

How To Read Scientific Papers

Professor Beste Filiz Yuksel
University of San Francisco
CS 686/486



UNIVERSITY OF
SAN FRANCISCO

CHANGE THE WORLD FROM HERE

Second half inspired by S. Keshav (2007) How To Read a Paper, *ACM SIGCOMM* 37(3) 83-84 and R.S. Larmee (2011) How To Read a Visualization Research Paper, *IEEE Computer Graphics and Applications* 31(3) 78-82.

Why is this important?

1. Short-term: how can you apply this in industry *right now*?
2. Long-term: science and technology usually takes a few decades to filter into the mainstream. Universities and institutions of research is where we build cutting edge technology to shape the long-term future.

Layout of a traditional scientific paper (excluding reviews)

Abstract

Introduction

Related Work

Method

Results

Discussion

Conclusion

References

(Sometimes Appendix)

Abstract

Summarizes paper.

Main contributions.

Does not provide details – do not be fooled by the abstract. Just a taster.

Introduction

Introduction is difficult– can be vital, can be flowery, hardest part of a paper to write.

Might not be best part of the paper to read on first pass (disagree with Keshav here), particularly in HCI, glance over it would be good. Some authors list out contributions in Introduction and summarize paper, some don't.

Related Work

To do with originality and significance.

First of all, have the authors done their homework? Do they know the work in their field? If they're missing critical work – big red flag.

Second, how original is their work. Is it a small incremental step from previous work or a significant contribution?

Science often takes small steps (not always!) but that is not always well-received by reviewers and grant foundations.

Methods

*To do with **validity**.*

Are their methods valid? Some questions to be asking (not exhaustive):

- Do they have a control condition?
- Do they have enough subjects?
- Have they controlled their independent variables?
- Have they clearly stated their dependent variables?
- Have they clearly stated their hypothesis?
- Are the dependent variables clearly evaluated?

HCI Statistics help: <http://yatani.jp/teaching/doku.php?id=hcistats:start>

Results

Results is a time to simply declare your results. No discussion here please.

Quantitative data – statistical analysis. E.g. performance or behavioral data.

p values – significant or not? – the magic number that every scientist hopes for.

Qualitative data – can't do statistical analysis but evaluation is still important.

Likert scales in questionnaires.

Interview data.

In HCI you usually get a mixture of both, we are evaluating human interaction afterall so need some subjective data.

Discussion

Advised to read on the first pass.

Significance of paper.

Contributions really clarified.

Takeaway message – how does this paper help researchers?

This is really where a paper can shine.

Authors can discuss the implications of their findings, limitations, factors to take into account. This is really where you can paint a picture of your results, where you can convince the reader of the importance of your work.

Conclusion

If you're *really* pressed for time, you can read this instead of discussion, although very dangerous to rely on this alone.

A good conclusion should summarize the main points of the paper. No more than a paragraph or two.

References

As you're reading the paper, mark any interesting papers here that you can go back and read.

I would recommend that you at least search for and read the abstract of several key papers cited by the paper to check for originality and significance.

Google Scholar is your friend, can look at papers that cite that paper as well to look at significance, how many times it's been cited, can search for keywords to look at related papers etc.:

<https://scholar.google.com/>

Appendices

Usually additional information that would be helpful in either evaluating and/or replicating the paper.

The Three-Pass Approach to Reading Scientific Papers

Instead of reading the paper all the way through and struggling through, you can use 3 passes:

1. The first pass gives you a general idea about the paper.
2. The second pass lets you grasp the paper's content, but not its details.
3. The third pass helps you understand the paper in depth.

- Keshav (2007)

However, in some HCI papers it might serve you better to read the Discussion instead of the Introduction.

First Pass

Take about 10 mins:

1. Carefully read the title, abstract, and introduction
2. Read the section and sub-section headings, but ignore everything else
3. Read the conclusions
4. Glance over the references, mentally ticking off the ones you've already read - Keshav (2007)

I would also add, the discussion might be more helpful than the introduction in some cases.

And look at the publication venue. Top tier/second tier etc.

First Pass

At the end of the first pass, you should be able to answer the five Cs:

1. *Category*: What type of paper is this? A measurement paper? An analysis of an existing system? A description of a research prototype?
2. *Context/Originality*: Which other papers is it related to? Which theoretical bases were used to analyze the problem?
3. *Correctness/Validity*: Do the assumptions appear to be valid?
4. *Contributions*: What are the paper's main contributions?
5. *Clarity*: Is the paper well written?

- Keshav, 2007

Second Pass

Should take about an hour:

In the second pass, read the paper with greater care, but ignore details such as proofs. It helps to jot down the key points, or to make comments in the margins, as you read.

1. Look carefully at the figures, diagrams and other illustrations in the paper. Pay special attention to graphs. Are the axes properly labeled? Are results shown with error bars, so that conclusions are statistically significant? Common mistakes like these will separate rushed, shoddy work from the truly excellent.
2. Remember to mark relevant unread references for further reading (this is a good way to learn more about the background of the paper).

Second Pass

After this pass, you should be able to grasp the content of the paper. You should be able to summarize the main thrust of the paper, with supporting evidence, to someone else.

Third Pass

This pass requires great attention to detail. You should identify and challenge every assumption in every statement. Moreover, you should think about how you yourself would present a particular idea. This comparison of the actual with the virtual lends a sharp insight into the proof and presentation techniques in the paper and you can very likely add this to your repertoire of tools. During this pass, you should also jot down ideas for future work.

Can take 4 or 5 hours for beginners! 1 hr for expert.

- Keshav, 2007

Highly recommended when presenting a paper in class.



Concept vs. Implementation

The difference between concept and implementation:

A concept is an idea or thought. An implementation is the actualization, or realization of a concept (or idea).

An Example:

- **Concept:** a writing utensil. A writing utensil is a concept. It is a tool that can be used to communicate with others using symbols that are drawn by a person (or animal).
- **Implementation:** a pencil. A pencil is a writing utensil, or more specifically, it is the implementation of a writing utensil. With a pencil, graphite is used to write down on a piece of paper.

Furthermore, there are many different implementations of a writing utensil. There are also pens, markers, palm pilots, etc.

Concept vs. Implementation

1. The Concept: What, conceptually, are the authors trying to achieve? What is the goal of the work? This can also be reformulated as: What is the contribution of the paper? (What's new here?)

2. The Implementation: How is the concept realized? How do the authors support their hypothesis? How do they implement the concept?

Class Exercise: comparing two papers

Contributions/significance

Originality

Validity

Clarity

Any problems?

Anything that is done well?