent and accommodative learning styles; and the text should be adjusted according to a certain learning style. My proposals are best suitable for one learning style only. My proposals are not based any scientific studies and it encourages us to perform the similar comparative studies as Mauranen did when she compared how the Finnish economists and American economists write their scientific article.

Acknowledgements: Foundation for Economic Education supported this study.
References


http://www3.interscience.wiley.com/cgi-bin/fulltext/122385301/PDFSTART


