2016

Gender differences in commenting of science blogs

HOANG NGUYEN

Master thesis in the International

Information and Knowledge Management

Master's Programme

Supervisors: Kim Holmberg and Gunilla Widén

Faculty of Social Sciences, Business and Economics

Åbo Akademi University

Åbo 2016

Abstract

Nowadays, the public receives a lot of scientific products from such open sources, some seek information and knowledge, others access for citation work or just being interested in up-to-date scientific news. This study explores scholarly communication in the context of gender differences to explore the impact between public and science. To measure the impact of science on various aspects of society, altmetrics as the modern metrics on social platforms is introduced and applied. Due to the potential investigation in social discussions about science, science blogging is chosen to examine the public impact and differences in communication between males and females. Specifically, this research investigates (1) the connection between the volume of online attention and the number of blog comments, (2) typical comments on science blog entries, and (3) gender differences in commenting behavior via sentiment analysis. The results indicated that there was no connection between altmetric attention score and number of comments on blogs; and knowledge sharing is the core motivation as well as main subject of discussions on science blogs. In addition, the findings revealed that women are more emotional in comments, especially towards female blog authors, while men tend to be neutral in discussions, but their comments are more likely negative on blog posts of female bloggers. Ultimately, the study concludes with discussions and comparison from current findings with previous research, limitations and implications suggested for supporting future research.

Acknowledgements

Allow me to begin by expressing my gratitude to one of my supervisors, Dr. Kim Holmberg. In the very first days of my master program, I already noticed his course titled "*Informetrics*" that later on inspired me to kick off this study. During the time working on this thesis, he provided me much support to keep me on the right track, opened other possibilities whenever I got stuck, checked my crappy draft diligently, and just discussed with me about career or professional matters. In overall, those lines can't express all of my gratitude to Kim, I was lucky to meet him, I was grateful to have him as my supervisor and I owe him forever for his kindness to me.

I also want to thank people from the research project "*measuring the societal impacts on open science*" at RUSE, University of Turku. They and Kim provided me valuable datasets that assisted my investigation remarkably.

If Kim inspired me to coin this study, Professor Gunilla Widén is the one who motivated me and kept me on track throughout 1 year of this thesis. To me, she is not only a supervisor, but also a mentor, a teacher and an adviser in all matter of education. The thesis seminars led by her provided me opportunities to learn as well as exchange ideas with other fellow students, to realize that I really need those mates during the study. Thanks for her flexibility in schedule arrangements that created much space for students, thanks for her patience for a lazy student like me. Without Gunilla, my studies would go nowhere and I'm also forever in her debt.

Another person who later helped me to complete this thesis is Patrick Gallagher, who checked language of my final manuscript to ensure it is flawless. I would like to express my respect as well as gratitude to this man.

As I said the thesis seminars were important and I would like to thank my mates in the seminars: you were always there to listen to my sleepy presentations patiently, I have learnt so much from you guys. An African proverb says: "*If you want to go fast, go alone. If you want to go far, go together*". Thanks to all of you once again and hopefully in the future we still have chances to walk together.

My family is obviously also a big support to me. Thanks to mom for calling me every two weeks from the other side of the Earth, to just check up on me, to just worry about my weight, to ask about my happiness and my studies and whether I satisfy my teachers or still am a trouble maker, as she used to know me in high school. Thanks to dad who worked very hard to be able to support me financially throughout my studies in Finland.

Last but not least, it would be remiss if I forgot to mention you here, my lovely girlfriend. This study is about measuring societal impact, but nothing can never measure your love, your companionship, your support, your precious time spending with me, because I already know our partnership is priceless and you are the worthiest treasure I have ever had. Thank you for these years by all of favors from the bottom of my heart.

Without all of you, this would not have been possible.

Table of Contents

List of Figures and Tables

Chapter 1. Introduction

Bertrand Russell wrote that "*Even if the open windows of science at first make us shiver after the cozy indoor warmth of traditional humanizing myths, in the end the fresh air brings vigor, and the great spaces have a splendor of their own*". Almost a hundred years ago, Russell already recognized the importance of open science. Traditionally, research results have been published in journals that are behind paywalls. This is, however, changing as open access is gaining attraction from the public, including different classes of the society. In addition, open science is a way of progressing scientific discoveries that can quickly spread new information and knowledge, provide transparency to the public and motivate social collaboration in research.

This movement towards open science has presented scientific products to broader audiences; enabling interaction in two ways between researchers and the public. Researchers publish their research in open access journals and utilize social networking sites to promote their publications. There are also some communities that are passionate about summarizing and reviewing scientific publications, for example science bloggers that can provide a bridge between scientific publications and the public.

Open science brings scholarly communication to the public. Anybody can read and recommend open access articles by sharing them on their social networking sites, and mention them in blogs so that others can comment on the blog entries. All these mentions and interactions help to build the reputation of open access research articles. These mentions could also be used to evaluate that fame and attention received. Altmetrics are the accumulative counts of these mentions on different platforms. These counts reflect the activity around research products, especially at the article-level on social platforms; how many times an article has been tweeted on Twitter, how many times it has been shared on Facebook, or mentioned on science blogs.

This study (1) contributes to investigations of the meaning of altmetrics, focusing on the impacts of open scholarly discussions in science blogs, and (2) maps gender differences in those discussions.

1.1 Overview of the study

To create a foundation for examining the meaning of the societal aspect of open science, this study begins by reviewing relevant studies. A general concept of open scholarly communication is introduced first. The following part in the literature review focuses on altmetrics, which is one of the core factors to investigate in this study. Here the current understanding of the meaning of altmetrics as well as the so called altmetric attention score are described. Next, science blogging is presented as a platform for examining open scholarly communication and public commenting of science blogs. This section discusses motivations of science blogs, how blog authors influence the public, and the level of engagement of blog readers. The last part in this theoretical background focuses on gender differences in online communication, as that is also one of the areas of interest in this study; how gender differences appear in science blogs, in communicative style, as well as emotion and language use.

This study seeks to answer the following research questions:

1. Is there a connection between the volume of online attention and blog comments?
2. What type of comments are typical for blog entries that mention scientific articles?
3. What kind of gender differences exist in the sentiments of the commenting of blog entries that mention scientific articles?

The first question aims to explore the meaning of altmetrics by investigating comments to blog entries that mention scientific articles. More specifically, correlations between the altmetric attention score, sentiment strength detected in the comments, and the quantity of the comments will be analyzed.

In order to answer the second research question a codebook is built for a deeper analysis of the blog comments. This work categorizes different types of comments on different subjects such as comments about blog authors, comments about blog posts, and interaction with others. This will provide an understanding of the type of comments that are typical for science blogs.

The last question goes deeper into gender differences in both commenters and blog authors. First the gender of the commenters and the authors will be identified where possible. Then the results from the sentiment analysis will be compared in different setups of commenters and authors.

1.2 Motivation of the study

The initial thoughts to coin this study were influenced by a course titled "*Informetrics*" at Åbo Akademi university. This course provided insights into informetrics, bibliometrics, webometrics, scientometrics, altmetrics and of quantitative aspects of information that is disseminated online. The course impressed me not only with the new knowledge, but also spectacular characteristics of using these kinds of metrics in assessing digital information and knowledge. In addition, another course called "*Information Management*" also enriched me with new knowledge about knowledge management, especially personal knowledge management on Web 2.0. This motivated me to conduct research on commenting on science blogs.

I approached a research project called "*Measuring the societal impact of open science*" at the Research Unit for the Sociology of Education at University of Turku. The project leader provided me with guidance as well as data to develop my study to explore even deeper into academia than I expected with my original plan for this thesis.

That is how this study started, but motivation for it has been increasing while working on other studies. After reading more relevant literature and digging deeper into the data, with support from my supervisors I could shape this study in the way I wanted to conduct research. Altmetrics attention score is no longer as abstract to me as it was during the informetrics course. I have also understood that science blogging is a potential platform to study and that it can be even more interesting to investigate open science than the more popular social networking sites such as Facebook or Twitter. Additionally, I have learned that understanding gender differences in terms of information sociology does not mean gender discrimination in any way. This understanding can help information users to understand differences in the online communication characteristics between men and women.

There were also some challenges during this study. Sometimes, I was lost in filtering messy data that affected the reliability of the results. My research direction was not always straight; several times I almost drifted to irrelevant topics that would have distorted the study. Nevertheless, those were valuable experiences about conducting research, and gaining new knowledge during this study has nurtured my passion to finalize it.

Chapter 2. Theoretical Background

This chapter presents relevant earlier studies. First, open scholarly communication is presented. Next, the development of metrics from traditional citation-based metrics to the new altmetrics is discussed. Third, an introduction of science blogging is given, and how the platform can potentially create or reflect societal impact of science. Lastly, differences between genders in communication and linguistic behaviors in commenting science blogs are discussed.

2.1 Open scholarly communication

Scholarly communication is the process of creating, evaluating, disseminating and connecting, as well as preserving research products systematically for future work (Klain-Gabbay & Shoham, 2016). Another modern functionality of open scholarly communication is raising awareness of scholarly products through for instance science blogging or microblogging. De Roure (2014) suggested that e-publications such as e-articles are used as social objects in this digital era for the purpose of sharing knowledge, citation, and open discussion. Moreover, those kinds of scholarly products enable information users, such as people in library and information sciences and researchers in general, to create social and research networks, to share information remotely, and more importantly, to measure attributes related to their own reputation based on public awareness. Therefore, a connection between the online metrics of usage or attention, and more traditional indicators of scientific impact have been demonstrated in earlier studies; such as citation counts association with reading counts on Mendeley, tweets, mentions on Facebook or research blog posts, even though they completely disregard the quality of the publication.

In recent years, scholarly communication has changed remarkably since innovative new ways of communication have been introduced; Wikis or Google docs for collaborative writing, instant messaging for conversation, shared workspaces to share images, and videos or other forms of documentations such as research blogs for references (Rowlands, Nicholas, Russell, Canty & Watkinson, 2011; Holmberg & Thelwall, 2014, p. 3). With advances in information technology, scholarly communication is supported in both formal ways (e.g. traditional peer-review, conferences), and informal ways (e.g. traditional library reference, personal conversations, interexchange of scholars writing). In addition, open scholarly communication on for instance scholarly social networking sites or on science blogs have contributed to a change in scholarly communication. Not being solely dependent on printed publications offers more possibilities for feedback and exchange of thoughts and ideas from both scholars and laymen across the world (Klain-Gabbay & Shoham, 2016).

Social networking sites, such as Twitter and Facebook, are also environments that can be used for academic purposes, as they promote efficient sharing and communication of scholarly works. In addition, as earlier research has shown that highly tweeted research articles tend to receive more citations later on (Eysenbach, 2011), they may have some potential for scientific impact measurement (Holmberg & Thelwall, 2014). Additionally, a previous study pointed out "*higher metric scores and higher citations for articles with positive altmetric scores*", even though with very low correlations (Thelwall, Haustein, Lariviere, & Sugimoto, 2013, p.1). Other studies have also found some correlation between citations and measures of online attention that scientific articles have received (Holmberg & Thelwall, 2014). This attention could be for instance tweeting links to academic articles or open peer review reports of that article, discussing scientific articles on blogs, mentioning or sharing scientific literature on Facebook. The accumulated amount of this attention could reflect the degree of attention or awareness the public has shown towards scientific research.

Advances in scholarly communication have been significant; from printed publication to electronic journal access, from traditional peer-review to open discussion that is rapid and global, and from private or individual communication to collaborative, developed form of communication that spread research products widely to the public that are then able to participate in the open conversation.

On the other hand, there is criticism against online metrics for research evaluation. Due to the open access to data, information and knowledge may lead to underestimation of research quality, productivity as well as other impacts or overall value (Haustein, Bowman & Costas, 2015). The consequence of this may cause ethical issues with copyright, plagiarism (e.g. transferring contents from printed publications to online resources without references), or boosting research productivity with the cost of quality.

Social web usage for communication may be very individual and it may differ among different social groups, cultures, and societies. A case study found that doctoral students and department staff at one university in Finland regularly used social web applications, but only few of the respondents used them for scholarly communication (Gu &

Widén-Wulff, 2011; Mohammadi, Thelwall & Kousha, 2015). On the other hand, a study about the researchers at an Indian university indicated that researchers utilize social platforms such as Facebook and ResearchGate for academic communication purposes (Mohammadi et al. 2015). Apart from common social networking sites like Twitter or Facebook, other social media sources for attention or usage metrics of online scholarly communication do exist. For example, social bookmarking sites or social reference managers such as CiteUlike or Mendeley can be a source of readership metrics (Mohammadi, Thelwall, Haustein & Lariviere, 2015).

2.2 From traditional metrics to altmetrics

Apart from Google Scholar, Microsoft Academic Search and Scopus (from Elsevier), Web of Science (WoS) maintained by Thomson Reuters has been the main database of scholarly literature across the world (Mas-Bleda, Thelwall, Kousha & Aguillo, 2014). WoS covers various multidisciplinary subjects from 1900 to the present with full details of authors, citations, indexes for easy dissemination, referencing, citing and searching relevant scientific documents. With the alternative sources, such as Google Scholar which provides large collection of scientific documents by automatically crawling web sites (Mas-Bleda et al. 2014), people have more ways to find scientific literature, and with that, scientific information and knowledge. In the era of Internet and information technology, scholars also inherit the benefits of rapid information dissemination. Researchers utilize the web not only for communication, but also for retrieving, publishing and assessing research products. Scientific literature is being published and shared on open repositories such as ArXiv, RePEc, SSRN, or academic social networking sites (e.g. Mendeley, Academia.edu, ResearchGate), and even common social networking sites such as Facebook, Twitter and Linkedin (Mas-Bleda et al. 2014).

Changes in the way of accessing scholarly documents can lead to changes in users' behavior as well as impact assessment of science. Mas-Bleda et al. (2014) summarized four ways of measuring impact of research from web usage: (1) conventional citations through web or web citations, (2) web usage statistics such as downloads and views, (3) web links, URL citations or mentions, and (4) tracking new impact indicators (also known as altmetrics) from social media, including tags, comments, bookmarks, conversations, readers, tweets, and blog posts.

Priem, Taraborelli, Groth, and Neylon (2010) summarized methods for assessing impact of research (Figure 1):

Figure 1. Four ways to measure impact of scholarly products (Priem et al. 2010)

Besides traditional impact assessment, such as evaluating research usage, peer-review, and conventional citation-based evaluations, altmetrics (short for alternative metrics) have been suggested as a complementary source of data about the impact academic products have made. There is, however, still some questions about whether this new evolution brings more benefits or more obstacles. The following sections including *impact through usage*, *peer-review*, *citation impact*, and finally *altmetrics – the new meta* will discuss respectively the developments in impact assessment, highlighting the strengths and weaknesses of each type of data.

2.2.1 Impact through usage

Colledge and her colleagues (2015, p.2) stated that "*viewing activity is one of the earliest indicators of interest in research*", thus counting of views may reveal something about the accumulated interest towards research. This is also a way for an alternative measurement of research excellence of scientific documents that are non-published, non-cited or non-referenced. For example, in arts and humanities, people are interested in observing aesthetic of artworks rather than citing them (Colledge et al. 2015). Researchers are not the only ones who are interested in science, as many people beyond academia seek information and knowledge everyday online. In addition, studying IP (Internet Protocol) addresses of website visitors and usage data such as views, reads or downloads, can reveal users' locations and the frequency of visits. From these statistics, one could make predictions of trends that are geographically established and based on the interests of information users.

Different data sources, including usage data, could be important in predicting citations as well as sharing (Colledge et al. 2015). Furthermore, the hottest topics of a research area are possible to be tracked with view counts, thus revealing the information seeking behaviors. However, the biggest obstacle to using usage metrics is questionable reliability because of easy manipulation. It is not hard to find a free automatic-click software on the Internet for boosting numbers of click views or download counts. Research products are complex outputs and they need to be considered from various sources for evaluation.

2.2.2 Peer-review

Peer-review performed by experts or expert judgement is considered as the top assessment of research quality (Mohammadi & Thelwall, 2013). Unlike web usage metrics, peer evaluation assesses academic literature qualitatively. Peer review is the process of assessing the quality of a manuscript before publishing it. A scientific publisher relies on expert peer review to maintain quality of the journals. The review process can be in single blind review, double blind review, open review, or collaborative peer review. In a single blind review, the identity of the reviewers is hidden from the author. In a double-blind review, both the reviewer and the author are anonymous to each other. An open review allows reviewers and authors to be known to each other and the review could be published openly online. Collaborative peer review is another way to assess scientific literature and it is done in collaboration among reviewers, and often the reviewers can decide whether to remain anonymous or not. In general, single blind review and double blind review are the most common types of peer review as with them, fairness and unbiased in reviewing a scholar product is guaranteed. For example, Chinese authors are not normally treated fairly if they are judged by open review (71% of scientific papers from China were rejected even without review) due to their bad reputation in plagiarism, unreliable findings or invalidation of methodology, and possibly other discriminating reasons (Primack, Arcelz & Koh, 2015).

In addition, there are both compliments and critics for open review; some praise plagiarism prevention and the encouraged sharing of different aspects, others criticize it for a less honest evaluation or politeness in evaluation. However, peer judgements consume much time and effort; according to Elsevier, a submitted manuscript is normally in the review process for 80 days before the final decision and the eventual publication after a couple of more weeks (Elsevier peer review policy and publication times, 2010). Later therefore, the impact of peer review is not counted as “how many” but instead “how positive”. Thus, the main purpose of peer-review is to mark quality, to check errors as well as plagiarism, to access the connections between theories, results and conclusions, all part of a standard scientific article.

2.2.3 Citation impact

Dating back to 1963, the roots of citation analysis is derived from the birth of the Science Citation Index (SCI), founded by Eugene Garfield (Desrochers et al. 2015). A citation creates a link between two scientific documents to show a connection, to indicate use of ideas from the earlier work, and to give credit to that work. Investigating citing behavior and citations' characteristics can evaluate the potential to apply citation analysis on research assessment (Haustein et al. 2015). A few studies concluded the people with a higher degree of expertise in their field tend to select fewer documents, read more, and cite more, that expresses a very specific information seeking behavior as well as referencing behavior (Lariviere, Sugimoto & Bergeron, 2013). Therefore, one of the most familiar citation measurement tools, H-index has been used widely by universities. The tool is a quantitative indicator to evaluate an individual researcher's performance by considering factors such as how active he or she has been publishing and how many citations those publications have received. For instance, a researcher that has an h-index of 8 means this author has published at least 8 papers which were cited at least 8 times each. Therefore, this is a common metric at author-level to measure productivity and citation impact of researchers (Desrochers et al. 2015).

Desrochers et al (2015, p.3) asked: "*Do citations measure scientific excellence?".* To answer the question, an overview of factors surrounding citation impact needs to be examined. There are several factors affecting citation counts, such as degree of internationalization and interdisciplinary of authors, language of literature, reputation of publisher or editorial board, number of references, and length of abstracts (Didegah & Thelwall, 2012). Moreover, the "*Tit-for-Tat*" is also applied in citing behavior; authors tend to cite their former citers and supervisors (Didegah & Thelwall, 2012). Multidisciplinary and multinational cooperation between authors lead to higher rates of citations. Didegah & Thelwall added that, language also does matter for citations and it causes an imbalance between non-English and English scientific literature; there is a clear difference in citation counts for scholarly products between authors from non-native English nations and native English speaking authors. Furthermore, the authors who are from highly ranked educational or research institutions receive more citations than their peers from lower ranked institutions (Didegah & Thelwall, 2012).

While Sud and Thelwall (2014) emphasized the importance of citations on research collaboration and as the sole indicator of value, Desrochers and her colleagues (2015, p.3) asked: "*Do you think that citation counts are the best indicators for the assessment of research*?". Collaborative research brings interdisciplinary aspects into research, but citation-based indicators are not always appropriate due to self-citations, thus, these are often removed from analyses of citations. In addition, citations are mainly used in journals and scientific articles, rather than books or other kinds of scholarly products (Desrochers et al. 2015). Citation-based metrics may reveal something about the citing behavior as well as citation predictions, but measurements based on citations need to be complemented with other types of metrics before concluding the quality of a scientific product. As a consequence, a neo evolution of metrics has been introduced to support the traditional metrics.

2.2.4 Altmetrics – the new meta

Nowadays, it is undeniable that scholarly communication is shifting to the online world and researchers use online sources to share and discuss research products (Holmberg, Didegah, Bowman, Bowman, & Kortelainen, 2015). Moreover, there are new ways that may be used to assess impact or influence of science. The complete range of various data sources include citations, peer-review assessments, and those often connected to altmetrics, namely number of tweets or retweets, social bookmarks, number of hits or shares and so on. Holmberg et al. (2015, p.2) wrote that altmetrics are used in two ways: *"(1) to describe the metrics resulting from the traces of the use and production of research products made in online environments, and (2) as an umbrella term used to identify the research field in investigating the meaning and application of these metrics and events"*. In short, Thelwall & Wilson (2014, p. 3) defined simply that altmetrics at article-level is "*an indicator that counts how often an article has been mentioned in a specific social web*". In fact, altmetrics has been investigating to be an alternative tool or complementary for estimating the influence of societal impact of open science.

Another benefit that altmetrics has is the potential to bring an enhancement to filtering of scholarly literature. In addition, open discussions on social platforms, such as science blogs may contain various online peer-reviews or summary of a scientific paper that could complement traditional peer-review. Apart from traditional research assessment, altmetrics provides other metrics that may reveal impact measurement of research products that are not peer-reviewed, such as unpublished papers, by checking indicators such as reading counts, blog posts, social bookmarks, tags, comments, tweets, academia forums or other online research collaborative discussions (Priem et al., 2010).

The visibility and awareness of research products is being summarized by altmetric attention scores. Unlike citation based metrics, altmetrics can show the traces outside academia, traces that reflect research influence mainly on social media platforms as well as other online sources. How are those altmetrics counted? According to Altmetric LLP (https://www.altmetric.com), a London-based startup company founded by Euan Adie in 2011 that provides services and data for tracking and analyzing online scholarly research activities, altmetric attention scores are derived from automated algorithm that calculate weighted counts of attention indicators from different online platforms. *Altmetric* serves for profits; it supplies products and commercial Application Program Interface (API) for customers such as publishers, institutions, researchers or funders, and free tools (Altmetric Bookmarklet, Institutional Repository badges, and API for research) as well as non-commercial license of the API which makes it possible to retrieve altmetric data. The standard weighted counts that Altmetric.com uses to calculate their Altmetric Attention Score are listed in Table 1.

News	8
Blogs	5
Twitter	1
Facebook	0.25
Wikipedia	3
Policy Documents (per source)	3
Q&A	0.25
F1000/Publons/Pubpeer	1
YouTube	0.25
Reddit/Pinterest	0.25
LinkedIn	0.5

Table 1. Altmetric weighted scores (cited from Altmetrics Support, 2015)

The altmetric attention score is always rounded to the nearest whole number. For instance, an article that is mentioned on two Facebook posts (0.5 point) will be rounded up to 1 point. Mendeley or CiteULike counts are not used towards the altmetric attention score due to problems with auditing the data, but the number of readers is shown on the individual altmetric page of the article. Altmetric attention scores may fluctuate to some degree over time; they may rise with more shares or decrease because tweets or posts are deleted. But the altmetric attention scores are not ideal as impact measurement metrics. Although different sources are given different weights, these weights may not reflect the social characteristics of the sources. For example, a tweet on Twitter about a scientific article, which is a very easy to share or retweet, both receive 1 point, while an article on Wikipedia, which usually takes much more effort and longer time to produce, receives 3 points. It is unclear whether this difference sufficiently reflects the increased value or importance of Wikipedia articles in comparison to tweets. Therefore, in terms of accumulation, Twitter is an ideal choice for boosting altmetric scores of an article due to speedy spam, regardless of its actual excellence or quality.

Figure 2 shows the total altmetrics scores for a specific article, including the mentions from social platforms. These are shown in a colorful circle, and each color represents a specific source. For instance, light blue expresses how widely the publication has been tweeted and dark blue indicates how many times it has been mentioned on Facebook. It could be said that the altmetric attention score is just that, an indication of the received attention and not of the quality of the article.

Figure 2. An example of altmetric attention score provided by altmetric.com to a publication

Recently, the imbalance between the altmetric attention score and scientific quality was highlighted, by a case that is now identified as *#Creatorgate.* A paper by Chinese authors on PLOS ONE reached a high altmetric attention score of 1884 points by March 3th, 2016 (Figure 3). The main reason to rocket the altmetric attention score of the paper was the controversy surrounding the use of the term "*creator*" as in a reference to a deity that designed the human hand. The paper later was retracted by the publisher because of this. As both negative and positive evaluations are counted towards the altmetric attention score and in altmetrics in general, the metrics can thus reflect some form of attention or awareness, rather than research quality. This is a big challenge for altmetrics. Negative reviews for a journal article and spamming from automated bot generators on online platforms all contribute towards boosting of altmetrics. Twitter is an ideal environment for automated bots due to real-time tweeting or retweeting immediately without checking contents, although anti-spam service has been developed. According to Akimoto (2011, p.2), "*Twitter-bots are small software programs that are designed to mimic human tweets*". Identifying bots on for instance Twitter is difficult, because they can "*reply to other users when they detect specific keywords*" or "*randomly tweet preset phrases such as proverbs*" (Akimoto, 2011, p.2). Apart from bots, there is an ethical issue that researchers and publishers have to face. They can promote their own papers on social media by for instance spamming, using bots, and draw their friends and colleagues' support to share their research products further. With that, there is an inappropriate manipulation of altmetrics that Altmetric.com is facing, and they are already counteracting against such behavior by attempting to examine data manually. This is of course a time-consuming effort (Adie, 2013).

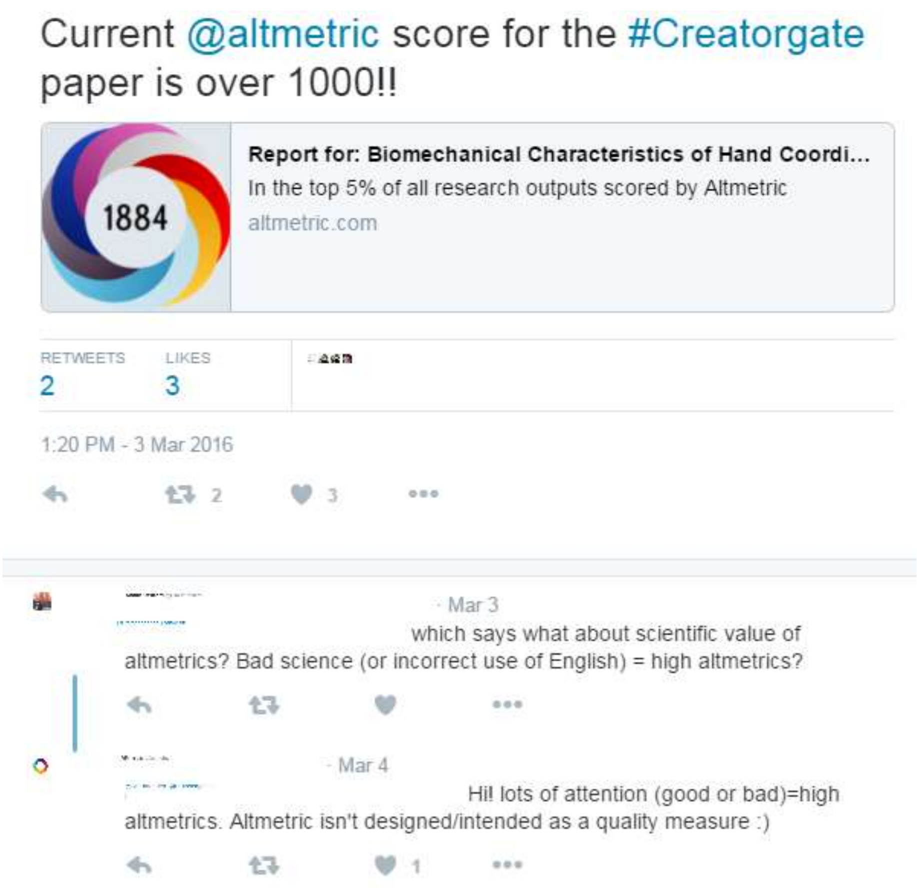

Figure 3. A controversial altmetric attention score for the *#Creatorgate* paper

Furthermore, Sugimoto (2015) raises a concern about the corruption of boosting altmetric scores; she suggested that libraries should start to educate researchers and university administrators about the use of altmetrics for researchers' activities such as merit evaluation, promotion and tenure. Furthermore, altmetrics may be prone to the "*Matthew effect*" of science (Bernal, 2013, p. 2). This means that research publications with more citations and higher altmetric scores will generate even more citations, attention or altmetrics scores, while other are left far behind.

2.3 Science blogging

Blogs are usually considered as an open space where bloggers keep their own diaries, or in the case of a community blog, publish bulletin announcements. Among numerous open resources and modern social media platforms, science blogs have their own position in scholarly communication as well as in altmetrics measurements. Science blogs produce posts or articles that are relevant to scientific news or research reviews from scholarly publication (Shema, Bar-Ilan & Thelwall, 2014). Science blogs disseminate information, URLs or citations perhaps slower than Twitter or Facebook, but they provide more 'user-friendly' texts to readers due to scientific content summarization, usually done by journalists, interested writers or editors who are experts in public writing. Moreover, bloggers or journalists are able to provide insightful reviews via blog posts along with a platform for open discussion without the limitations of space.

Although research blogging is relatively easy to start, maintaining attractive contents, contributions and views is another story (Finlay et al. 2013). The next sub-sections discuss impact, motivations, and citations in science blogging, bloggers' influence, and level of engagement in blog participation.

2.3.1 Impact, motivation and citations

Science blogging has been suggested to be a valuable indicator of scholarly impact. Several studies have been conducted to confirm the effect of research blogs on altmetric sources. According to Thelwall and his co-authors (2013, p.4), discovering data from the almetrics tracking provider revealed that "*three altmetrics (tweets, FbWalls, blogs) tend to associate with citations at the level of individual journals*". In addition, their data revealed "*a relatively strong correlation between blog and news outlet mentions and citations*". On blogs about science published by good reputation publishers such as PLOS ONE and Nature, page views are fairly high reflecting the scientific blogs' attention via usage (Shema et al., 2014).

The motivation of creating research blogs is not merely an interest in science, bloggers desire to disseminate their own thoughts of science to influence readers widely, plus share their skills as well as expert reviews. ResearchBlogging.org (RB) and Science-Blogs.com (2006) are places for enthusiastic writers write about their work as well as cites others' work, and peer-reviews.

There are blog citations and blog mentions on science blogs, despite of far less common use than traditional citations. Kousha, Thelwall, & Rezaie (2010) defined distinctly between *blog citations* and *blog mentions;* that blog citations do citing works as academic citation and appear on blog posts, whereas blog mentions are kinds of references or literature collections relevant to a blog's content. The scholarly communication on research blogs is informal and open; the tone in discussions or blog posts is mostly "reader-friendly", and there is no obligation of peer-reviewing in this social platform. Despite of that, Shema et al. (2014) reviewed some studies on real occupations of bloggers and found that, 59% of bloggers participated in academic community, 43% of the SciLogs blog authors work in academic institutions, and 45% of them hold PhD qualification. Researchers may exchange ideas via science blogs or practice their writing skills without worrying about complex process such as traditional peer-review, and in overall, they feel that their contribution and interaction appear in a public community.

However, some obstacles still exist. Only a few on scientific articles appear on blogs, and the dissemination speed is much lower than other social platforms such as Twitter. Though blog posts are not published quickly because much effort is demanded to write a science blog posts, meanwhile Twitter is much more simple. Next, the blog data is a problem. An assumption on specific aggregators for blog data misleads the impact of articles that appear on blogs when the data collection is out of range, and then relying on inaccurate information source causes doubts to scientists. Third, there is no sustainability or systematic preservation of blog environments. Blogs may lose comments any time, crash, or move without leaving any trace (although there is a possibility to crawl archives). Lastly, like any social web platforms, blogs are also facing spamming trouble; and as a scholarly impact indicator for altmetrics, unethical boosting of altmetric attention scores on science blogs can always happen (Shema et al., 2014).

For some readers, science blogging is entertaining reading with unexpected news or results from various research, but others consider those blogs as just personal corners of bloggers to express their own opinions without authoritative information sources. Research blogging is like an online communication version of traditional open peer-review, Shema and her colleagues (2014, p.3) said: "*If peer-reviewed journals citations are "frozen footprints", then citation in blogs, and altmetrics in general, are footprints in quicksand. In spite of these limitations, we consider blogs to be an especially promising altmetric source*".

2.3.2 Bloggers' influence

Bloggers on social platforms have an influential ability towards the public in their own way. Due to being a kind of "*content generators of digital social platforms*", they are capable to play a role as "*traditional opinion leaders, such as openness to messages, and taking the role of discussant, and influential and valued among social contacts*" (Uzunoğlu & Kip, 2014, p. 2). The leadership of sharing information publicly, especially on science blogs motivates readers, viewers, or followers who are also interested in science as well as have a good command of knowledge to engage in open discussions. According to Khan and his colleagues (2015), the influence of a blogger is measured by four factors: recognition (quantity of comments received), activity generation (number of comments posted), novelty (inverse proportion of outgoing links), and eloquence (comments' length). Additionally, they also proposed metrics to evaluate a blogger's influence such as MEIBI & MEIBX that trace the blogger's activity and support time-aware identification, or blog ranking metric (BI-impact) which measures a blogger's interaction on blog posts and the contents, and shares counts of linking from social networking like Twitter, Facebook, Google Plus. The engagement of both sides as blog authors and commenters is not always positive; for example, an aggressive blogger may have wars with commenters on his/her blog post due to disagreements in some contexts. Moreover, Shema, Bar-IIan, & Thelwall (2012) evidenced the deeply doubtful thoughts of science bloggers on the mentioned research findings that may influence skeptical opinions to readers.

2.3.3 Level of engagement

According to Haustein and her colleagues (2015), activities of those who participate on blogs are classified into three levels of engagement: *access*, *appraise*, and *apply*. *Access* refers to a way of approaching a research product, or a scholar objective such as scientific blog. Moreover, *accessing content* also comprises views hits, downloads, read counts, bookmarking counts, or storage of a research document from various social platforms as well as other open sources. *Appraising* indicates the acts of mentioning a scientific product on open platforms such as social networking sites, micro blogs, science blogs, mainstream media, Wikis, and other open access documents. Mentioning can appear as a linking URL on Twitter in a tweet, a brief post on Facebook, or even a citation on blog post. Appraising is an apparent act, so normally there would be reaction in form of comments, open discussions, reviews, or rates. Furthermore, in term of sentiment, appraising can be classified into compliments as positive, critics as negative and unbiased as neutral. *Apply* is considered as the strongest degree of engagement in scholarly context. Haustein et al. (2015) described apply as an act of adaption or transformation of a research product. The act probably involves applying methods, theories, frameworks, and findings into a new document, or even into the real life. At this level of engagement, it can be recognized via for example, content of a blog post or reviews based on discussion and limitations of a scientific paper, self-lessons adapted from its methods and results. Additionally, apply can also refer to information and knowledge collaboration in future research and experiments in scholarly context.

2.4 Gender differences

This section aims to provide readers a pre-comprehension before empirical research in this study. It is all about differences in communication behavior between males and females. First, how distinguished genders' appearance is discussed, for example, exposure versus hidden profile information between males and females on social media. Next, a distinction and different characteristics of the two genders in communication behavior and engagement is presented, for instance, how and why there are differences in communicative manners between men and women. Finally, an explanation of contrast in emotion and language use of different genders is given (e.g. women express more emotion than men).

2.4.1 Appearance on blogs

There are gender differences of blog authors in different blog formats involvement. Hollenbaugh (2010, p.2) reviewed several prior works and summarized that "*more women than men maintain personal journal blogs*" and that "*women are more likely to be inspired to post by personal experiences*". In contrast, men dominate science blogs authors, as was found on several platforms such as Scilogs, RB, and medical blogging (Shema et al. 2015). In addition, research on bloggers from UK found that women are interested in social contexts, while men are interested in information and opinions. Furthermore, women are more likely to disclose profile information than men, although women are more respectful of privacy and more active in communication (Hollenbaugh, 2010; Guadagno, Eno & Okdle, 2010; Pedersen & Macafee, 2007). Shema et al. (2012) investigated studied blog authorship on ResearchBlogging and concluded that the gender disparities were found to be simiar with Wikipedia contributors (only 12% are female contributors). Moreover, they also found that 72% of those blogs are written by at least one or two male authors (Shema et al., 2012). Therefore, the popularity of blogs authored by male is higher and received more inlinkings as well as outlinkings than female's. Female bloggers only represented a minority of open discussions on British blogging, except for topics about sexual confession which was discussed by 100% women (Pedersen & Macafee, 2007).

The online style between the two genders are also distinctly different. While male is biased to aggressive attitude as more adversial, self-promotion, sarcasm, female tends to be supportive and appreciative in a public community. As such, people normally guess other users' or bloggers' genders via online style expression when their profiles are anonymous. Additionally, a study about Bristish blogging by Pedersen & Macafee (2007) confirmed that women's blogs focus more on creative work and personal content generation, while men express opinions or link to other interesting or funny sites more than building their own blogs' contents. On academic blogging, bloggers often show their backgrounds for reliability or just to show their area professions; however, feminists still feel insecured and they tend to be guarded by hidden profiles for avoidance of bad intentions from online strangers such as stalkers or sexist abuse (Pedersen & Macafee, 2007).

2.4.2 Engagement and online communicative style

Atai and Chahkandi (2012) reviewed studies to investigate male and female online users initiate to an online dicussion. Although mixed results were found, a majority revealed findings that males tend to engage in online discussions or conversations in a negative way, while women were found to support discussions or prolong conversations in positive way. Atai and Chahkandi (2012) also indicated a stronger flame behavior in men on online social platforms that males are easily involved in fighting more than females. However, Herring (1994) believed that men represent as the symbol of promoting value of freedom from censorship, opening discussions as well as knowledge pursuit. Another study of Witmer & Katzman (1997) pointed out that when women involve in a fight, especially with men who have inflammatory or sacarm manners, they are also easy to lose control and express bad langage. The above opinions are dependent on their own studies, different individuals may express in their own ways to engage in their favourite topics.

Sussman and Tyson (2000) hypothesed that men would engage in communication more frequently than women. However, their findings are in the reverse order based on an investigation of postings' length on online blogs and Internet newsgroups. Women initiate discourse more than men, especially in feminist topics. Females tend to anticipate information as well as propose open questions to extend the longevity of postings, while males only participate in opinionated conversations that measure posts' contents, especially on "male-linked issues". On the other hand, Herring (1994) reported that males dominate online interaction by producing longer posts more frequently than females in various discussion topics. In terms of scientific discussions, for example, open discussions on science blogs, females are likely to be more affirmative on female-related studies, library service or education, while males are interested in engagement in topics of politics, sports, computing, and philosophy discourse (Herring & Paolillo, 2006). However, Guiller and Durndell (2006) pointed out that towards individuals perspective, sometimes prolonging discussions or reaction on postings is just "for fun" rather than serious engagement in a discussion.

2.4.3 Emotion and language use

Women tend to express more explicitly emotional expression than men, although this can vary from culture to culture as well as different social contexts and degree or frequency of emotional expression (Thelwall, Wilkinson & Uppal, 2009). An emotional expression could be in a negative or a positive way, but men and women have different facets to express their emotions. For instance, female tends to be depressed, panic, show empathy or confused in a vulnerability case, while male are biased to show anger in a similar situation (Thelwall et al. 2009; Brody & Hall, 2014). In the positive way, female behavior is normally expressed in joy, happiness, whereas men are kind of proud, or show self-pride (Thelwall et al. 2009). In ordinary social life, women also have a tendency to use positive emotion frequently, for example, caring activities such as childcare, or chatting with neighbours. Similarly, Herring (2000) concluded that women are represented as good and soft by saying "*thanks*", appreciate or apologize more, additionally, they are easy to be upset by offensive comments or other kinds of "violations of poliness". In contrast, men preferably hide their emotion in most daily situation (Thelwall et al. 2009). As such, those emotion expression behaviors are also applied distinctly based on gender on online social spaces such as social networking sites, blogs, or forums. Therefore, a gender recognition is possible to explore characteristics of emotion behavior in socialization.

When communicating online, it is sometimes challenging to know the real meaning as well as attitude or implication via text-based communication without facial expressions. *Emoticons*, or *smileys,* created from "*emotion*" and "*icon*", are normally added into comments to support online users' expression. Witmer & Katzman (1997) found from a study of online messages that females tend to use more graphical accents or emoticons, normally positive "*smiley*" emoticons in conversation to support their emotion expression than males. Emoticons support the impression or attitude of messages' receivers. Moreover, Thelwall and his co-author (2009) also confirm a difference of using emoticons between genders; female who use affiliative language regularly attach positive or joyful emoticons with their language expression on social discussions such as comments and posts, whereas male are totally reversed with negative or even offensive emoticons. On the other hand, males usually desire to hide their real feelings behind screens, including common negative feelings such as sadness, depression or disappointment (Huffaker & Calvert, 2005). Interestingly, Lee (2003) analyzed instant messaging dialogues among students in Stanford University and found that males rarely use emoticons to each other, but they produce more emoticons when communicating to females, while females use emoticons as an equal amount of frequency to both genders.

It is derived from natural emotion behavior of genders to expose their language use in different styles in social communication. Women are intentional in showing positive engagement (*affiliative language*) via their language expression, while men tend to use *assertive languge* directly in stronger degree, including critism, sacarm (Thelwall et al. 2009). Generally, women show a positive face on social communication than men due to the soft style of socialization which leads to effective co-operation in a community. This is not only applied in off-line interaction, the online world is also influenced by a similar pattern. Guiller and Durndell (2006) carried out a study of Internet group discussions and found that a female-only group discussion showed more emotional and supportive language in a community, while a male-only group shows less emotion expression and is more likely to use positive comments on a blog. Regardless this study, norm and social value may vary in different cultures or different contexts, so it is not clear yet to draw a conclusion of which gender has a “better” performance in expressing their emotion through language use, even though women are assumed to be successful communicators based on their natural communicative behavior.

Chapter 3. Methodology

This chapter includes two parts; data presentation which describes the process from collecting raw data to refining sample selection for study; and methods presentation that explain how the study analyses are carried out.

The study aims to explore gender differences from social discussions on science blogs which mention scientific articles produced by authors or co-authors from Finnish academic institutions. As such, the raw data was collected from a project in a Finnish university, and the data was filtered based on three criteria: 1) Digital Object Identity (DOI), 2) altmetric attention score, and 3) availability of comments on science blogs.

The core characteristic of methods includes gender identification, content analysis, and sentiment strength detection. As its core, the data was explored by tools such as Gender-API for gender identification, SentiStrength for sentiment strength detection and analysis, while content analysis is pre-defined by a codebook to analyze comments' contents and gender differences in commenting behaviors.

3.1 Data

The raw data in this study was provided by RUSE (Research Unit for the Sociology of Education), University of Turku, Finland, from their current research project "*Measuring the societal impact of open science*". The project is an investigation of tracing altmetrics data that may record open science activities to explore potential awareness, attention as well as societal impact derived from Finnish research. Initially, the data involve three social media platforms Twitter, blogs, and news; but after specifying this study, the author focuses on blog platform and uses only the data for blogs.

Overall, the raw data for blogs include 338 Finnish scientific articles in the field of science and technology, that are mentioned 937 times on various blogs. In the context of Finnish research, those articles are produced not only by Finnish researchers, but also by researchers who work for Finnish institutions or by co-authors from Finland. The data covers full records of articles' altmetric attention score, articles' DOI, citations as well as mentions on blogs, date, time, link, and summary of blogposts, blog authors and their description. Nevertheless, some details are not accurate, such as blog authors' description, and the original DOIs in the field of science and technology that is required to re-crawl data.

The following sub-sections explain how the raw data are filtered and describe the sample selection for this study.

3.1.1 Filtering data

In the very first stage of filtering data, the altmetric attention score which was explained in section 2.2.4 is a milestone to select data for this study. According to Altmetric Support (2016), if an article is evaluated at scores of 20 or more, it is supposed to be "*far better than most of its contemporaries*". Being aware of this score level does not indicate "good" or "bad" of an article, instead, attention measurement is the core value. Therefore, all of the articles that received a score of at least 20 altmetrics score are kept, making the sample 199 articles out of a total of 338 articles from the raw data. The next stage is to determine the articles' field of science based on their original DOIs in order to extract the records precisely as well as to narrow the appropriate data. Web of Science by Thomson Reuters was used to search those DOIs, and the given results were only 136 DOIs in the field of science and technology over total of 199 DOIs. Additionally, those 136 articles are mentioned on 567 blogs. The last stage is to choose blogs which contain additional comments for further analysis. In this phase, 567 blogs were crawled manually including some links that were no longer existing that were searched for from Internet Archive (https://archive.org). Eventually, there were only 69 scientific articles mentioned on 184 blogs which contain additional comments, and this is the refined data for sample selection of this study.

In summary, the filtering data process passes three stages including altmetric score ($\geq$ 20), DOI (field of science and technology) and blog posts which contain comments. The number of articles mentioned and number of blogs are both reduced gradually after refining data for precision and feasibility of further analysis in the following steps in this chapter.

3.1.2 Sample selection

After filtering data, the usable data contains 69 articles mentioned on 184 blogs that would be chosen for comments sample selection. One thousand comments is the ideal number to fit this study according to the author's supervision. Therefore, a random sample from the total of 4311 comments on 184 blogs was extracted systematically to a limit of 1000 comments. The number of comments on each blog varies. Some have hundreds of comments (the most one has 652 comments including replies), others have only 1 comment.

Initially, the average number of comment on each blog is counted by 4311 comments over 184 blogs. The results were around 23 comments. Top comments are not always the most recent comments, sometimes they were on the top positions persistently for the large number of likes, replies, or nominated by the page author. Next, the blogs were put in order from the largest number of comments to the smallest number of comments, then 23 top position comments from each blog were extracted until the total extraction reaches 1000 comments. Most of the comments are English (942 over 1000 comments), the rest include Polish (33 comments), German (10 comments), Spanish (9 comments), Swedish (6 comments). Thanks to the availability of SentiStrength, all of those languages can be processed by its own adequate language versions that enhance the quality of contents analysis as well as the sentiment strength evaluation. In addition, all extracted comments are relevant to scientific topics mentioned on the blogs. Comments created by the "Reply" button, were skipped because they turn to private conversations rather than public communication; and the next comments would be chosen to fill up to 23 comments on each blog (for blogs that have exact or fewer 23 comments, all comments will be taken without "reply" comments).

All of the comments have been privately archived by the author, making a total of 1000 comments on 62 blogs (32 articles mentioned on these blogs) which have the largest number of additional comments. This way of sample selection would choose the most attentive comments (top comments) from the most interactive blogs (blogs with the most additional comments).

3.2 Methods

The core characteristics of methods in this study are exploratory along with supportive tools such as Gender API for gender detection, SentiStrength for sentiment strength detection, and IBM Watson analytics for data visualization. This section starts with sentiment strength analysis which is executed by SentiStrength for each comment. Next, the codebook for content analysis is created to classify comments in 4 main groups of comment about bloggers, comment about blog post, interaction with other comments, and other. Finally, gender identification of both blog authors and commenters was done using Gender-API before visualizing all findings on IBM Watson analytics.

3.2.1 Sentiment strength detection and analysis

Sentiment analysis indicates how positive, negative, or neutral emotion people express via their comments. The SentiStrength tool (http://sentistrength.wlv.ac.uk) was used to determine sentiments of each comment. The tool uses a lexicon with pre-defined sentiment-word-lists mostly derived from Linguistic Inquiry and Word Count (LIWC) to process linguistic analysis based on its own algorithm rule. The sentiment analysis results produce two scores of weighted scales for analyzed texts to optimize detection mixed emotions in a sentence. The scores are between -5 (extremely negative) to -1 (not negative) for negative, and between 1 (not positive) to 5 (extremely positive) for positive. There is no value 0 because the scores 1 and -1 represent a neutral sentiment. To evaluate the sentiment analysis, combined sentiment scores (i.e. negative scores plus positive scores) are calculated and compared. There are three cases of combined value: equal 0, larger than 0, and smaller than 0; they are represented as neutral, positive, and negative sentiment respectively.

However, SentiStrength performs high accuracy for only short texts. For long texts, a summarization is required before testing to guarantee the accuracy rate. The tool was originally built for English texts, but since it has been developed, it is now available in many other European languages despite limitations. Moreover, SentiStrength also has lists of 115 common emoticons collected from social web, traditional English idioms, slang words, booster words, and question words, all those are to improve the accuracy of emotion evaluation or sentiment analysis.

The core of SentiStrength focuses on adjective words which express emotion, some nouns or pronouns could be skipped or be left as neutral when analyzing a sentence. For example, a sentence “I felt frightened, but it was an awesome experience for me” could provide two scores as -4 for negative scale and +4 for positive scale because two words “frightened” and “awesome” are emphasized to predict sentiment, while the other words are measured as neutral value. The tool also has a related-topic function to predict sentiment in a particular circumstance. For instance, applying the above sentence into a topic of watching a movie, “frightened” would be measured as a positive value. For the texts in term of negating negative (i.e. “I don’t hate them”), the sentiment result stands for neutral value, not positive value. Additionally, emphasized words with repeated letters (i.e. sooooo coooool), or adding emoticons repeatedly, (i.e. sooooo coooool <3 <3 !!!!!) is set to boost the sentiment value to very positive score or very negative score (Thelwall & Buckley, 2010). Those kinds of words normally appear on social media via public comments to express explicitly the real emotions of people by texts.

Despite the automatic word correction function of SentiStrength, the sentiment analysis still struggles with ambiguous words such as sarcasm, regional dialects or manipulated words, and idioms for jokes. The tool can only correct simple spelling mistakes (i.e. "hepl" = "help", "dat" = "that", "asap" = "as soon as possible") to evaluate sentiment strength, as it is impossible to recognize various complex chatting language or lingo on social web that is created by anyone (i.e. "nab" means "noob", "ffs" or "fgs" means "for fuck's sake" or "for god's sake"). To optimize sentiment strength and assess the accuracy of SentiStrength for a large data set, the tool offers 10-fold cross-validation function that repeats 10 times by each time and 10% of the data are tested and the 90% remained are trained, then reporting the total results.

3.2.2 Content analysis

This method is basically developed to classify the content of a sample selection as well as to show the results of sentiment analysis on the comments. There is a total of 4 groups to differentiate the types of comments in the codebook (Table 2) including comments about blogger, comments about blog post, interaction with other comments, and other kinds of comments. In addition, the four classes of sentiment analysis are positive (P), negative (N), neutral (U), and mixed sentiments (M). In each group, there are sub-groups to classify type of comments by their sentiments. For example, group A is comment about blogger, then sub-group A1 is praising the blogger (positive), A2 is criticizing the blogger (negative), A3 and A4 are respectively comment about blogger's personal stuffs and other comment about blogger (both neutral or mixed sentiment). To support the reliability of inter-coding, apart from the author self-evaluation, the sentiment strength detection software SentiStrength is used to validate the sentiment analysis (section 3.2.1) in sub-groups according to explicitly defined groups.

GROUP	TYPE OF COMMENT	DESCRIPTION
A	Comment about the blogger	Given opinions all about the bloggers regardless of how good or bad of the blog post was.
A1	Praising the blogger	Praising the blogger rather than the content of the blog post.
A2	Criticizing the blogger	Criticizing the blogger rather than the content of the blog post.
A3	Comment about blogger's personal stuffs	Comments about blogger's background, age, gender, appearance, etc.
A4	Other comment about blogger	Any other comment about blogger that does not fit the above classes.
B	Comment about blog post	Given all relevant interaction to blog post's content.
B1	Praising the blog post	Positive evaluation to the blog post's content.
B2	Criticizing the blog post	Negative evaluation to the blog post's content.
B3	Summarizing the blog post, quoting key points	Given a brief overview of the blog post, repeat key points of the blog post.
B4	Further discussion to the blog post	Neutral discussion related to issues within the blog post for sharing knowledge.
B5	Other comment about blog post	Any other comment about blog post that does not fit the above classes.
C	Interaction with other comments	This group is for only comments that react to other comments regardless of the blog post as well as blogger.
C1	Praising other comments	Praising other commenters rather than the blog post or blog author.
C2	Criticizing other comments	Criticizing other commenters rather than the blog post or blog author.
C3	Other interaction	Any other interaction with other comments that does not fit the above classes.
D	Other	Other kinds of comments that do not fit the above classes.
D1	Spam	Irrelevant annoying comments appeared repeatedly.
D2	Self-promotion	Comment sounds like marketing that is irrelevant to the blog post's context.
D3	Linking	Comment contains hyperlinks, citations.
D4	Other	Other comments do not fit the above classes

Table 2. The codebook of 4 groups for comments' content analysis with description

Although each group in the codebook is classified clearly, the data is sometimes reflected in mixed groups. That means some pre-defined groups are not mutually exclusive. For instance, a comment may be given multiple codes that belongs to multiple groups such as both comment about blogger and interaction with other comments at the same time. However, the author considers group B (comment on blog post), group A (comment about blogger), and group C (interaction with other comments) are exclusive, meaning that comment is coded in group B would not belong to group A or C. Only duplicate codings are between group D and the other groups. This was made as a priority method of content analysis to classify distinctly between the relevant comments to blog posts and without reference to the blogs' contents.

The sentiment is tested by both self-evaluation of the author and SentiStrength, then it is applied into each group. For example, a "good manners" comment such as "This blog is outstanding, love it !!!" about blog post is coded as BP, where B means comment on blogger and P stands for positive sentiment. Given multiple codes probably happen if a comment matches multiple groups as well as types of sentiment; for instance, a given comment "I agree with Pete, this blog post is trash" would be coded as CP and BN.

3.2.3 Gender identification

After gathering all samples of comments, gender identification is processed based on the data retrieval of the commenters as well as the blog authors. The main method to detect genders is dependent on Gender API (https://gender-api.com) which provides gender identification based on the first name. The site is free for registered user for the first 500 requests of gender identification, after that a free member can request only 20 names each day. The author relies on this provider for its wide range of data that records the most common names across the world (supporting 178 countries). Moreover, when a first name is tested, statistical information about the name is shown such as accuracy (percentage value of gender identification), gender (determined gender), samples (the number of records used for gender identification), and duration (how long to determine the gender identification). Among those factors, accuracy is the most determined information. If the accuracy value is more than 50% for a name, the gender is determined; otherwise, it is unknown or it is considered as a unisex name.

To support the reliability of gender identification, an alternative method given. Based on the extracted comments and the blogs' links, profiles of the commenters and the blog authors are archived manually if they are available. Thus, those people's profile pictures and their public information about gender are used as additional sources to verify their genders. Additionally, the bloggers' academic status, when available, is also explored by this way for further analysis.

Chapter 4. Results

This chapter reports findings of four analyses based on sentiment strength detection, impacts of altmetric score, content analysis and gender differences. As the methods' structure, sentiment strength detection on 1000 comments by SentiStrength is reported initially, then the results are applied in the following section of altmetric attention score's impacts. Next, the content analysis categorizes comments into different groups from the pre-defined codebook including comment about blogger, comment about blog post, interaction with other comments, and others. Finally, exploration on gender differences was carried out on commenters and blog authors. This investigation relies on sentiment strength detection to determine emotional attitudes of commenters on blog posts rather than contents analysis.

4.1 Sentiment strength detection

There is a balance between negative and neutral discussions, with 359 negative comments and 389 neutral comments, whereas positive comments were found 252 times.

Due to the mixed emotion in sentiment strength detection of Sentistrength, all texts were analyzed under two scales, that can be combined to a unique result as plus score for positive sentiment, minus score for negative sentiment, and exact zero score for neutral sentiment. Table 3 shows the report of sentiment detection including combined scales to provide an insight into social commenting behavior according to commenters' emotion expression. While combined results at +3 takes only 2,78% among total of positive comments, +2 and +1 combined scales share the rest with 29% and 68,22% respectively; and no extremely positive score (+4) was found as positive sentiment. In negative comments, there was with only one extremely negative score (-4), and most comments are found in the less extreme negative groups. For example, scores at -2 and -1 have proportion of 30,08% and 61,28% respectively. The average combined scores were calculated from the average positive scales and negative scales detected in each comment. This gave the average results of how much weighted scales as well as emotional expression were detected from those discussions.

SENTIMENT STRENGTH	COMBINED SCORES	PROPORTION	TOTAL OF COMMENTS	AVERAGE SENTIMENT SCORE
POSITIVE	+3	2,78%	252	1,35
	+2	29%		
	+1	68,22%		
NEGATIVE	-4	0,28%	359	-1,48
	-3	8,36%		
	-2	30,08%		
	-1	61,28%		
NEUTRAL	0	100%	389	0

Table 3. Summary of 1000 comments' sentiment strength detection

In general, negative average combined scores (-1,48) was found as a little more extreme than the positive result (1,35). The average scores show that there is no strong sentiment in the comments, that indicate a tendency for neutral or less extreme attitudes of commenters in social discussions.

4.2 Altmetric attention scores vs. number of comments

As one of research direction, this study investigates the connection between quantity of discussions and altmetrics scores. The Figure 4 shows a comparison values between altmetrics attention score versus number of comments on the science blogs. The number of comments is shown descending in range 10 - 652 on each blog in total of 62 blogs, and altmetric scores of mentioned articles were attached to those blogs adequately. The correlation coefficient was calculated based on the two variables, and the given result was around 0,18. This is a strong indication that there is no connection between altmetric attention scores and number of comments because a correlation coefficient near 0 indicates no correlation.

Correlation coefficient = 0,178869

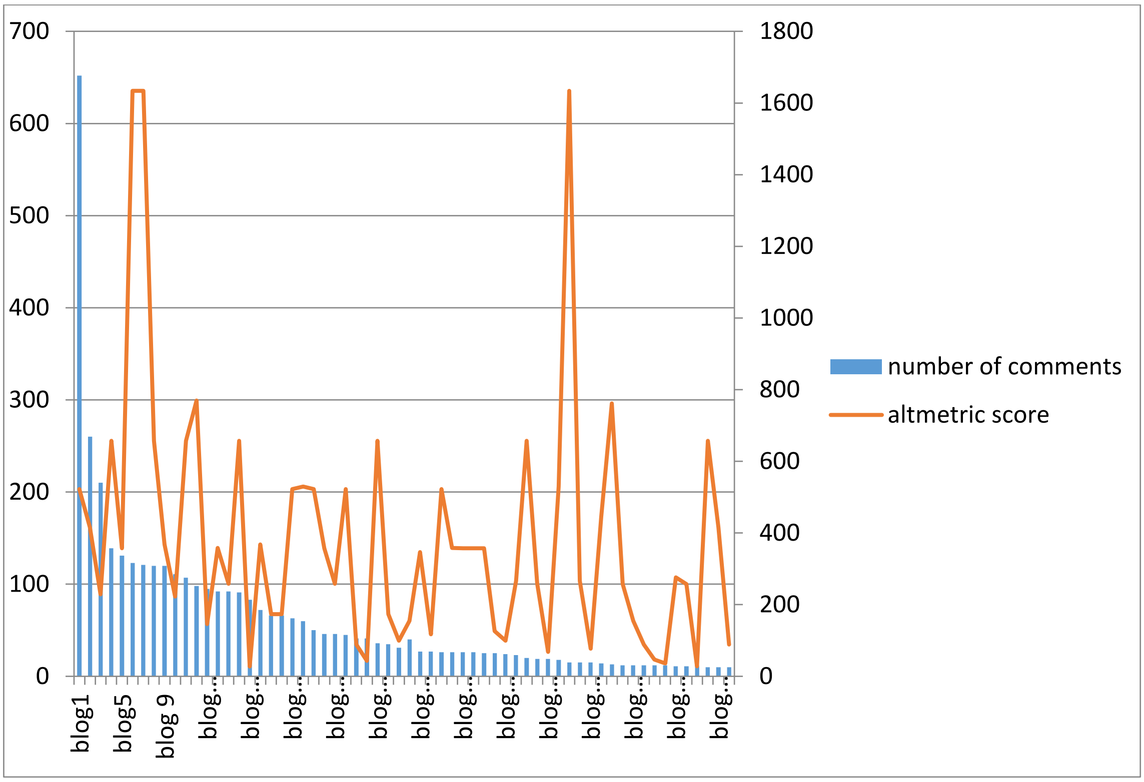

Figure 4. A comparison of values between altmetric attention scores and number of comments

As can be seen, number of comments do not match the trending of altmetric attention scores. Blog which has the largest number of comments (652) reaches medium altmetric attention score (around 500), while some blogs with lower degree of discussions (over 100 comments, or even below 50 comments) rank at the top score (over 1600), and the smallest number of comments in a blog does not mean the lowest altmetric attention score.

4.3 Content analysis

A content analysis of all comments from science blogs were extracted. Each comment was codified in groups and some also belong in multiple categories; therefore, the total value of percentages in the summary codebook may exceed 100%. The results were produced based on the author's manual coding along with the support of SentiStrength to improve the accuracy of content analysis. The codebook's results in this section also give typical examples for comments in each group. Most examples are quoted directly from the sample.

GROUP	TYPE OF COMMENT	PROPORTION	EXAMPLE
A	Comment about blogger	3,4%	
A1	Praising the blogger	1,6%	Well put Anne.
A2	Criticizing the blogger	1,3%	Stop copying istockphoto
A3	Comment about blogger's personal stuffs	0,1%	Alice, you must have the looking of your mother then
A4	Other comment about blogger	0,4%	What time are you scheduled to be on News Hour?
B	Comment about blog post	62,7%	
B1	Praising the blog post	9%	Very exciting! Lots of good stuff in the supp info.
B2	Criticizing the blog post	11,8%	So this isn't actual imaging of peo-ple bodies? How disappointing.
B3	Summarizing the blog post, quoting key points	2,4%	To quote the article "it's a horren-dously complex system" and yet Al Gore and the climate change fa-natics say it settled science.
B4	Further discussion to the blog post	34%	Autism describes a range of ef-fects. It isn't surprising that many genes are implicated.
B5	Other comment about blog post	5,5%	But...but...vaccines.
C	Interaction with other comments	25,3%	
C1	Praising other comments	2,6%	Debate is good!
C2	Criticizing other comments	8,9%	The above comments (from Hughes and Dole) don't address the article's main point
C3	Other interaction	13,8%	Did you read the paper?
D	Other	17,9%	
D1	Spam	3,3%	Por que…Por que…Por que…
D2	Self-promotion	0,4%	I make $82/h while I'm traveling the world. Check it out, w*w^w . Best96 . c^o*m-
D3	Linking	11,4%	http://globalcompostproject.org
D4	Other	2,8%	Phuket

Table 4. Results of comments after codifying in each group with typical examples

The results (Table 4) show that group B as comment about blog post dominates with 62,7% in overall; followed by group C and D as interaction with other comments and other interaction by 25,3% and 18% respectively, while group A as comment about blogger makes the smallest piece at 3,4%. Due to the mutual exclusive between group A, B, and C, the percentages for comparison between them are absolutely distinctive. Consequently, the results express clearly the gap between group B with the rest, that people still pay attention and discuss about the blog post's contents rather than other subjects such as previous comments or the blog author.

GROUP	TYPE OF COMMENT/SENTIMENT	PROPORTION	AVERAGE SENTIMENT SCORES
A	Comment about blogger	3,4%	
	Positive	1,2%	1,42
	Negative	0,8%	-1,5
	Neutral	1,4%	0
B	Comment about blog post	62,7%	
	Positive	16,9%	1,37
	Negative	23,9%	-1,5
	Neutral	21,9%	0
C	Interaction with othe comments	25,3%	
	Positive	6%	1,27
	Negative	9,5%	-1,45
	Neutral	9,8%	0
D	Other	17,9%	
	Positive	3,2%	1,375
	Negative	5,1%	-1,47
	Neutral	9,6%	0

Table 5. Proportion of sentiment analysis applied in the codebook

The findings of sentiment detection applied in the sentiment codebook (Table 5) provide differences in commenting behavior in each group, although these differences are fairly small. As for the proportion of comments, in group A there are slightly more neutral comments (1,4%) than positive (1,2%). In comments about the blog posts, negative (23,9%) comments are most frequent, followed by neutral (21,9%) and positive (16,9%). In groups C and D most comments are neutral (9,8% and 9,6% respectively), closely followed by negative and positive. It is not surprising that neutral comments dominate in group D because it contains all kinds of comments which are irrelevant or unbiased to blog contents as well as no signs of social interaction with others. However, based on the sentiment detection, on average the comments overall are slightly more negative.

4.4 Gender differences

Gender detection was carried out using Gender API tool on both blog authors and commenters. Therefore, the gender differences' results are shown and compared from two perspectives; blog authors and commenters. If the accuracy of the gender detection was equal or under 50%, individual's profile was manually explored if it was available. In other cases, the gender was determined as unknown.

GROUP	TYPE OF COMMENT/SENTIMENT	MALE	FEMALE	UNKNOWN
A	Comment about blogger	3,9%	3,6%	2,66%
	AP	1,36%	0%	1,33%
	AN	0,98%	0,9%	0,53%
	AU	1,56%	2,7%	0,8%
B	Comment about blog post	63,55%	64,86%	60,9%
	BP	17,24%	16,22%	16,22%
	BN	24,37%	19,82%	24,47%
	BU	21,64%	28,82%	20,21%
C	Interaction with other comments	25,34%	21,62%	26,33%
	CP	6,04%	5,4%	6,12%
	CN	9,16%	8,11%	10,37%
	CU	10,14%	8,11%	9,84%
D	Other	18,71%	14,41%	15,69%
	DP	4,29%	2,7%	2,93%
	DN	4,48%	3,6%	5,05%
	DU	9,94%	8,11%	7,71%
TOTAL		100%	100%	100%
	Positive (P)	28,93%	24,32%	26,6%
	Negative (N)	38,99%	32,43%	40,42%
	Neutral (U)	43,28%	47,74%	38,56%

Table 6. Gender differences in commenting attitude of commenters

Differences in the commenting attitude of commenters based on their gender is shown in Table 6. Overviewing the total sentiment results, there is no differences between male and female, although there is a remarkable inequality in the number of men (513 commenters) vs. women (111 commenters), the rest is unknown gender for 376 commenters.

Zooming in different groups from content analysis some differences are revealed. Regarding comments about bloggers (group A), based on manual coding and SentiStrength detection, female showed no positive comment towards bloggers, while male had 1,36% of all male commenters in the same context. In addition, females tend to post more neutral comments than male (2,7% vs. 1,56%), and they share the similar ratio of negative comments. Group B shows evaluation of blog content, that both men and women have similar findings of positive attitude, but men seem to post more critical thoughts about the blog content than women do (24,37% vs. 19,82%), while women tend to be clearly more neutral. The last two groups are more equal. The highest percentage (24,37%) of males belongs to negative comment about blog contents, while female reaches the top rate (28,82%) in term of given neutral opinion to the blog content. As such, commenters seem to be interested in blog content more than anything else, but their gender expresses somewhat different attitudes in this subject.

There are 62 science blogs written by the same number of blog authors (35 males, 19 females, and 8 unknown) who are also investigated in their gender differences from different perspectives in this study. Due to the domination of male blog authors, the total of comments in blog post written by men is also the largest at 565 comments, women bloggers received 293 comments, and 142 comments are given in blog posts of unknown bloggers. This is also the same situation in commenters' perspectives, that males contribute the largest comments in both males' and females' blog posts (305 and 162 comments respectively), whereas female commenters have far lower numbers of comments (40 in males' blogs and 56 in females' blogs).

Table 7 presents comparisons of male and female commenters on blog posts according to gender of blog authors. The differences in gender from two perspectives of commenters and blog authors are conducted based on sentiment strength detection. There are some significant differences in commenting behavior of commenters according to different gender of blog authors. Females tend to be more neutral social discussions on blogs of male authors (-0,02 as total of combined sentiment score), while they show more positive emotion if the blog authors have the same sex (+0,14). By contrast, men's comments show more emotion in negative way if blog authors have a different sex (+0,27 as male commenters – male bloggers vs. -0,65 as male commenters – female bloggers). The common thing between male and female commenters is that they tend to be more positive in discussions on blogs of authors who have the same gender. As shown, the total of combine sentiment scores for male commenters – male bloggers are +0,27 and for female commenters – female bloggers are +0,14.

BLOGGER	COMMENTER	POSITIVE	NEGATIVE	COMBINED SCORE	NUMBER OF COMMENTS
Male	Male	1,09	-0,82	0,27	305
	Female	1,08	-1,1	-0,02	40
	Unknown	0,73	-1,23	-0,5	220
Female	Male	0,64	-1,29	-0,65	162
	Female	1,7	-1,56	0,14	56
	Unknown	1,43	-0,56	0,87	75
Unknown	Male	1,36	-0,71	0,65	46
	Female	1,33	-1,75	-0,42	15
	Unknown	0,62	-0,1	0,52	81

Table 7. Average sentiment score towards gender between commenter and blog author

Along with gender, profession of blog authors is also examined to compare the differences in receiving responses on their blogs. bloggers were explored as 3 main professions as citizen scientist who is interested in science but not working in science fields, faculty (professor, researcher) who works in science for academic institutions, and scientific journalist who is writer, editor, staff of a journal or a publisher.

The difference in gender's profession is that there were no citizen scientists as women bloggers, instead, they are news journalists and reporters who work for news journals rather than scientific journals. In total of 19 female bloggers and 35 male bloggers, most bloggers are scientific journalists in both men (13 people) and women bloggers (9 people); the least founded professions of male are faculty (professor, researcher) (5 bloggers) and female news journalist and reporter (2 bloggers).

As shown in the Table 8, blogs written by scientific journalist receive the largest number of comments, including 214 responses for male bloggers and 119 responses for female bloggers, followed by 192 comments on blog posts of male citizen scientists and 102 comments to female researchers. Faculty (researcher or professor) receives the least comments among male bloggers, while journalist and reporter have the similar situation among female bloggers.

BLOGGER	PROFESSION	POSITIVE	NEGATIVE	COMBINED SCORE	NUMBER OF COMMENTS
Male	Citizen scientist	1,25	-0,83	0,42	192
	Faculty (professor, researcher)	1,47	-1,53	-0,06	53
	Scientific journalist	0,88	-1,05	-0,17	214
	Unknown	1,23	-0,54	0,69	106
Female	Faculty (professor, researcher)	1,5	-1,59	-0,09	102
	Scientific journalist	0,28	-0,45	-0,17	119
	News journalist & reporter	1,67	-1,36	0,31	26
	Unknown	1,8	-1,03	0,77	46
Unknown	Citizen scientist	1,27	-1,3	-0,03	37
	News journalist & reporter	0	-1	-1	3
	Unknown	0,08	-0,48	-0,4	102

Table 8. Average sentiment score towards blog authors' gender differences in receiving comments according to their profession

In addition, the findings show on emotional commenting behavior on blogs based on differences of gender as well as profession of blog authors. There was no remarkable difference between males and females who comment blogs written by scientific journalists (both at -0,17) and professor or researcher (-0,06 for male vs -0,09 for female). Even professional in science such as scientific journalists or faculty researchers are also responded negatively. This reveals that the tendency of critical concerns may happen to any profession inevitably regardless expert or non-expert authors. On the other hand, both male citizen scientists and female news journalists and reporters received positive comments (+0,42 and +031 respectively). Nevertheless, unknown bloggers in both gender receive positive comments with the highest degree of average sentiment score (0,69 for male bloggers and 0,77 female bloggers) versus others.

Chapter 5. Discussion

Based on the literature review and the study's outcomes, this chapter discusses and compares all relevant matters. It starts by investigating altmetric attention scores, that were used to examine the connection of sentiment analysis and comments. The discussions emphasized the main impact of altmetric attention scores towards public attention then assessed the potential usage of altmertics in an aspect of research quality measurement. The next section discusses how the public reacts to science blogging based on findings on content analysis and literature reviews. After that, gender differences are explored deeper from the perspectives of commenters and of blog authors. This analyzes the above findings and compares differences as well as results from earlier studies. Finally, limitations of this study are discussed before suggesting directions for future research.

5.1 Altmetrics measures volume of attention

In this study, the altmetric attention score is not only used for refining data, but also as a measurement tool for attention in social discussions on science blogs. Although altmetrics is a new development of metrics in measuring attention from social aspects and not yet considered as a standard measurement, it brings interesting findings to this study.

Both altmetric attention score creator (Altmetric) and earlier research indicate that this score on research products tends to be a social attention measurement rather than a quality evaluation. In addition, Thelwall & Wilson (2014, p. 3) referred to altmetrics in article-level as "*an indicator that counts how often an article has been mentioned in a specific social web*"; and the data in this study is comprised of scientific articles mentioned on science blogs.

To examine the above characteristic of the altmetric attention score, sentiment strength was measured on 1000 comments from science blogs in order to see the controversy between quality and attention. In terms of attention, emotional thoughts such as strong sentiment is usually mentioned, but why less emotional responses such as neutral sentiment also raise a big count in altmetric score? There was a discussion about automated bots which automatically generate sharing, commenting, spamming (Akimoto, 2011) to boost altmetric attention scores. Additionally, self-manipulating altmetric attention scores by quick self-spamming, self-commenting or self-sharing from researchers is also a method to improve the altmetric attention score. In this study, spam, self-promotion, linking, and others all combine up to 18% in total of comments (Table 4) excluding real neutral discussions such as further discussions related to blog post's content, knowledge exchange with other commenters, or neutral communication with blog authors.

The first research question examines whether there is a connection between the volume of online attention and the blog comments, and the answer is no connection. As discussed, blogs are the second highest altmetric attention score generator per post, and this study also examines the association between altmetric attention score and number of comments on science blogs. The correlation gave strong evidence for no connection between altmetric attention score and how many comments on blogs. This means that blogs which mention high altmetric attention score articles do not always receive many comments and vice versa, at least not in this study. As such, the alternative metric correlates with frequency of social attention, for instance, how many shares, mentions, links, citations there are, but it does not reflect deeper attention such as comments, thoughts, or reviews on public posts.

In addition, as discussed in theoretical background, the altmetric attention score reflects the amount of attention, but has no correlation with research quality. Therefore, the quantity of social attention did not define the comments' sentiment which may reflect the quality of research. In fact, Cat Williams (2016), announced on the official site of Altmetric that "*the altmetric score is now the altmetric attention score*". The new name has been suggested to describe the accuracy of altmetric score's meaning. Although this indicator grants attention, the public is not supposed to rely on only altmetric scores, and they need to consider between the positive and negative sides of a research product in their own way before diving into usage. Moreover, with visibility of altmetric score, people can check which social platforms such as Facebook, Twitter, Mendeley, science blogs, etc. contributed the most towards the score of a specific academic article.

However, when observing the details of altmetric attention score, it is scaled qualitatively. News and blogs are the most weighted in generating altmetric score per post due to the perceived complexity of contents. According to Table 1, a scientific article mentioned on New York Times will gain more altmetric attention scores than the same article mentioned on a blog post, and a science blog contributes more than a tweet.

In general, self-manipulating or boosting the still lowers the standard quality of altmetric attention scores. There were surprising findings from different aspects revealed surrounding altmetric attention score to measure attention. Those issues may affect the reliability in evaluating for instance researchers, especially when institutions assess researchers' performance.

5.2 Public impact of science blogging

Science blogging is like a bridge between the scholarly world and the public. Investigating this platform may reveal the influence of academia on people as well as the reverse responses. A study (Paul-Hus et al. 2015) argued that the response to scholarly communication from social media audience is even more reliable than tracking social media metrics. Thus, this study also examines the level of engagement of commenters along with blog authors' influence on the public. As discussed in the theoretical framework, science blogs are normally written and summarized in a "reader-friendly" tone that easily reaches a wider range of audiences than scientific journals. There are generally two kinds of science blog readers, some seek the blogs for interesting scientific news as entertainment purpose, and others read blogs for reviewing research articles as academic purpose.

As reported (Table 4), commenters tend to discuss more about blog posts (about 63% in total of comments) than comment about blog authors or interaction with other commenters. Further discussion for knowledge exchange (group B4) is the hottest subject of commenting about blog posts that makes 34% of the total. Scientific news is the most likely reason for visiting science blogs, and knowledge sharing is probably the biggest motivation in social discussions on this open platform. This answers the second research question that commenters in this study are motivated by knowledge sharing and discussing knowledge beyond the blog posts.

On the other hand, bloggers' influence towards the public is not really evidenced as only 3,4% of responses are about blog authors of the total of comments. As such, the recognition as one of the main factors to measure blog authors' influence (Khan et al. 2015) is considered as low attention in this study. In addition, some comments about blog authors including outgoing links (group A1,D3) also express a low degree of novelty in measuring bloggers' influence. Generally, in this study, most of the core factors for evaluating blog authors' influence have no remarkable impact on readers. This indicates a weak inspiration from blog writers towards the public. However, this does not mean that the quality of blog posts is also low.

Throughout the content analysis in this study, three levels of engagement from public as access, appraise, to apply (Haustein et al. 2015) were recognized. The first level is the way people access scholarly products via reading science blogs, then they express thoughts via discussions on those blogs. The next level is not only appraising something, but also contains criticizing or neutral responses. The average combined sentiment scales indicated that criticizing comments were always slightly more extreme than praising comments in any categories of content analysis. Furthermore, writing a blog post about a scientific product is also a way of applying, which is as the highest level of engagement, and blog visitors in this study comment the most about knowledge collaboration that expresses the strong recognition of appraise (positive or negative) level of engagement from public perspective.

5.3 Gender differences

In this study, the gender differences are measured by examining how commenting is influenced by both the gender of commenters and blog authors. In general, all commenters tend to be rather neutral in discussions, but there were gender disparities in different subjects such as comments about blog posts or comments about blog authors. Women commentators in general are more emotional in language use than men in both a positive and negative way, especially when blog authors are also female, but they tend to be neutral when discussing on male's blog posts. By contrast, men are more positive if blog authors have the same gender, while they tend to criticize on female bloggers' posts.

A previous study (Pedersen & Macafee, 2007) suggested that commenting behavior is one essential factor to identify gender on blogs via their communication style. This work also added that while male tends to express aggressive attitude, sarcasm, and negative manner, females are more likely to be supportive and respectful towards other people. Content analysis in the context of gender revealed that both men and women are interested most in discussions about the blog posts, but they did not have the same manners in those discussions. Men's comments were biased to negative attitude, whereas women commenters tend to be closer to neutral. This finding may agree partly with a study of Atai & Chahkandi (2012) that men's attitudes tend to have aggressive behavior or engage in online discussions in a negative way, while women tend to support conversations and praise others. Nevertheless, an earlier study by Herring (1994) considered that male's manners were derived from their desire of supporting knowledge or debating others' points. Although different findings of motivation in social discussions were discovered, most reviewed studies support the conclusion that men use bad language or lose control in expressing language for their aggressive manners.

Furthermore, the content analysis results (Table 4) also indicated that further knowledge discussion (i.e. commenting on a blog post' content) is the dominant subject in commenting about blog posts. Knowledge exchange may stimulate male's motivation of flaming or debate engagement in scientific topics, that may draw them into posting more negative comment on blog posts. On the other hand, male commenters expressed a different facet in commenting about blog authors. The results (Table 6) displayed a disparity in gender for comments about bloggers; while women commenters had no commendation to bloggers and still keep neutral discussions, compliments were detected in men's comments in the same manner. This finding contrasted with the gender difference in natural emotion and language use. Women tend to show more explicitly emotional expression than men, although this could be in positive or negative way (Thelwall et al. 2009). Another study (Herring, 2000) indicated that women easily tend to express positive manners by saying "*thanks*", appreciate or apologize more, while males tend to hide those kinds of emotional expressions (Thelwall et al. 2009). It has to be noted that the number of comments about blog authors only take more than 3% in over total of 1000 comments, and as communication style may vary regard to gender, therefore, it is inadequate to compare this small finding with previous studies.

Due to the inequality in number of female and male commenters, this study examined the average scores from sentiment analysis on gender to enhance the quality of results. The findings (Table 7) showed that female commenters tend to be more neutral in social discussions if blog posts were written by men, but they showed more emotional expression in both a positive and negative way if the blog authors are female. Thus, female commenters' average combined sentiment on blog posts written by female bloggers is 1,7 as positive and -1,56 as negative, while the similar scales in men's posts are only 1,08 and -0,82. Accordingly, a study about instant messaging dialogues by Lee (2003) found that females produce more emoticons in their messages to support their emotional expression regardless of the receivers' gender, and the natural emotional exposure of women explained their extreme language use. On the other hand, male commenters were bias to praise slightly if bloggers have the same gender, while they tend to show a negative attitude on female bloggers's posts. This probably indicated the different writing styles between male and female bloggers that influence the manners of commenters differently according to their gender.This could be explained by the gender differences in information behavior, especially in textual conversation without knowing the real emotion (Witmer & Katzman, 1997).

From the perspective of blog authors, the number of male blog authors are nearly double compared to female bloggers due to the sample selection. Therefore, the number of received comments have a similar ratio. A previous study (Shema et al., 2012) also discussed blog authorship on ResearchBlogging, showing that men are the main blog contributor in both sole authorship and co-authorship blogs, while the number of female bloggers are far behind. In addition, the profession of blog authors was also investigated along with their gender. Scientific journalist is the main profession of both male and female blog authors. However, blog posts of scientific journalists who supposedly have good writing skills as well as scientific knowledge receive the largest number of negative responses, and the most positive ones are received by bloggers that are faculty members such as professor or researcher regardless bloggers' gender. Nevertheless, sentiment analysis (Table 8) revealed something else. There was no significant sentiment difference from the aspect of blog authors' gender. While professors and researchers receive neutral comments on average, scientific journalists tend to receive neutral responses.

5.4 Limitations

Despite triangulation of methods that produced the findings, several limitations still exist.

First, the sample selection phase filtered inadequate ratio between males' and females' comments due to focusing on attention characteristic regardless gender. Thus, the number of male's comment is five times larger than female's, and this lead to imbalance and unfair bias in analyzing gender differences. In addition, although "reply" comments were already filtered, some extracted comments also have contents of replying that seen to be private conversations. Moreover, another hole in the sample selection are duplicates. The unknown blog authors and unknown commenters might be selected repeatedly creating duplicate data in sample selection because an unknown person may have multiple comments that were extracted to the sample. Therefore, this affected the quality of data, analysis of gender differences as well as content analysis.

Next, the sentiment analysis was also not perfect. Although SentiStrength was applied properly in this study for its functional utility as well as availability of various languages, the method still has obstacles. While SentiStrength is recommended to analyze short texts, there were a bunch of long comments chosen from the sample selection. The summary of comments sometimes does not express fully the original texts' meaning and summarized words may distort sentiment results because the tool detects sentiment strength and produces scores based on the analysis of every single word. Additionally, SentiStrength has limited capabilities to analyze lingos or symbols which are usually found in natural language. This may have skewed the results towards neutral. Moreover, sarcasm as well as ironic words in different contexts of social discussions can cause a misunderstanding for SentiStrength. In general, those causes of inaccuracy in assessing sentiment analysis may have affected the study's outcome.

Moreover, the study was designed to abandon data of unknown commenters as well as unknown bloggers. If included, these could have revealed disparities in appearance style on blogs between males and females. Although with support of sentiment analysis by SentiStrength, the manual codifying of content analysis has been done by the author alone, therefore, a bias in categorizing content analysis may have happened.

5.5 Future research

This study discusses the importance of science blogging as a potential platform to investigate scholarly communication, especially in the context of societal impact. The datasets in this study is based on a Finnish research project which focuses on Finnish science and education. Researchers can forecast the trending of people's interest in Finnish scientific topics, and evaluate how the Finnish research is viewed as well as reviewed by the public.

Additionally, based on the limitations above, other types of studies, especially exploratory studies and quantitative studies could improve their methodology by exploiting all usable data, and by filtering duplicates in sample selection that may affect the research's outcome. Moreover, the database for the sentiment analysis is necessary to enhance, because more and more lingos, new terminology or symbols will be created with the evolution of communication.

Furthermore, the future of altmetrics is still a big question mark. It is necessary to investigate more research on this new subject. Altmetrics' developers may think about anti-boosting-score features such as reviewing posts' authors before generating scores. For example, blog authors or Twitter accounts who have good conduct or have a remarkable number of followers will contribute more to the altmetric attention scores than others. Nowadays, with the domination of user-generated content platforms, massive information and big data created by those users are still messy. Observing beyond the social attention, altmetrics is not yet a standard tool for evaluation of science until the quality side can be accomplished. Many people do not even know what altmetrics is, and it is necessary to optimize the tools and methods before applying them in other areas.

Chapter 6. Conclusion

Science blogging as part of open scholarly communication takes advantage of enduring discussions. As such, this creates a good opportunity to examine deeper into public reach and the public's response to open research products. Altmetrics played an important role in tracing digital information from social platforms but its real conception is still attached with a question mark. Thus, in this study, altmetric attention scores were collected and analyzed along with social discussions to explore its standard function in assessing the impact of public attention. As results, there were no surprising or any reversed findings against previous research. This agreed with the results that altmetrics still need further development to accomplish a standard measurement tool which is able to take into account research quality besides public attention. Nevertheless, this study investigated deeper by examining the relationship of altmetric attention score and number of comments. The outcome showed no connection between altmetric attention score and number of comments.

Moreover, exploring deeper into social discussions' content, the study revealed commenting behavior on science blogging. People seem to care the most about knowledge sharing; they desire to debate, share and discuss further knowledge that support topic of the blog post they read and have commented. This also indicates the motivation of participating in science blogs as well as of extending the discussion. Furthermore, considering gender differences in the context of that communication in this study brings a different look from previous studies. Both female and male tend to be neutral in their commenting behavior. However, examining weighted scales of sentiment in their comments, women remain more emotional in comments, especially if blog authors are the same gender. Despite neutral discussions in general, male commenters tend to criticize more than women.

There is a need to carry out more research in order to introduce transparent and functional altmetrics, and to understand more about differences in communication style between males and females. Open access has been done very well in presenting and updating new knowledge to public, but this is just the beginning of the future ahead and finding a way to assess this source is still a big challenge.

Bibliography

Altmetric Support (2015) *How is the Altmetric attention score calculated?* Available at: https://help.altmetric.com/support/solutions/articles/6000060969-how-is-the-altmetric-score-calculated- (Accessed: 21 June 2016).

Altmetric Support (2016) *Putting the Altmetric attention score in context*. Available at: https://help.altmetric.com/support/solutions/articles/6000060970-putting-the-altmetric-attention-score-in-context (Accessed: 6 July 2016).

Adie, E. (2013) *Gaming altmetrics – Altmetric*. Available at: https://www.altmetric.com/blog/gaming-altmetrics/#comment-jump (Accessed: 22 June 2016).

Akimoto, A. (2011) *Japan, the Twitter nation | the Japan times*. Available at: http://www.japantimes.co.jp/life/2011/05/18/digital/japan-the-twitter-nation/#.V0Ic6JF96Uk (Accessed: 22 May 2016).

Atai, M.R. and Chahkandi, F. (2012) 'Democracy in computer-mediated communication: Gender, communicative style, and amount of participation in professional listservs', *Computers in Human Behavior*, 28(3), pp. 881–888. doi: 10.1016/j.chb.2011.12.007.

Bernal, I. (2013) 'Open access and the changing landscape of research impact indicators: New roles for repositories', *Publications*, 1(2), pp. 56–77. doi: 10.3390/publications1020056.

Brody, L.R. and Hall, J.A. (2014) *Gender and emotion in context*. Available at: https://imshever.files.wordpress.com/2014/09/brodyhall2008.pdf (Accessed: 23 May 2016).

Colledge, L., Phil, D. and James, C. (2015) *5 ways usage metrics can help you see the bigger picture*. Available at: https://www.elsevier.com/connect/5-ways-usage-metrics-can-help-you-see-the-bigger-picture (Accessed: 22 May 2016).

De Roure, D. (2014) 'The future of scholarly communications', *Insights*, 27(3). doi: 10.1629/2048-7754.171.

Desrochers, N., Bowman, T.D., Haustein, S., Mongeon, P., Quan-Haase, A., Paul-Hus, A., Costas, R., Larivière, V., Pecoskie, J. and Tsou, A. (2015) 'Authorship, patents, citations, acknowledgments, Tweets, reader counts and the multifaceted reward system of science', *Proceedings of the Association for Information Science and Technology*, 52(1), pp. 1–4. doi: 10.1002/pra2.2015.145052010013.

Didegah, F. and Thelwall, M. (2012) *Predictive indicators of research citation impact in S&T fields: A case study of Nanoscience and Nanotechnology*. Available at: http://s3.amazonaws.com/academia.edu.documents/30526736/didegah_predictive_236.pdf?AWSAccessKeyId=AKIAJ56TQJRTWSMTNPEA&Expires=1463952204&Signature=VdLM1rNcBeUEhJMHuey%2BruDEIh0%3D&response-content-disposition=inline%3B%20filename%3DPredictive_Indicators_of_Research_Citati.pdf (Accessed: 22 May 2016).

Elsevier (2010) *Peer review policy and publication times*. Available at: http://www.journals.elsevier.com/social-science-and-medicine/policies/peer-review-policy-and-publication-times (Accessed: 20 June 2016).

Eysenbach, G. (2011) 'Can Tweets predict citations? Metrics of social impact based on Twitter and correlation with traditional metrics of scientific impact', *Journal of Medical Internet Research*, 13(4), p. 123. doi: 10.2196/jmir.2012.

Finlay, S.C., Hank, C., Sugimoto, C.R. and Johnson, M. (2013) 'The structure of the Biblioblogosphere: An examination of the linking practices of institutional and personal library Blogs', *Journal of Web Librarianship*, 7(1), pp. 20–36. doi: 10.1080/19322909.2012.714305.

Gu, F. and Widén-Wulff, G. (2011) 'Scholarly communication and possible changes in the context of social media', *The Electronic Library*, 29(6), pp. 762–776. doi: 10.1108/02640471111187999.

Guadagno, R.E., Eno, C.A. and Okdie, B.M. (2010) 'Personal Blogging', *Handbook of Research on Social Interaction Technologies and Collaboration Software*, , pp. 292–301. doi: 10.4018/978-1-60566-368-5.ch026.

Guiller, J. and Durndell, A. (2006a) 'Students' linguistic behaviour in online discussion groups: Does gender matter?', *Computers in Human Behavior*, 23(5), pp. 2240–2255. doi: 10.1016/j.chb.2006.03.004.

Guiller, J. and Durndell, A. (2006b) 'Students' linguistic behaviour in online discussion groups: Does gender matter?', *Computers in Human Behavior*, 23(5), pp. 2240–2255. doi: 10.1016/j.chb.2006.03.004.

Haustein, S., Bowman, T.D. and Costas, R. (2015) *Interpreting ' altmetrics ': Viewing acts on social media through the lens of citation and social theories*. Available at: https://arxiv.org/ftp/arxiv/papers/1502/1502.05701.pdf (Accessed: 14 May 2016).

Herring, S. (1994a) *Gender differences in computer-mediated communication: Bringing faimilar baggage to the new frontier*. Available at: http://cpsr.org/prevsite/cpsr/gender/herring.txt/ (Accessed: 2 June 2016).

Herring, S. (1994b) *Politeness in computer culture: Why women thank and men flame*. Available at: http://ella.slis.indiana.edu/~herring/politeness.1994.pdf (Accessed: 2 June 2016).

Herring, S.C. (2000) *Gender differences in CMC: Findings and implications*. Available at: http://cpsr.org/issues/womenintech/herring/ (Accessed: 2 June 2016).

Herring, S.C. and Paolillo, J.C. (2006) 'Gender and genre variation in weblogs', *Journal of Sociolinguistics*, 10(4), pp. 439–459. doi: 10.1111/j.1467-9841.2006.00287.x.

Hollenbaugh, E.E. (2010) 'Personal journal bloggers: Profiles of disclosiveness', *Computers in Human Behavior*, 26(6), pp. 1657–1666. doi: 10.1016/j.chb.2010.06.014.

Holmberg, K., Didegah, F., Bowman, S., Bowman, T.D. and Kortelainen, T. (2015) 'Measuring the societal impact of open science – presentation of a research project', *Informaatiotutkimus*, 34(4).

Holmberg, K. and Thelwall, M. (2014) 'Disciplinary differences in Twitter scholarly communication', *Scientometrics*, 101(2), pp. 1027–1042. doi: 10.1007/s11192-014-1229-3.

Huffaker, D.A. and Calvert, S.L. (2005) 'Gender, identity, and language use in teenage Blogs', *Journal of Computer-Mediated Communication*, 10(2), pp. 00–00. doi: 10.1111/j.1083-6101.2005.tb00238.x.

Khan, H.U., Daud, A. and Malik, T.A. (2015) 'MIIB: A metric to identify top influential Bloggers in a community', *PLOS ONE*, 10(9), p. e0138359. doi: 10.1371/journal.pone.0138359.

Klain-Gabbay, L. and Shoham, S. (2016) 'Scholarly communication and academic librarians', *Library & Information Science Research*, 38(2), pp. 170–179. doi: 10.1016/j.lisr.2016.04.004.

Konkiel, S., Sugimoto, C.R. and Williams, S. (2016) *The use of Altmetrics in promotion and tenure*. Available at: http://er.educause.edu/articles/2016/3/the-use-of-altmetrics-in-promotion-and-tenure (Accessed: 22 May 2016).

Kousha, K., Thelwall, M. and Rezaie, S. (2010) 'Using the web for research evaluation: The integrated online impact indicator', *Journal of Informetrics*, 4(1), pp. 124–135. doi: 10.1016/j.joi.2009.10.003.

Larivière, V., Sugimoto, C.R. and Bergeron, P. (2013) 'In their own image? A comparison of doctoral students“ and faculty members” referencing behavior', *J Am Soc Inf Sci Tec*, 64(5), pp. 1045–1054. doi: 10.1002/asi.22797.

Lee, C. (2003) *How does instant messaging affect interaction between the genders?* Available at: http://web.stanford.edu/class/pwr3-25/group2/pdfs/IM_Genders.pdf (Accessed: 2 June 2016).

Mas-Bleda, A., Thelwall, M., Kousha, K. and Aguillo, I.F. (2014) 'Do highly cited researchers successfully use the social web?', *Scientometrics*, 101(1), pp. 337–356. doi: 10.1007/s11192-014-1345-0.

Mohammadi, E. and Thelwall, M. (2013) 'Reference tools', *Scientometrics*, 97(2), pp. 383–395. doi: 10.1007/s11192-013-0993-9.

Mohammadi, E., Thelwall, M., Haustein, S. and Larivière, V. (2015) 'Who reads research articles? An altmetrics analysis of Mendeley user categories', *Journal of the Association for Information Science and Technology*, 66(9), pp. 1832–1846. doi: 10.1002/asi.23286.

Mohammadi, E., Thelwall, M. and Kousha, K. (2015) 'Can Mendeley bookmarks reflect readership? A survey of user motivations', *Journal of the Association for Information Science and Technology*, 67(5), pp. 1198–1209. doi: 10.1002/asi.23477.

Paul-Hus, A., Sugimoto, C.R., Haustein, S. and Larivière, V. (2015) 'Is there a gender gap in social media metrics?', .

Pedersen, S. and Macafee, C. (2007) 'Gender differences in British Blogging', *Journal of Computer-Mediated Communication*, 12(4), pp. 1472–1492. doi: 10.1111/j.1083-6101.2007.00382.x.

Priem, J., Taraborelli, D., Groth, P. and Neylon, C. (2010) *Altmetrics: A manifesto*. Available at: http://altmetrics.org/manifesto/ (Accessed: 22 May 2016).

Primack, R.B., Arceiz, A.C. and Koh, L.P. (2015) *Is peer review just a crapshoot?* Available at: https://www.elsevier.com/connect/is-peer-review-just-a-crapshoot (Accessed: 20 June 2016).

Rowlands, I., Nicholas, D., Russell, B., Canty, N. and Watkinson, A. (2011) 'Social media use in the research workflow', *Learned Publishing*, 24(3), pp. 183–195. doi: 10.1087/20110306.

Shema, H., Bar-Ilan, J. and Thelwall, M. (2012) 'Research Blogs and the discussion of scholarly information', *PLoS ONE*, 7(5), p. e35869. doi: 10.1371/journal.pone.0035869.

Shema, Hadas, Bar-Ilan, J. and Thelwall, M. (2014) 'Do blog citations correlate with a higher number of future citations? Research blogs as a potential source for alternative metrics', *Journal of the Association for Information Science and Technology*, 65(5), pp. 1018–1027. doi: 10.1002/asi.23037.

Shema, H., Bar-Ilan, J. and Thelwall, M. (2014) Available at: https://www.researchtrends.com/issue-37-june-2014/scholarly-blogs-are-a-promising-altmetric-source/ (Accessed: 23 May 2016).

Shema, H., Bar-Ilan, J. and Thelwall, M. (2015) 'How is research blogged? A content analysis approach', *Journal of the Association for Information Science and Technology*, 66(6), pp. 1136–1149. doi: 10.1002/asi.23239.

Sud, P. and Thelwall, M. (2014) 'No citation advantage for monograph-based collaborations?', *Journal of Informetrics*, 8(1), pp. 276–283. doi: 10.1016/j.joi.2013.12.008.

Sugimoto, C. (2015) *Discover the future of research: 'Attentio is not impact' and other challenges for altmetrics*. Available at: https://hub.wiley.com/community/exchanges/discover/blog/2015/06/23/attention-is-not-impact-and-other-challenges-for-altmetrics#comment-2097762855 (Accessed: 22 May 2016).

Sussman, N.M. and Tyson, D.H. (2000) 'Sex and power: Gender differences in computer-mediated interactions', *Computers in Human Behavior*, 16(4), pp. 381–394. doi: 10.1016/s0747-5632(00)00020-0.

Tenn⓪ntJon (2016) 'Current @altmetric score for the #Creatorgate paper is over 1000!! Https://t.co/Q8Bj8AQvkl', *Twitter*, 3 March. Available at: https://twitter.com/protohedgehog/status/705503086312353792 (Accessed: 22 June 2016).

Thelwall, M. and Buckley, K. (2012) Available at: http://www.scit.wlv.ac.uk/~cm1993/papers/SentiStrength3Preprintx.pdf (Accessed: 18 July 2016).

Thelwall, M., Haustein, S., Larivière, V. and Sugimoto, C.R. (2013) 'Do Altmetrics work? Twitter and Ten other social web services', 8(5).

Thelwall, M., Wilkinson, D. and Uppal, S. (2009) 'Data mining emotion in social network communication: Gender differences in MySpace', *Journal of the American Society for Information Science and Technology*, 61(1), pp. 190–199. doi: 10.1002/asi.21180.

Thelwall, M. and Wilson, P. (2014) Available at: https://arxiv.org/ftp/arxiv/papers/1510/1510.08877.pdf (Accessed: 22 May 2016).

Uzunoğlu, E. and Kip, S.M. (2014) 'Brand communication through digital influencers: Leveraging blogger engagement', *International Journal of Information Management*, 34(5), pp. 592–602. doi: 10.1016/j.ijinfomgt.2014.04.007.

Williams, C. (2016) *The Altmetric score is now the Altmetric attention score*. Available at: https://www.altmetric.com/blog/the-altmetric-score-is-now-the-altmetric-attention-score/ (Accessed: 22 October 2016).

Witmer, D.F. and Katzman, S.L. (1997) 'On-line smiles: Does gender make a difference in the use of graphic accents?', *Journal of Computer-Mediated Communication*, 2(4), pp. 0–0. doi: 10.1111/j.1083-6101.1997.tb00192.x.