Relationships Between Traditional Metrics and Altmerics: A Case Analysis of PLoS*

Geleneksel Metrikler ile Altmetri Metrikleri Arasındaki İlişkiler: PLoS Analizi Örneği

*Müge AKBULUT****

Abstract

The possibility of creating and easily distributing something via a digital platform creates an enormous material volume. With the increase of scientific publications it becomes harder for scholars to choose the most related and significant resources from the others (Henning & Gunn, 2012). Bibliometrics measures the impact of scholars between each other, but the impact of articles across the entire web is not limited by this. On the other hand, altmetrics measures the impact of articles from the web, alongside bibliometric data. The aim of this study is to investigate the probable relationships between traditional metrics and altmetrics by analysing the PLoS Article-Level Metric (ALM) dataset.

***Keywords:** Bibliometrics, altmetrics, Article Level Metric (ALM)*

Öz

Dijital ortamda herhangi bir şey yaratma ve kolayca herkese sunma olanağı devasa bir materyal hacmi yaratmaktadır. Bilimsel yayın sayısı arttıkça da akademisyenler için en ilgili ve dikkate değer kaynakları diğerleri arasında seçmek zorlaşmaktadır (Henning ve Gunn, 2012). Bibliyometrik ölçümler akademik dergilerde akademisyenlerin birbirlerine olan etkilerini ölçmektedir. Fakat makalelerin tüm Web üzerindeki etkisi bununla sınırlı değildir. Altmetri (altmetric) ise mevcut bibliyometrik ölçümlerin yanı sıra makalenin tüm Web üzerindeki etkisini içermektedir. Bu çalışmanın amacı; PLoS Article Level Metric (ALM) veri setinin analiz edilerek geleneksel metrikler ile altmetri metrikleri arasındaki olası ilişkileri incelemektir.

***Anahtar sözcükler:** Bilbiyometri, altmetri, Article Level Metric (ALM)*

Introduction

The number of papers published in scientific journals is gradually increasing in parallel with developments in science and technology. Scholars engaged in research need to follow more journals and papers every day. A web environment which is globally accessible to everyone leads to benefits from a shared intelligence obtained through the contributions of users. In this environment, in which users gradually become participants, the media allows a more flexible and open system for cooperation and sharing instead of focusing only on the call to consume (O'Reilly, 2005; Shirky, 2008). At some point, mass amateurisation will cause a filtration problem which is bigger than that found in traditional environments. Therefore, solutions used before might be inadequate (Shirky, 2002). Scholars are making use of filtration to follow academic literature. However, traditional filtration methods are becoming gradually more dysfunctional as the environment continues to diversify. Traditional performance measurement systems are desirably adequate in comparison with current technology. Now, other ways of handling this problem should be focused and a number of different sources should be taken into account. The recent increase in online academic indexing systems has enabled new filtration opportunities. The introduction of altmetrics is significant, as it could rapidly impact on academic filtration within the digital ecosystem.

Scholars carry out their daily tasks on Web. According to some studies, there are more than 40 million papers listed on some online softwares like Zotero and Mendeley (more universal than PubMed). Today, conversations about an invention are pondered and discussed on blogs and within social media (Mollett et al., 2011). Nearly one third of scholars are also Twitter users (Priem & Costello, 2010). Besides, the number of scientific citations attributed to articles on Twitter is more than 58,000. Researchers are following other researchers whom they think to be important opinion leaders via Twitter instead of reading a bulk of pages by various authors in a peer-reviewed journal.

All these interactions are reflected in scientific communication processes. Articles which are dog-eared and have not been quoted can be found and recounted in some online environments like Mendeley, Zotero and CiteULike (Howard, 2012).

Individuals have an opportunity to reach all kinds of online environments in which they can share a variety of articles, images, videos and so on. They can make use of such opportunities for research and cooperation. Thus, these environments may provide a useful measure of the impact of scientific research, or performance.

Scientific communication has developed along with digital technologies. As communications have moved to an electronic environment, not only articles but also a number of new structures (data sets, analysis, reference managers, blogs, social networks, social marking, discussion lists and so on), have begun to appear. Hence, the homogeneity of channels has decreased while the diversification of channels has increased. Informal communication sparks debates since it is significant in scientific

communication. Researchers share their own studies with each other within online environments, potentially changing the direction of the research. More diverse environments mean more debates and discussions. If there are many ideas, then there are many individuals who discuss these with each other. Informal systems, which have been at the centre of research communications in the past, have now migrated to and are being followed online.

There are a great many more amateurs than professionals, as consumers also become producers who can easily communicate with each other in online environments (Surowiecki, 2004). This activity could be utilized to measure the impact of scientific research, or performance.

This study will focus on correlations between altmetrics and traditional metrics. It also discuss how altmetrics could affect traditional filters and whether it could be employed for forecasting.

Literature Review

As a result of the increasing opportunities provided by digital environments, researchers have begun to make use of different methods to measure scientific performance impact. One of the most significant of these methods is webometric. Webometric, first put forward by Almind and Ingwersen, applies infometric methods on the World Wide Web (1997, p.404). Webometrics was further used to draw conclusions about the authenticity of the research and academic papers digitally available. In a research paper written by Owen Thomas and Peter Willet during his webometric analysis of library and Information Science department (LIS), their study stated that using citations as a measure of the excellence of an academic paper was not a possible solution (Thomas and Willet, 2000). There have been a number of studies about webometrics usage and usage records of online article access. In Bollen and friends' study (2009) they produced a component analysis of 39 Scientific Impact Measures; they concluded their study by stating that a single metric cannot prove to be an accurate measure for the scientific authenticity of data. However, webometrics data are only periodically manually collected periodically, as they are limited by automatic mining. This method is not functional to a large extent because publishers are not willing to open their sources for an extensive usage (Haustein and Siebenlist, 2009). Therefore, usage metrics are only automatically collected on very small scale.

According to the studies of Procter and friends, 80% of academicians are estimated to have social media user accounts (2010). There are also some studies about how much altmetrics data we have and how they are distributed. Altmetric is basically a subset of webometric and scientometric. (Priem, Groth and Taraborelli, 2012). One of such studies was carried out by in which they found some amazing results which included that a particular altmetric source varied over a particular community and

period of time.(Priem, Piwowar and Hemminger, 2012) In studies in which comment systems based on journals and 'rapid impacts' are focused, there are many asymmetric distributions. In a study of Schriger and others, prepress articles in the repository and rates of being tweeted are compared. According to the results of this study, it has been revealed that sample articles from the arXiv preprint repository were tweeted at the rate of 95% (Schriger, Chehrazi, Merchant, and Altman, 2011). Wardles points out those citations in Wikipedia do not coincide with the Journal Citation Report, although there are a few more citations in more effective journals (Wardle, 2010).

Altmetric data has opened new avenues for filtering. Gone are the days when peer-review systems used to one of the best techniques for filtering data. Altmetric has revolutionized the way we filter data because it not only filters data according to citations but also takes into account the impact of that article or paper another most significant benefit of altmetrics is that it's speed has enabled people to produce real-time collaborative filtering and recommendations (Priem, Taraborelli, Groth and Neylon, 2010). The revolution of altmetric is not limited to just filtering. It has brought significant changes to forecasting as well. According to the studies done, we are getting closer to using altmetric for forecasting purposes. Since they are extremely beneficial for assessing impact of scholarly output, that idea is being used to determine the popularity of fields and how they are increasing and this data is being used to forecast future trends (Priem, Taraborelli, Groth and Neylon, 2010).

We cannot forget the numerous advantages of altmetric but also we cannot ignore some of the disadvantages or limitations as well. Data quality is one of the main disadvantages where altmetric is concerned and due to that it leads to some limitation for the altmetric. The most common being the bias data. Studies show that by measuring the impact of the scholarly output online, a huge chunk of the population who reads the papers from other sources are being neglected. As majority of the population who reads the resources online are of a certain age, thus the altmetric impact is biased in this way. (Bornmann, 2014)

Traditional Filters and Almetrics

Counting the number of citations, peer-review and impact factor could be listed out as traditional citation methods which are often used in order to measure the scientific research performance.

Counting Citation Numbers

While evaluating the potential impact of a scientific study, the most commonly used method is counting citation numbers (Kear & Colbert-Lewis, 2011). Such methods of citation counting are necessary and beneficial; however they are not enough to measure the scientific impact. Even H-index, which helps count how many times an

article has been cited, is obtained slower than peer review. Counting the first citation of an academic study takes years (Al, 2008). By focusing on the citation numbers, only the academic use of scientific studies could be counted. Yet, it should be considered that a great many people who are not scholars also benefit from such kinds of academic studies. In this point, the academic impacts of the scientific studies cannot be centered because such persons do not write and make use of citations in scientific journals.

Another disadvantage of citation counting is related with the fact that academic citations could be analyzed much later that the actual time. In this context, almetrics ensures counting citation numbers much faster.

It is also possible to trace the impacts of the academic studies on the other fields outside of academy through almetrics. In this way, the impacts of the studies which have not been cited or reviewed could also be measured.

Peer-Review

Performance indicators, which are created as a consequence of experts opinions while reviewing an academic study, are also often used (Aksnes & Taxt 2004; Smith & Eysenck 2002). In an academic study in which Reale, Barbara and Costantini (2007:224) analyzed the harmony between bibliometric indicators and expert opinions, it has been revealed that the impact factor values of the journals where articles are published are not independent of peer review and there is a relationship between these two variables.

Impact Factor

Impact factor is calculated by dividing the average number of citations to the articles published in the last two years to the number of articles published in the last two years (Garfield 1994). Unlike impact factor, almetrics reveal the own impact of an article. However, impact factor is the value of the journal in which an article is published. Impact factor is open to manipulations whereas almetrics has the potential of being more systematic and reliable. Almetrics could detect the manipulative activities and correct them by making use of the multiplicity and statistical force of large data.

The filters of Future: Almetrics

Scientific studies are revealed in a variety of ways through digital media. For example, by employing executable papers, researchers could reanalyze the data which are used during research (Nowakowski et al., 2011). Besides, instead of citing the whole article, only a certain part of that article could be cited by means of nanopublication or semantic publishing. Interpretations and explanations could be added as well.

As almetrics is diverse on its own, it is highly effective in terms of measuring the impact in this diverse almetrics. Through API, data could be collected in a short time.

There are algorithms not just for collecting data but also for interpreting these data. It is also possible to analyze the content beyond semantic by making use of almetrics (username, timestamp, tags and so on).

Conclusion

In order to understand whether almetrics values could measure the real impact or not, it is essential to carry out more academic studies by making use of comprehensive and large data sets. It is substantially obvious that almetrics brought a new dimension to measuring impact values. Although it is a more flexible system in comparison with peer-reviewed journal system, it is not employed for measuring scientific performance. However, when the crisis of current filtering systems and the evolution of academic communication are taken into account, almetrics needs to be analyzed more because of its speed, value and flexibility. More comprehensive studies should be carried out in order to find out whether almetrics relatively reveals the real impact or not. Application developers should develop systems through which the semantic data of almetrics could be interpreted and the questions of 'how' and 'why' could be answered except for number information. Such systems should also have a developed structure by means of which the abnormalities of these systems could be focused.

Method

The objective of this study is to seek any correlations between traditional and altmetric metrics by analysing PLoS Article-Level Metric (ALM) data sets. This study also aims to test whether there are statistically significant correlations among articles' use frequency and their impact factors, total citation numbers and half-lives.

This paper addresses the following research questions:

◊   Is the impact of an article on the Internet sensitive enough to forecast an increase in citations?

◊   Are impact factors and total citation numbers of the journals in which there are articles having high impact (altmetrics values) values higher?

◊   Do 18 metric statistics which designate the altmetrics values of the articles show changes over the years?

Essential data sets were downloaded from the PLoS Article-Level Metric (ALM) web page (http://article-level-metrics.plos.org/plos-alm-data/) on January 2013. Data sets included 78386 articles which were published in eight PLoS journals (PLOSBiology, PLOSClinicalTrials, PLOS Computational Biology, PLOSGenetics, PLOSMedicine, PLOS Neglected Tropical Diseases, PLOS ONE ve PLOSPathogens) between 2003-2013. Citation numbers were manually extracted from Web of Science (WoS) for all articles matching the journal "PLoS *". Detailed Web of Science (WoS) citation information was exported for the

22k PLoS articles, 500 articles at a time (maximum permitted by ISI website). For impact factor, JCR (Journal Citation Report) data sets were united and transferred to Numbers, SPSS and Tableau programs to evaluate them after making necessary arrangements. Citation numbers of the articles have been gained from Web of Science (WoS) database published by Institute for Scientific Information in USA. Journal Citation Reports (JCR) published by ISI has been used to find out impact factors of the articles along with citation numbers, self-citation rates and so on.

Findings

After an article has been published, the probability of an article, which is in the first quarter in tweeter in a month, to be in the first quarter in citations is 17 times more. Therefore, forecasting element could be considered as one area of utilization for almetrics. The first objective is not to guess citation numbers. If correlation numbers are not high, it does not mean that there is a failure. It is obvious that these metrics bring a new dimension to measuring impact values.

The correlations of different metrics including usage statistics of articles (views in HTML, PDF downloads, XML downloads), citations statistics (PubMed, CrossRef, Scopus), social share (Facebook, twitter), bookmarkings (CiteULike, Connotea, Mendeley) have been analyzed. As could be seen in Figure 1, there is a large set which is relatively revealed through high correlation coefficients in bookmarking, citation metrics and social share metrics. There is a statistically significant relationship between the number of articles downloaded as PDF and the number of HTML (Hyper Text Markup Language) views (Pearson's r=0,854, p < 0,01).

	Bookmarking	Social Share	Citation	Usage
Usage	0.3	0.3	0.5	1.0
Citation	0.7	0.7	1.0	0.5
Social Share	0.3	1.0	0.7	0.3
Bookmarking	1.0	0.3	0.7	0.3

Figure 1. Correlation Between Derived Factors

HTML view of a document indicates the popularity of that document. Such an analysis has also been repeated for PDF download and XML download statistics through which how many times a document have been read could be revealed. 'decrease model' are indicated for these three access types. However, XML downloads do not show similar characteristics.

View statistics have been used for information distribution. View statistics of 5000 articles which were published at least one year ago and the almetrics values of which are the highest have been focused. These documents are found in 6 different PLoS journals (PLoS Biology, PLoS Computational Biology, PLoS Genetics, PLoS Medicine, PLoS One ve PLoS Pathogens). The view numbers of an old document is lower. Such a decrease is much faster compared to the periods following the first and second months after that article has been published.
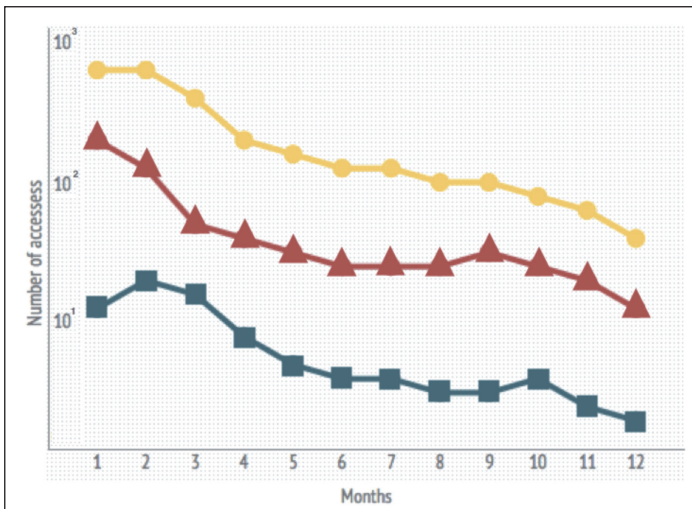

Figure 2. The Decay Patterns of Web Accesses

In Figure 2, access numbers for each document have been analyzed along with average HTML access numbers. In the first three months when 5000 documents the almetrics values of which are the highest have been published, the cumulative numbers of their HTML views have also been focused. What these findings mean should be discussed further.

Results and Recommendations

Expert supervision could be carried out through crowdsourcing. Instead of waiting for peer review of a scientific article, such a document could be reviewed by focusing on

its impact, interpretations, discussions and bookmarks. PLoS ONE, BMC Research Notes and BMJ Open make use of this method to quicken peer review process.

Another problem of almetrics is that there are many users of almetrics. For instance, after an article is published, writer's father or relatives might like that article and this could cause a kind of manipulation while evaluating the article. In order to prevent such a problem, there are a number of ways. In his study, Nelson (2011) suggests that different values should be benefited to avoid this problem. Considering that an article is cited much in Mendeley while it is not shared or liked much in Facebook, it could be asserted that the scientific value of such an article is high. Nelson suggests that Facebook should be taken into account as a negative value.

There are also some studies which shows that Almetrics could be deceived Taraborelli, 2008, Neylon and Wu, 2009; Priem and Hemminger, 2010; Groth and Gurney, 2010). Emilio Delgado López-Cózar et al. (2012) carried out a study in which it was indicated that some indicators like citation, like, tweet in web tools were easy to manipulate. It is also possible to come across with some academic studies which point out that it is not so easy to deceive Almetrics since it is based on big data.

The speed of almetrics enables to create filtering systems based on collective work and real-time tips. Instead of being a member of dozens of pages, a user can simply obtain information about the most significant study of the week. In particular, when blogs and pre-print servers are connected, it might be more beneficial because it narrows down the communication loop from years to days.

Acknowledgements

References

Aksnes, D. W., & Taxt, R. E. (2004). Peer reviews and bibliometric indicators: a comparative study at a Norwegian university. Research evaluation, 13(1), 33-41. Retrieved January 13, 2015 from http://rev.oxfordjournals.org/content/13/1/33.full.pdf.

Al, U. (2008). Türkiye'nin Bilimsel Yayın Politikası: Atıf Dizinlerine Dayalı Bibliyometrik Bir Yaklaşım. Ankara: Hacettepe Üniversitesi, June 2008. Retrieved January 13, 2015 from http://yunus.hacettepe.edu.tr/~umutal/publications/dissertation/c2.pdf.

Almind, T. & Ingwersen, P. (1997). Informetric analysis on the World Wide Web: Methodological approaches to 'webometrics'. Journal of Documentation, 53(4). pp. 404-426. Retrieved January 13, 2015 from http://comminfo.rutgers.edu/~muresan/IR/Docs/Articles/jdocAlmind1997.pdf

Bollen, J., Van de Sompel, H., Hagberg, A. & Chute, R. (2009). A principal component analysis of 39 scientific impact measures. PLoS ONE 4. doi:10.1371/journal.pone.0006022. Retrieved January 13, 2015 from http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0006022.

Bornmann, L. (2014). Do altmetrics point to the broader impact of research? An overview of benefits and disadvantages of alt metrics. Journal of Informatics, 8, 895-903.

Delgado López-Cózar, E., Robinson-García, N., & Torres-Salinas, D. (2012). Manipulating Google Scholar Citations and Google Scholar Metrics: simple, easy and tempting. arXiv preprint arXiv:1212.0638. Retrieved January 13, 2015 from http://arxiv.org/abs/1212.0638.

Garfield, E (20 June 1994). The Thomson Reuters Impact Factor. Thomson Reuters. Retrieved January 13, 2015 from http://wokinfo.com/essays/impact-factor/.

Groth, P & Gurney, T. (2010). Studying Scientific Discourse on the Web using Bibliometrics: A Chemistry Blogging Case Study. In: Proceedings of the WebSci10: Extending the Frontiers of Society On-Line, April 26-27th, 2010, Raleigh, NC: US. . Retrieved January 13, 2015 from http://wiki.few.vu.nl/sms/images/9/9c/Websci10-FINAL-29-4-2010f.pdf

Haustein, S. & Siebenlist, T. (2011). Applying social bookmarking data to evaluate journal usage, Journal of Informetrics, 5 (3), 446-457, ISSN 1751-1577, 10.1016/j.joi.2011.04.002. Retrieved January 13, 2015 from http://www.sciencedirect.com/science/article/pii/S1751157711000393.

Henning V. & Gunn W. (2012). "Impact factor: Researchers should define the metrics that matter to them.". Retrieved January 13, 2015 from http://www.guardian.co.uk/higher-education-network/blog/2012/sep/06/mendeley-altmetricss-open-access-publishing.

Howard, J. (March 12, 2012). Tracking scholarly influence beyond the impact factor. Chronicle of Higher Education. Retrieved January 13, 2015 from http://chronicle.com/blogs/wiredcampus/tracking-scholarly-influence-beyond-the-impact-factor/35565.

Kear, R. & Colbert-Lewis, D. (2011). Citation searching and bibliometric measures. College & Research Libraries News, 72 (8), 470–474. Retrieved January 13, 2015 from http://crln.acrl.org/content/72/8/470.full.pdf+html.

Mollett, A., Moran, D. & Dunleavy, P. (2011). Using Twitter in university research, teaching and impact activities: A guide for academics and researchers. Retrieved January 13, 2015 from http://blogs.lse.ac.uk/impactofsocialsciences/files/2011/11/Published-Twitter_Guide_Sept_2011.pdf

Neylon C, Wu S (2009). Article-Level Metrics and the Evolution of Scientific Impact. PLoS Biol 7(11): e1000242. doi:10.1371/journal.pbio.1000242. Retrieved January 13, 2015 from http://www.plosbiology.org/article/info:doi/10.1371/journal.pbio.1000242.

Nowakowski, P. et al. (2011). The Collage Authoring Environment. Procedia Computer Science, 4, 608-617. Retrieved January 13, 2015 from http://www.sciencedirect.com/science/article/pii/S1877050911001220.

O'Reilly, T. (2005). What Is Web 2.0: Design patterns and business models for the next generation of software. Retrieved January 13, 2013 from http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html?page=1.

Priem, J. & Costello, K (2010). How and why scholars cite on twittter?. Retrieved January 13, 2015 from https://docs.google.com/presentation/d/15J41q9EzK3CiMMoIPKY545ACndORpa6Wtj7 5v1YjrzQ/present?pli=1&ueb=true#slide=id.i0.

Priem, J. & Hemminger, B. (2010). Scientometrics 2.0: Toward new metrics of scholarly impact on the social Web. First Monday 15. Retrieved January 13, 2013 from http://firstmonday.org/ htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2874/2570.

Priem, J., Piwowar, H. A., & Hemminger, B. M. (2012). Altmetrics in the wild: Using social media to explore scholarly impact. arXiv preprint arXiv:1203.4745.

Priem, J., Taraborelli, D., Groth, P., & Neylon, C. (2010). Altmetrics: A manifesto.

Thomas, O., & Willett, P. (2000). Webometric analysis of departments of librarianship and information science. Journal of information science, 26(6), 421-428.

Procter, R., Williams, R., Stewart, J., Poschen, M., Snee, H., Voss, A. & Asgari-Targhi, M. (2010). Adoption and use of Web 2.0 in scholarly communications. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 368: 4039 –4056. doi:10.1098/ rsta.2010.0155. Retrieved January 13, 2015 from http://rsta.royalsocietypublishing.org/ content/368/1926/4039.short#cited-by.

Reale, E., Barbara, A. & Costantini, A. (2007). Peer review for the evaluation of academic research: lessons from the Italian experience. Research Policy, 16(3),216-228.

Schriger, D., Chehrazi, A., Merchant, R., & Altman, D. (2011). Use of the Internet by Print Medical Journals in 2003 to 2009: A Longitudinal Observational Study. Annals of Emergency Medicine In Press, Retrieved January 13, 2015 http://www.sciencedirect.com.libproxy.lib.unc.edu/ science/article/B6WB0-51S6FN7-1/2/2edccd79138b204deee17536d13a2eae.

Shirky, C. (2002). Weblogs and the Mass Amateurization of Publishing. Retrieved January 13, 2015 from http://www.shirky.com/writings/herecomeseverybody/weblogs_publishing.html.

Shirky, C. (2008). Here comes everybody, the power of organizing without organizations. New York, NY: Penguin Group USA.

Smith, A. & Eysenck, M. (2002). The correlation between RAE ratings and citation counts in psychology. Retrieved January 13, 2015 from http://cogprints.org/2749/1/citations.pdf.

Surowiecki, J. (2004). The Wisdom of Crowds: Why the Many are Smarter than the Few. Retrieved January 13, 2015 from http://www.diplomacy.edu/resources/books/reviews/wisdom-crowds-why-many-are-smarter-few.

Taraborelli, D. (2008). Soft peer review: social software and distributed scientific evaluation. In: Hassanaly, P. andRamrajsingh, A. and Randall, D. and Salembier, P. and Tixier, M., (eds.) Proceedings of the 8th International Conference on the Design of Cooperative Systems, Carry-le-Rouet, 20-23 May 2008. (pp. pp. 99-110). Institut d'Etudes Politiques d'Aix-en-Provence: Aix-en-Provence, France. Retrieved January 13, 2015 http://eprints.ucl.ac.uk/8279/.

Wardle, D. (2010). Do "Faculty of 1000" (F1000) ratings of ecological publications serve as reasonable predictors of their future impact? Ideas Eco Evo 3. Retrieved January 13, 2015 from http://library.queensu.ca/ojs/index.php/IEE/article/view/2379.