Are there biases in decisions to tweet on scientific papers? A plea for conducting an experimental *Twitter* study. Technical note

Lutz Bornmann; Robin Haunschild; Alexander Tekles

Lutz Bornmann ✉
https://orcid.org/0000-0003-0810-7091

Administrative Headquarters of the Max Planck Society, Germany
bornmann@gv.mpg.de

Max Planck Institute for Solid State Research
l.bornmann@fkf.mpg.de

Robin Haunschild
https://orcid.org/0000-0001-7025-7256

Max Planck Institute for Solid State Research
Heisenbergstraße, 1
70569 Stuttgart, Germany
r.haunschild@fkf.mpg.de

Alexander Tekles
https://orcid.org/0000-0001-8765-9331

Ludwig-Maximilians-Universität Munich, Department of Sociology
Konradstr., 6. 80801 Munich, Germany

Administrative Headquarters of the Max Planck Society
alexander.tekles.extern@gv.mpg.de

Abstract

Tweets are used as alternative metrics (altmetrics) to measure the impact or attention of research. However, *Twitter* data can only be used for research evaluation purposes, if biases do not influence tweet decisions on papers. The existence of biases can only be reasonably investigated using an experimental design with controlled (marginal) manipulations. In this comment, we propose to undertake an experimental approach to study the decision of scientists to 'tweet' on a paper. We describe the design of a study that might allow the experimental investigation of tweet decisions including randomized variations and theoretically derived mechanisms for explaining the empirical results. The described study design should be adaptable to other social media platforms (e.g., *Facebook* or *ResearchGate*). This comment is intended to be a plea for using an experimental design to investigate biases in tweet decisions. It is an advantage of tweets –in contrast to citations– that an experimental approach can be applied to investigate the decision of scientists to communicate on papers.

Keywords

Altmetrics; *Twitter*; Experiments; Gender differences; Retweet preferences; Social media; Social networks; Scholarly communication.

1. Introduction

Following the structural functionalism tradition in sociology, **Merton** (1942; 1973) formulated the ethos of science that includes several norms guiding actions of scientists. Norms specify actions that are justified as adequate or correct by a specific group of people (**Coleman**, 1990). The existence of the ethos of science distinguishes science from other sectors in society (**Wyatt**; **Milojevic**; **Park**; **Leydesdorff**, 2016); the norms enable scientists to fulfill the function of the science sector: the production of reliable and valid knowledge (**Luhmann**, 1992). Several studies investigating the ethos of science have revealed that scientists' actions seem to be guided by these norms. For example, the results by **Anderson** (2000) confirm that the "subscription to the Mertonian norms is very high, especially among faculty" (p. 457). However, on the other side, many studies revealed that scientists' actions are simultaneously characterized by norm-violating behavior. These studies are frequently grounded in the social-constructivist theory tradition (**Berger**; **Luckmann**, 1966), relativistic philosophy, and historiographic science studies (see **Cole**, 1992; **Mulkay**, 1979). These approaches assume non-normative actions of scientists, and emphasize the various forms of deceptions and biases in scientists' actions (see **Small**, 2020).

> ' Since the beginning of the use of citation data in research evaluation, many possible systematic biases have been proposed '

One area in the science of science field (see an overview of research in this field in **Fortunato** *et al.*, 2018) in which systematic biases have been frequently investigated are citation decisions of scientists. Since the beginning of the use of citation data in research evaluation, many possible systematic biases have been proposed. Overviews of these biases can be found in **Didegah** and **Thelwall** (2013), as well as in **Tahamtan** and **Bornmann** (2018; 2019). For example, the results of some studies suggested that citation decisions are dependent on the gender of the citing and cited authors, respectively. Investigations of systematic biases in science concern two of five norms that are part of the ethos of science: universalism and disinterestedness.

(1) Universalism demands that scientific work should not be ignored (or praised) based on irrelevant criteria such as the social status or the nationality of the scientist.

(2) Disinterestedness demands a neutral and impersonal stance of the scientists; that means, for example, the suppression of enthusiasm for own contributions (see **Ziman**, 1996). Both norms can be seen as efforts to avoid systematic biases in the action of scientists (**Tennant** *et al.*, 2018).

We wonder whether metrics which have been proposed as alternatives to bibliometrics (so called altmetrics) are concerned by biases that have been observed for citation data.

The results by **Van-Noorden** (2014) show that many researchers use *Twitter*

> "to follow discussions on research-related issues, and 40% said that it is a medium for 'commenting on research that is relevant to my field'" (p. 127).

Twitter is a web-based microblogging system

> "allowing users to post short messages. It has social networking features, letting users connect with each other" (**Mas-Bleda**; **Thelwall**, 2016, p. 2012).

Similar to citations of scholarly papers, tweets of papers are used to evaluate the impact or attention of papers (**Colledge**, 2014). *Twitter* counts are used as alternative to traditional bibliometrics besides other social media counts such as *Facebook* counts. Companies like *Altmetric.com* and *Plum Analytics* maintain databases and tools to use *Twitter* counts and other altmetrics to measure impact or attention of research. However, tweets should only be used for research evaluation purposes if systematic biases do not influence decisions to tweet, and tweet decisions by researchers follow the norms in the ethos of science.

It is an advantage of tweets –in contrast to citations– that an experimental approach can be applied to investigate the decision of scientists to communicate on papers. In the following section, we describe the possible design of a study that allows the experimental investigation of tweet decisions including randomized variations and –based on the normative and social-constructivist theories– theoretically derived mechanisms for explaining the empirical results (**Dablander**, 2020; **Lance**; **Guilkey**; **Hattori**; **Angeles**, 2014). Insights about biases in tweet decisions are urgently needed, since *Twitter* data are used in research evaluation studies since recently.

2. Design

Most of the studies that have investigated decisions to cite papers are based on an observational design. According to **Austin** (2011),

> "an observational study has the same intent as a randomized experiment: to estimate a causal effect. However, an observational study differs from a randomized experiment in one design issue: the use of randomization to allocate units to treatment and control groups. In observational studies, the treated subjects often differ systematically from untreated subjects" (p. 402).

In studies of gender bias in citation decisions, the cited papers published by men might differ systematically from the cited papers published by women, for instance, in the field-specific alignment (see **Parker** *et al.*, 2020). Such differences lead to the disadvantage of observational studies that biases in citation decisions cannot be ascertained without any doubts: the outcomes of the studies on citation decisions may be due to actual biases or a result of uncontrolled confounders. One of the few bias studies in scientometrics that applied a quasi-experimental design used papers that were published twice in order to investigate the influence of the journal on citation counts (**Larivière**; **Gingras**, 2010). However, the randomized assignment of papers to a treatment and control group could not be realized in this study.

> ' Universalism demands that scientific work should not be ignored (or praised) based on irrelevant criteria such as the social status or the nationality of the scientist '

According to **Falk** and **Heckman** (2009),

> "controlled variation is the foundation of empirical scientific knowledge" (p. 535).

In the recent past, the results of some field experiments have been published in social sciences that included controlled manipulations of conditions. Below, we present three exemplary studies to provide an insight into experiments based on various internet platforms; some other studies can be found in **Salganik** (2017). We used the following three studies as exemplars (best practices) for designing a possible study based on the *Twitter* platform.

(1) **Salganik**, **Dodds**, and **Watts** (2006) investigated the paradox that some hit songs are significantly more popular than an 'average' song, but their success cannot be predicted. The authors programmed a 'music market' in which people could download various songs. In the experimental setting, **Salganik** *et al.* (2006) manipulated the condition whether the people were informed or not about the song choices of other people. The results show that this information (the social influence) increased the inequality and the unpredictability of the songs' success.

(2) **Muchnik**, **Aral**, and **Taylor** (2013) ran a randomized experiment on a social news aggregation website (similar to *Digg. com* and *Reddit.com*) to investigate whether knowledge of the aggregated opinions of others distorts decision-making. About 100,000 comments submitted on the site were randomly assigned to an up-treated group (the comments received an up-vote), a down-treated group (the comments received a down-vote), and a control group (receiving neither an up- nor a down vote). The results of the experiment by **Muchnik** *et al.* (2013) show that

> "whereas positive social influence accumulates, creating a tendency toward ratings bubbles, negative social influence is neutralized by crowd correction … Our findings suggest that social influence substantially biases rating dynamics in systems designed to harness collective intelligence" (p. 650).

(3) **Van-de-Rijt**, **Kang**, **Restivo**, and **Patil** (2014) came up with three experimental designs in which they controlled the allocation of success. One study focused on the encyclopedia website *Wikipedia.org*. A randomly selected subset of very productive editors on this platform received

> "status awards from community members in recognition of their dedication" (**Van-de-Rijt** *et al.*, 2014, p. 6935).

This award was the condition that was manipulated in the study: the randomly selected treatment group received the award and the control group not. The results show that 40% of the editors in the treatment group received additional awards (from other editors). In the control group, the percentage was statistically significantly lower (31%). The result therefore demonstrated the existence of a causal link between past and future success.

An experimental *Twitter* study could start from the premise of the normative view of science (**Merton**, 1942; 1973). This view, according to **Cole** (1992), includes the perception that

> "the reaction of the scientific community to newly published work was believed to be determined by the content of that work" (p. 6).

The focus on content is mainly linked to the successful operation of the norms of universalism and disinterestedness in science. The opposite view of science, which is mainly rooted in the social-constructivist theory tradition (**Berger**; **Luckmann**, 1966), assumes that the reaction of the community is inherently subjective, based on irrelevant criteria for assessing research (such as the gender or nationality of the publishing author). The lack of reliability of reviewers' ratings in journal peer review processes (**Bornmann**; **Mutz**; **Daniel**, 2011) might be an indication of the existence of these subjective elements in the assessments of papers. Based on the normative and social-constructivist views of the causal dependencies of assessments in science, an experimental design using *Twitter* data could be used to decide which of the two conflicting views about assessments of papers might be correct. These views can be investigated by analyzing whether certain factors (e.g., the gender of a paper's authors) can significantly influence the assessment of tweets on papers.

For the experimental study, new *Twitter* accounts should be created that tweet about scientific papers in a manipulated way. Papers published in *Scientometrics* could be used because this journal is one of the major journals in bibliometrics. The manipulation is necessary to facilitate an experimental design and to estimate a causal effect. The account profiles should be as similar as possible regarding the photo, description, and background image. In order to prevent the accounts from looking like inactive or fake accounts, each account should start to follow some researchers from the

field of scientometrics and should retweet one tweet by each researcher. Some accounts should have popular US American names and others popular Chinese names. The use of typical names from two countries with an enormous paper output is intended to enable the investigation of a possible national bias in tweeting about papers and to avoid any misrepresentation of individuals with uncommon names.

> ' We wonder whether metrics which have been proposed as alternatives to bibliometrics (so called altmetrics) are concerned by biases that have been observed for citation data '

Every account tweets on papers that are selected from all papers published in *Scientometrics*. The tweets include the names of researchers from the scientometric community (by using @name) so that these researchers can retweet the tweets ('success') or not ('failure'). The researchers should be randomly selected from the pool of authors with a paper in *Scientometrics*. The sampled authors should be restricted to those with an active *Twitter* account. The different accounts that are created for the experimental setting are used for two settings.

(1) In the first setting, the experimental manipulation refers to the profile name. The same tweet (i.e., the same text referring to the same, randomly selected paper) is posted from accounts with different national appearance (Chinese or US American). Each tweet includes the name of a randomly selected scientometrician: "Paper entitled [title of the paper] @[*Twitter* handle of the randomly selected scientometrician] [URL of the paper]". The successful operation of the norms of disinterestedness and universalism would suggest that the tweets from the 'US American' account (occasion 1) are retweeted with a similar probability as the tweets from the 'Chinese' account (occasion 2). If –according to the social-constructivist view– subjective elements drive the decision to retweet, the tweets from the 'Chinese' account are retweeted with a statistically significantly lower or higher probability than those from the 'US American' account.

(2) The second experimental setting focuses on the gender of the corresponding author of the tweeted paper. For this setting, papers should be selected for which the corresponding authors' name clearly indicates his/her gender. The experiment reveals whether irrelevant criteria (i.e., the corresponding authors' gender) influence retweet decisions. The content of the tweet can be as follows: "Paper entitled [title of the paper] by [name of the female/male corresponding author] @[*Twitter* handle of the randomly selected scientometrician] [URL of the paper]". Some studies have shown that papers published by men received more citations than papers published by women (e.g., **Cole**; **Singer**, 1991; **Knobloch-Westerwick**; **Glynn**, 2013). However, evidence for the contrary is also available (e.g., **Lynn**; **Noonan**; **Sauder**; **Andersson**, 2019; **Strumia**, 2021). Based on these contradicting results on citation decisions, different retweet probabilities of tweets including male or female information might be detectable in the experiment (but only with a small effect size).

The collected data from the experimental settings should be anonymized by removing obvious personal identifiers such as *Twitter* profile names (**Salganik**, 2017). The removed personal identifiers should be stored separately in order to inform the involved researchers with study details and results when the study is finalized (see section ethical reflections about the planned study).

3. Statistics

The statistical analysis of the data from the experimental settings reveals whether the proportion of retweets differs between two occasions. A tweet is counted as retweeted when the randomly selected scientometrician mentioned in the tweet retweets the tweet within a certain period. Pairs of tweets are sent within two occasions which refer to the same paper each. As a result of each experimental setting (specified above), one receives data as exemplified in Table 1. Based on such data, the proportion of retweeted tweets can be compared between both occasions.

Table 1. Fictitious data that could result from an experiment (0 = the tweet is not retweeted, 1 = the tweet is retweeted)

	Occasion 1	Occasion 2
Tweet pair 1	1	0
Tweet pair 2	0	1
Tweet pair 3	1	0
Tweet pair 4	0	1
Tweet pair 5	1	0
Tweet pair 6	1	0
Tweet pair 7	1	1

Each line in Table 1 (reflecting the results of an experiment in a binary fashion) can be filled in a 2 × 2 table where the outcome of interest is a pair of results from 'occasion 1' and 'occasion 2', each either a 'success' (retweeted) or a 'failure' (not retweeted). Using tweet pair 7 from Table 1 as an example, Table 2 shows how a tweet pair (with retweeted or not retweeted results) is evaluated for the statistical analysis. Since the tweet pair 7 is retweeted on both occasions, the tweet pair (the paper) is categorized as retweeted-retweeted.

When all pairs in Table 1 have been filled in a table as the first pair in Table 2, one receives the 2 × 2 Table 3 that summarizes the complete data from Table 1. As the total in Table 3 reveals, seven papers have been tweeted twice. The results show, e.g., that four tweets on papers have been retweeted on occasion 1, but not on occasion 2.
The McNemar test can be used to

> "evaluate categorical data obtained in a true experiment (i.e., an experiment involving a manipulated independent variable). In such an experiment, the two scores of each subject (or pair of matched subjects) represent a subject's responses under the two levels of the independent variables (i.e., the two experimental conditions).

A significant result allows the researcher to conclude there is a high likelihood the two experimental conditions represent two different populations" (**Sheskin**, 2007, p. 817).

We calculated the McNemar test for the data in Table 3. As the results show the test statistic is 0.67 with the corresponding two-sided p value of 0.41, which provides no evidence of a (statistically significant) retweet difference between both occasions.

Power analysis methods can be used to compute the sample size for the experiments that are conducted. The results of the McNemar's test can be interpreted meaningfully with hindsight only when an a priori defined sample size is used (see **Acock**, 2018). Large samples are required in order to find small effects, and large effects can be detected with small samples. The power analysis can be applied to make the sample of the experiment as small as possible so that as few as possible researchers will receive the manipulated tweets (**Salganik**, 2017).

Table 2. Categorized data of tweet pair 7 in Table 1

		Occasion 1	
		Retweeted	Not retweeted
Occasion 2	Retweeted	1	
	Not retweeted		

Table 3. Summarized data of all tweet pairs in Table 1

		Occasion 1		Total
		Retweeted	Not retweeted	
Occasion 2	Retweeted	1	2	3
	Not retweeted	4	0	4
Total		5	2	7

4. Ethical reflections about an experimental *Twitter* study

Experimental studies should follow basically the ethical suggestions by **Salganik** (2017) which he has formulated in the chapter ethics in his book on social research in the digital age.

Scientometric studies are usually based on observational designs. Although scientometricians are interested in the estimation of causal effects, these estimations are not possible based on observational designs. Only randomized experiments in which units are allocated to treatment and control groups facilitate estimations of causal effects. In these studies, however, it is necessary to manipulate certain processes or objects to measure possible effects. In the experimental study on tweet decisions about papers, the tweeting process is manipulated by posting specific tweets from profiles with certain characteristics and including certain information. Although these tweets are manipulated postings for the receivers of the tweets, we do not expect significant harms for the receivers if the tweets do not contain any wrong or misleading information; manipulations concern tweeting with pseudonyms and mentioning correct information about the authors of papers or leaving them out.

After finalizing an experimental study, the scientists should be contacted (*Twitter* users and authors of the tweeted papers) who (unknowingly) have been involved in the study. The scientists can be contacted using the *Twitter* platform (based on personal and not public communications), email, or their address that they provide as author affiliation on their publications. They should be informed about the study design (especially the necessary minimal manipulations) and the results of the study. The opportunity to delete the corresponding tweets resulting from the study should be provided. Since the data of the experimental study should be stored for at least 10 years (see for example the guidelines for good scientific practice of the *German Research Foundation* and the *Max Planck Society*)[1], all personal data should be deleted after this period.

> We propose to undertake an experimental approach to the decision of scientists to 'tweet' on a paper

5. Conclusions

In this comment, we propose to undertake an experimental approach to the decision of scientists to 'tweet' on a paper. We describe the design of a posible study that allow the experimental investigation of tweet decisions including randomized variations and theoretically derived mechanisms for explaining the empirical results. The detailed design proposed in this comment can be adapted to other social media platforms (e.g., *Facebook* or *ResearchGate*). Using *Facebook*, the accounts could create posts that contain content as described in this comment. If other mentioned *Facebook* users share the post, this would be treated like a retweet on *Twitter*.

Users on *ResearchGate* can start discussions that contain comments like those we proposed for a *Twitter*-based study. Discussions on *ResearchGate* can be shared by other users. Such a sharing on *ResearchGate* could be treated as a retweet event on *Twitter*. Using *ResearchGate* would probably require more human work than studies using *Twitter* or *Facebook* because *ResearchGate* does not provide an application programming interface like *Twitter* or *Facebook*. Nevertheless, the increased human effort might be worthwhile because *ResearchGate* is explicitly oriented towards researchers in contrast to *Twitter* and *Facebook* that are primarily directed at the general public but also used by many researchers.

6. Note

1. See *https://zenodo.org/record/3923602*
and *https://www.mpg.de/199493/regelnWissPraxis.pdf*

7. References

Acock, Alan C. (2018). *A gentle introduction to Stata* (6th ed.). College Station, TX, USA: Stata Press. ISBN: 978 1 59718 269 0

Anderson, Melissa S. (2000). "Normative orientations of university faculty and doctoral students". *Science and engineering ethics*, v. 6, n. 4, pp. 443-461.
https://doi.org/10.1007/s11948-000-0002-6

Austin, Peter C. (2011). "An introduction to propensity score methods for reducing the effects of confounding in observational studies". *Multivariate behavioral research*, v. 46, n. 3, pp. 399-424.
https://doi.org/10.1080/00273171.2011.568786

Berger, Peter L.; **Luckmann, Thomas** (1966). *The social construction of reality: A treatise in the sociology of knowledge*. New York, NY, USA: Doubleday.
http://perflensburg.se/Berger%20social-construction-of-reality.pdf

Bornmann, Lutz; **Mutz, Rüdiger**; **Daniel, Hans-Dieter** (2011). "A reliability-generalization study of journal peer reviews: A multilevel meta-analysis of inter-rater reliability and its determinants". *PLoS one*, v. 5, n. 12, e14331.
https://doi.org/10.1371/journal.pone.0014331

Cole, Jonathan R.; **Singer, Burton** (1991). "A theory of limited differences. Explaining the productivity puzzle in science". In: H. Zuckerman; J. R. Cole; J. T. Bruer (eds.). *The outer circle. Women in the scientific community* (pp. 277-310). London, UK: W. W. Norton & Company. ISBN: 978 0 393027730

Cole, Stephen (1992). *Making science. Between nature and society*. Cambridge, MA, USA: Harvard University Press. ISBN: 0 674 54347 5

Coleman, James S. (1990). *Foundations of social theory*. Cambridge, MA, USA: Belknap Press of Harvard University Press. ISBN: 0674312260

Colledge, Lisa (2014). *Snowball metrics recipe book*. Amsterdam: Snowball Metrics Program Partners.
https://www.snowballmetrics.com/wp-content/uploads/0211-Snowball-Metrics-Recipe-Book-v7-LO.pdf

Dablander, Fabian (2020). *An introduction to causal inference*, 15 pp.
https://psyarxiv.com/b3fkw

Didegah, Fereshteh; **Thelwall, Mike** (2013). "Determinants of research citation impact in nanoscience and nanotechnology". *Journal of the American Society for Information Science and Technology*, v. 64, n. 5, pp. 1055-1064.
https://doi.org/10.1002/asi.22806

Falk, Armin; **Heckman, James J.** (2009). "Lab experiments are a major source of knowledge in the social sciences". *Science*, v. 326, n. 5952, pp. 535-538.
https://doi.org/10.1126/science.1168244

Fortunato, Santo; **Bergstrom, Carl T.**; **Börner, Katy**; **Evans, James A.**; **Helbing, Dirk**; **Milojević, Staša**; **Petersen, Alexander M.**; **Radicchi, Filippo**; **Sinatra, Robert**; **Uzzi, Brian**; **Vespignani, Alessandro**; **Waltman, Ludo**; **Wang, Dashun**; **Barabási, Albert-László** (2018). "Science of science". *Science*, v. 359, n. 6379, eaao018.
https://doi.org/10.1126/science.aao0185

Knobloch-Westerwick, Silvia; **Glynn, Carroll J.** (2013). "The Matilda effect - Role congruity effects on scholarly communication: A citation analysis of *Communication research* and *Journal of communication* articles". *Communication research*, 40, v. n. 1, pp. 3-26.
https://doi.org/10.1177/0093650211418339

Lance, Peter M.; **Guilkey, David K.**; **Hattori, Aiko**; **Angeles, Gustavo** (2014). *How do we know if a program made a difference? A guide to statistical methods for program impact evaluation*. Carolina: Measure Evaluation. ISBN: 978 0 692 23861 5
https://www.measureevaluation.org/resources/publications/ms-14-87-en/at_download/document

Larivière, Vincent; **Gingras, Yves** (2010). "The impact factor's Matthew effect: A natural experiment in bibliometrics". *Journal of the Association for Information Science and Technology*, v. 61, n. 2, pp. 424-427.
https://doi.org/10.1002/asi.21232

Luhmann, Niklas (1992). *Die Wissenschaft der Gesellschaft*. Frankfurt am Main, Germany: Suhrkamp. ISBN: 978 3 518286012

Lynn, Freda B.; **Noonan, Mary C.**; **Sauder, Michael**; **Andersson, Matthew A.** (2019). "A rare case of gender parity in Academia". *Social forces*, v. 98, n. 2, pp. 518-547.
https://doi.org/10.1093/sf/soy126

Mas-Bleda, Amalia; **Thelwall, Mike** (2016). "Can alternative indicators overcome language biases in citation counts? A comparison of Spanish and UK research". *Scientometrics*, v. 109, n. 3, pp. 2007-2030.
https://doi.org/10.1007/s11192-016-2118-8

Merton, Robert K. (1942). "Science and technology in a democratic order". *Journal of legal and political sociology*, n. 1, pp. 115-126.

Merton, Robert K. (1973). *The sociology of science: Theoretical and empirical investigations*. Chicago, IL, USA: University of Chicago Press. ISBN: 0 226 52091 9
https://sciencepolicy.colorado.edu/students/envs_5110/merton_sociology_science.pdf

Muchnik, Lev; **Aral, Sinan**; **Taylor, Sean J.** (2013). "Social influence bias: A randomized experiment". *Science*, v. 341, n. 6146, pp. 647-651.
https://doi.org/10.1126/science.1240466

Mulkay, Michael (1979). *Science and the sociology of knowledge*. London, UK: George Allen and Unwin. ISBN: 978 0 043010938

Parker, Philip D.; **Van-Zanden, Brooke**; **Marsh, Herbert W.**; **Owen, Katherine**; **Duineveld, Jasper J.**; **Noetel, Michael** (2020). "The intersection of gender, social class, and cultural context: A meta-analysis". *Educational psychology review*, v. 32, n. 1, pp. 197-228.
https://doi.org/10.1007/s10648-019-09493-1

Salganik, Matthew J. (2017). *Bit by bit: Social research in the digital age*. Oxfordshire, UK: Princeton University Press. ISBN: 978 1 400888184

Salganik, Matthew J.; **Dodds, Peter-Sheridan**; **Watts, Duncan J.** (2006). "Experimental study of inequality and unpredictability in an artificial cultural market". *Science*, v. 311, n. 5762, pp. 854-856.
https://doi.org/10.1126/science.1121066

Sheskin, David J. (2007). *Handbook of parametric and nonparametric statistical procedures* (4th ed.). Boca Raton, FL, USA: Chapman & Hall/CRC. ISBN: 978 0 429186196
https://doi.org/10.1201/9780429186196

Small, Henry (2020). "Past as prologue: Approaches to the study of confirmation in science". *Quantitative science studies*, v. 1, n. 3, pp. 1025-1040.
https://doi.org/10.1162/qss_a_00063

Strumia, Alessandro (2021). "Gender issues in fundamental physics: A bibliometric analysis". *Quantitative science studies*, v. 2, n. 1, pp. 225-253.
https://doi.org/10.1162/qss_a_00114

Tahamtan, Iman; **Bornmann, Lutz** (2018). "Core elements in the process of citing publications: Conceptual overview of the literature". *Journal of informetrics*, v. 12, n. 1, pp. 203-216.
https://doi.org/10.1016/j.joi.2018.01.002

Tahamtan, Iman; **Bornmann, Lutz** (2019). "What do citation counts measure? An updated review of studies on citations in scientific documents published between 2006 and 2018". *Scientometrics*, v. 121, n. 3, pp. 1635-1684.
https://doi.org/10.1007/s11192-019-03243-4

Tennant, Jonathan P.; **Dugan, Jonathan M.**; **Graziotin, Daniel**; **Jacques, Damien C.**; **Waldner, Francois**; **Mietchen, Daniel**; **Elkhatib, Yehia**; **Collister, Lauren**; **Pikas, Christina K.**;**... Colomb, Julian** (2018). "A multi-disciplinary perspective on emergent and future innovations in peer review [version 3; peer review: approved]". *F1000Research*.
https://f1000research.com/articles/6-1151

Van-de-Rijt, Arnout; **Kang, Soong-Moon**; **Restivo, Michael**; **Patil, Akshay** (2014). "Field experiments of success-breeds-success dynamics". *Proceedings of the National Academy of Sciences of the United States of America*, v. 111, n. 19, pp. 6934-6939.
https://doi.org/10.1073/pnas.1316836111

Van-Noorden, Richard (2014). "Scientists and the social network". *Nature*, v. 512, n. 7513, pp. 126-129.
https://doi.org/10.1038/512126a

Wyatt, Sally; **Milojevic, Staša**; **Park, Han-Woo**; **Leydesdorff, Loet** (2016). "The intellectual and practical contributions of scientometrics to STS". In: U. Felt; R. Fouché; C. A. Miller; L. Smith-Doerr (eds.). *Handbook of science and technology studies* (pp. 87-112): The MIT Press. ISBN: 978 0 262035682
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2588336

Ziman, John (1996). "'Postacademic science': Constructing knowledge with networks and norms". *Science studies*, v. 9, n. 1, pp. 67-80.