

Study on the Difference between Summary Peer Reviews and Abstracts of Scientific Papers

Chong Chen[†]
School of Government
Beijing Normal University
Beijing China
chenchong@bnu.edu.cn

Jingying Zhang, Xiaoyu Chu
School of Government
Beijing Normal University
Beijing China
zh-jy, chuxiaoyu@mail.bnu.edu.cn

Jinglin Zheng
School of Government
Beijing Normal University
Beijing China
738987984@qq.com

ABSTRACT

Summary peer reviews and abstracts represent opinions of reviewers and authors on the same scientific paper. Their focus and statement may be different. We propose primary measurement to compare them from readability and semantic function types. The results show that summary peer reviews highlight some distinct function types, and the terminology in peer reviews is not as dense as in abstracts. That means summary peer reviews can be complement to abstracts in literature searches, and can help readers understanding papers more thoroughly.

CCS CONCEPTS

Information systems → Information retrieval → Retrieval tasks and goals → Information extraction

KEYWORDS

Summary peer review, readability, function type, term density, functional component

1 Introduction

Scientific papers are the records of research activities. Critical information is contained in functional components of discourse, such as the research purpose, the problem definition, methods, experiments and contributions. Highlights of a study can be discovered from these components. Peer reviews are undoubtedly a quality source to reveal the highlights since reviewers are selected authorities with broad scholarly vision and strict taste. The summary reviews reflect peers' evaluation by comments on functional components of discourse. And thus the reviews are worthy of consideration in readers' paper selection. In order to identify the value of summary peer reviews from abstracts, we make comparison study on two problems, i.e., (1) the difference on functional components and readability of the two texts; and (2) the focus aspects highlighted by summary peer reviews.

2 Related Work

In order to demonstrate a scientific study, authors usually organize the content of their papers in compliance with well-established norms. For example, one common rhetorical structure of abstracts is Introduction-Method-Results-Discussion (IMRD) [1]. With the

trend of structuralizing knowledge in scientific papers, researchers try to extract information by dividing papers to semantic functional components [5]. Fine-grained functional components have also been defined. Wang et.al proposed Functional Units Ontology (FUO) based on functional unit theory [2]. The FUO includes 12 classes such as *background*, *theme*, *method*, *experiment* etc., and 28 subclasses. In fact, functional components can be obtained not only inside papers but also from their summary peer reviews. Reviewers judge innovation and contribution of a paper according to components such as the research purpose, the problems, the methods, etc.; and then summarize their positive or negative opinions in the peer reviews. Kang Dongyeop et al. created the first open dataset of review comments, PeerRead [3]. The characteristic of the accepted and the rejected papers have been compared from the vocabulary usage and psychology linguistics based on reviews [4]. The agreement between the sentimental polarity of the reviews and the acceptance has also been studied. But these studies did not compare the focus of reviews and abstracts.

3 Research Framework

An observation of this study is that, in summary peer reviews and abstracts, the focus and statement are different. The comparison is shown in Figure 1. Firstly, defining the function types of reviews and abstracts; Second, annotating sentences by types; Thirdly, analyzing how the summary peer reviews highlight a paper's value from readability and the focus aspects.

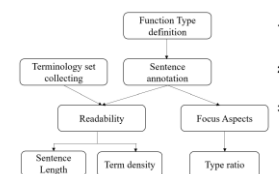


Fig.1 Research Framework

3.1 Functional Component Types

The type of functional components is called function type. Six types in scientific papers are defined in this study, i.e. *background*, *theme*, *process*, *result*, *contribution* and *strength*. The meaning of

each type is listed in Table 1. Sentences in summary peer reviews and abstracts are annotated to these types.

Table 1. The Definition of Function Types

Function Type	Meaning	Pattern examples
Background	Background of theories and applications; existing studies; unsettled gaps; necessity and significance of current study;	in order to (solve the problem) ...; ...remain unsolved; ...not (completely) studied yet;
Theme	Research scope; research goal; definition to the concerned problem;	(this study/article/paper) propose/ investigate/discuss/ demonstrate ...;
Process	Hypothesis; methods; experiments; theories and research perspectives;	Base on...proposed; First...Secondly...Last; (model/approach/method) be used/implemented...;
Result	Description and evaluation on the result, Hypothesis and methods;	(experiment/result/ simulation/) show/demonstrate...; ...(provide/give) a reference to ...;
Contribution	Contribution to the related theories or methods; Comparison with previous studies; insight obtained; future work;	The contribution (of this study/paper)...; (This study) improve...; The improvement (of this study/research) is...;
Strength	Claim the strength or highlights of the whole study and the current paper with summary description.	(idea/ concerned problem) new/novel/critical...; (experiment/data processing/research design) is firm/well;

3.2 Comparison Approach

Term density and function type ratio are defined to compare the difference between abstracts and reviews.

Term density is the average number of general terms or terminology appeared in sentences of each type of functional components. In contrast to terminology, general terms refer to widely-accepted concepts or named entities without distinct domain specificity. Term density and sentence length reflect the readability of a text since readers may feel tough to understand if the sentence is long and with lots of terms or terminology.

Type ratio counts percentage of sentences of each function type. It reflects the focused aspects by reviewers and authors.

4 Experiment

4.1 Dataset

1. Reviews and Abstracts. The dataset was collected from *Acta Psychologica Sinica* (journal.psych.ac.cn/), including a total 774 papers published from 2014 to 2019. The sentence number of summary peer reviews and abstracts is 2777 and 4397. Let symbol T , R , A , n , $avgLen$ and $T\%$ respectively denote the function types, the summary peer review dataset, the abstract dataset, the number of sentences in R and A , the average sentence length per type, and the type ratio. The details are shown in Table 2.

Table 2. Sentences and Types ratio of R and A

Function types T	# sentences n		$avgLen$		$T\%$	
	R	A	R	A	R	A
Background	124	532	39	44	4.5%	12.1%
Theme	638	504	30	37	23.0%	11.5%
Process	519	1011	30	49	18.7%	23.0%
Result	354	1724	47	68	12.7%	39.2%
Contribution	347	575	35	61	12.5%	13.1%
Strength	795	44	23	34	28.6%	1.0%

2. General Terms and Terminology. The general terms are directly taken the vocabulary of Chinese segmentation toolkit, Jieba. The terminology of psychology has been collected from three sources, i.e. the keywords of the 774 papers, the *Academic Hotspots of Psychology* in CNKI¹, and the *Chinese Terms in Psychology*². The terms from above sources were merged to a terminology set with 8,354 domain terms in total.

4.2 Results and Analysis

Readability can be reflected from sentence length and term density. As shown in Table 2, the average sentences length of R is shorter than that of A . Table 3 lists the term density of each T .

Table 3. Term Density of R and A

Function types T	General terms		Terminology	
	R	A	R	A
Background	10.4	12.1	4.4	4.9
Theme	9.0	10.3	4.6	5.0
Process	9.1	14.3	3.4	5.4
Result	13.8	19.3	5.5	7.5
Contribution	9.7	17.3	4.1	7.0
Strength	6.4	10.7	2.1	4.8

Focus Aspect is reflected by type ratio, as shown in column $T\%$ of Table 2. The percentage of type *background*, *process* and *result* in A is 12.1%, 23.0% and 39.2%, much higher than those in R , i.e. 4.5%, 18.7% and 12.7%. While on the other side, the percentage of type *theme* and *strength* is 23.0% and 28.6%, higher than 11.5% and 1.0% in abstracts. It means the reviewers emphasize how reasonable and innovative a study is. At the same time, the authors try to clearly demonstrate the research process and results.

5 Conclusion

Generally speaking, summary peer reviews focus on different aspects from abstracts, and express in an easy-to-read way. That would be helpful in readers' understanding and selecting papers. In the further study, it is necessary to find out how summary peer reviews could improve literature search.

REFERENCES

- [1] Naomi Graetz. 1985. Teaching EFL students to extract structural information from abstracts. In *J. M. Ulijn & A. K. Pugh (Eds.), Reading for professional purposes. Methods and materials in teaching language* (pp. 123–135). Amersfoot: Leuven.

¹ China Academic Journal Network Publishing Database

² <http://shuyi.cnki.net/>

- [2] Xiaoguang Wang, Menglin Li and Ningyuan Song. 2018. Design and Application of Scientific Paper Functional Units Ontology[J], *Journal of Library Science in China*, 2018,44(04):73-88.
- [3] Dongyeop Kang, Waleed Ammar, Dalvi Bhavana, van Zuylen Madeleine, Kohlmeier Sebastian, Hovy Eduard, Schwartz Roy. 2018. A Dataset of Peer Reviews (Peer Read): Collection, Insights and NLP Applications. NAACL, 2018, arXiv: <https://arxiv.org/abs/1804.09635>.
- [4] Hélène de Ribaupierre, Gilles Falquet. 2014. User-centric design and evaluation of a semantic annotation model for scientific documents. In *Proceedings of the 14th International Conference on Knowledge Technologies and Data-driven Business (i-KNOW '14)* , September 16 -19, 2014, Graz, Austria. Publisher ACM New York, NY, USA.
- [5] Wei Lu, Yong Huang, Yi Bu, Qikai Cheng. 2018. Functional structure identification of scientific documents in computer science[J]. *Scientometrics*, 115(1): 463-486.