Journal of Informetrics 10 (2016) 48–61

Contents lists available at ScienceDirect

Journal of Informetrics

journal homepage: www.elsevier.com/locate/joi

# National, disciplinary and temporal variations in the extent to which articles with more authors have more impact: Evidence from a geometric field normalised citation indicator

Mike Thelwall*, Pardeep Sud

*Statistical Cybermetrics Research Group, School of Mathematics and Computer Science, University of Wolverhampton, Wulfruna Street, Wolverhampton WV1 1LY, UK*

ARTICLE INFO

*Article history:*
Received 9 October 2015
Received in revised form
20 November 2015
Accepted 22 November 2015
Available online 18 December 2015

*Keywords:*
Scientific co-authorship
gMNCS, MNCS
Geometric Mean Normalized Citation Score
Geometric new crown indicator
Research collaboration

ABSTRACT

The importance of collaboration in research is widely accepted, as is the fact that articles with more authors tend to be more cited. Nevertheless, although previous studies have investigated whether the apparent advantage of collaboration varies by country, discipline, and number of co-authors, this study introduces a more fine-grained method to identify differences: the geometric Mean Normalized Citation Score (gMNCS). Based on comparisons between disciplines, years and countries for two million journal articles, the average citation impact of articles increases with the number of authors, even when international collaboration is excluded. This apparent advantage of collaboration varies substantially by discipline and country and changes a little over time. Against the trend, however, in Russia solo articles have more impact. Across the four broad disciplines examined, collaboration had by far the strongest association with impact in the arts and humanities. Although international comparisons are limited by the availability of systematic data for author country affiliations, the new indicator is the most precise yet and can give statistical evidence rather than estimates.



## 1. Introduction

Cooperation in research is promoted by many funding agencies in the belief that collaborative research tends to have more impact. This originates from the theoretical argument that interdisciplinary collaboration is often necessary to solve important societal problems (Gibbons et al., 1994) and is supported by studies showing that collaborative research is often more highly cited than comparable solo studies (e.g., Thurman & Birkinshaw, 2006). International collaboration seems to be also promoted for partly political purposes, such as the European Union funding programmes that require at least three different member states to be represented within a funding bid (EC, 2014). These initiatives have presumably contributed to an increase in research collaboration in most fields (Wuchty, Jones, & Uzzi, 2007). Nevertheless, the value of collaboration seems to vary between fields, nations and type (e.g., national vs. international) and so it is important to understand where it is beneficial so that it can be promoted when it is most useful and perhaps even discouraged when it is problematic.

Although, as discussed below, previous studies have assessed factors that influence the success of collaborations, at least as reflected in the citation counts of the resulting publications, it is difficult to get a clear understanding of differences between

* Corresponding author. Tel.: +44 1902 321470; fax: +44 1902 321478.
*E-mail addresses:* m.thelwall@wlv.ac.uk (M. Thelwall), p.sud@wlv.ac.uk (P. Sud).

http://dx.doi.org/10.1016/j.joi.2015.11.007
1751-1577/

collaboration types. This is because citation counts are highly skewed (e.g., Seglen, 1992) and so comparing arithmetic mean citation rates between different types of articles is unreliable, needing large sample sizes to give reasonable statistical power. Moreover, analyses need to harness large sets of articles in order to reliably distinguish between the average impacts of sets of articles with similar properties (e.g., articles with two authors compared to articles with three authors). Hence a more precise method is needed to compare the effects of collaboration on collections of articles, such as the geometric mean. This may not be enough, however, since the geometric mean can only be applied to articles from a single year and field because of differences in average citation counts. Hence, an indicator is needed that combines the ability of the geometric mean to deal with skewed data with the ability of the Mean Normalized Citation Score (MNCS) (discussed below) to combine citation counts from multiple fields and years. In response, this article introduces a variant of the MNCS, the geometric MNCS (gMNCS), and applies it to assess the effect of field, year and country on the extent to which the average citation impact of collaborative research articles varies with the number of authors. Geometric variants of several standard bibliometric indicators have previously been proposed, following their initial introduction (Zitt, 2012). These include geometric journal impact factors (Thelwall & Fairclough, 2015) and a basic average citation indicator for individual subjects and years (Fairclough & Thelwall, 2015).

## 2. Background

Academic collaboration is the combining together of the expertise of multiple people in the production of research (Katz & Martin, 1997). In practice, even apparently solo research projects are sometimes collaborative to some extent through informal discussions with colleagues and help from support staff. Whilst these are important parts of the research process, collaborations that combine a substantial amount of academic expertise from the contributors are of particular interest because of its promotion by funding agencies in the belief that it tends to produce better research. In practice, quantitative studies of collaboration almost always focus on work that leads to published findings (there are many qualitative exceptions, e.g.: Latour & Woolgar, 1979) and use the authorship list as a proxy for the set of people that have substantially contributed to a study. Other contributors are sometimes recognised in an acknowledgement (Cronin, McKenzie, & Stiffler, 1992; Cronin, 2001a) but these are rarely analysed on a large scale.

The authorship list is a simplification of the concept of collaboration because it may omit important contributors (ghost authorship: Gotzsche et al., 2007) and include non-contributors (gift/honorary authorship: Cronin, 2001b; Drenth, 1998; Smith, 1994). Scientists also do not have a uniform understanding of concept of research collaboration and frequently do not grant co-authorships to people that have helped in research (Laudel, 2002). Moreover, although the authors are normally assumed to have contributed equally, in most fields the first author probably contributes more than the others (Vinkler, 1993). This is not true in all fields, with exceptions including mathematics, business and economics (Levitt & Thelwall, 2013) and there is no agreed formula to estimate the likely relative contributions of authors based on their order in the authorship list. It is becoming more possible to detect the value of the different authors for a paper because some journals require specific information about individual contributions (Bates, Anić, Marušić, & Marušić, 2004) but this falls short of giving a formula to estimate the relative importance of each one. In this context, it seems reasonable to accept the simplification that all authors' contributions are equally important.

Academic collaboration leading to co-authorship can be of very different types. A common type is probably junior-senior co-authorship where the main author is a PhD student and the second author is their main supervisor. Here, the student may have done most of the work but the supervisor may have provided expertise in the form of ideas and overall guidance on topic areas and specific advice about the research design, methods, analysis, write-up and publication venue. The exact nature of the relationship may vary by discipline, however (e.g., Barnes & Randall, 2012). In contrast, some collaborations involve sets of experienced researchers that provide complementary expertise from different fields, subfields, or tasks (e.g., statistics, interviews), in order to conduct studies that they could not perform as well individually. Other collaborations may also be between researchers with essentially identical skill sets but with their combined insights helping to solve a problem that they could, in theory, have addressed individually. For a large scale bibliometric analysis of publications no method has yet been developed to distinguish between these types of contributions and so there is no alternative to treating all types of collaboration as the same. When interpreting the results, however, the different types of collaboration should be considered as possible explanations for any patterns found.

Many, but not all (Bornmann, Schier, Marx, & Daniel, 2012; Haslam et al., 2008), studies investigating the connection between collaboration and citation have found that articles with more authors tend to be more cited (e.g., Thurman & Birkinshaw, 2006; Vieira & Gomes, 2010). Most articles cannot be easily generalised, however, due to a focus on a set of publications with a specific attribute, such as originating from a single university, country, journal or field. There have also been variations in the types of collaboration examined, from a course grained comparison of solo with collaborative research, to comparisons of types of collaboration (e.g., intra-institutional, international) and different numbers of authors.

Not all types of collaboration have equal apparent impact. It seems that collaboration is particularly likely to generate higher (arithmetic mean) impact research if the collaborators are from different countries (Didegah & Thelwall, 2013; Glänzel, 2001; Katz & Hicks, 1997), except perhaps in the social sciences (Didegah & Thelwall, 2013), and for authors at prestigious universities (e.g., Gazni & Didegah, 2010). Domestic collaborations seem to have the same impact whether multiple institutions are involved or not, however (Didegah & Thelwall, 2013). Moreover, collaboration within an institution associates with lower impact papers that solo research in at least one field (Leimu & Koricheva, 2005). The advantage of