Coauthored doctoral dissertations: Quantifying the candidate's share

Hagen Nils T., *University of Nordland, Norway*

Abstract

The new norm for doctoral dissertations in many fields of modern science is a thesis comprised of several coauthored papers or manuscripts. The prospective candidate joins a research team, and the doctorate is then based on that individual's contribution to a collaborative effort. Such candidatures hinge on the verifiable academic achievement documented by an accumulation of partial authorships. The candidate's share of credit is rarely specified numerically, but can be estimated from information in each paper's byline. Quantification of the candidate's share of credit provides baseline data for the comparison of theses from established doctoral programs, and may contribute to reducing the ambiguity of existing doctoral guidelines by providing an attainable benchmark for new candidates. This presentation provides a general outline of the quantification procedure, expanded by detailed examples of credit partitioning for doctoral candidates working in small to medium sized research teams consisting of their supervisors and other collaborators.

Background

A modern doctoral thesis consists of several papers, for each of which the candidate may have one or more coauthors. The sum of the candidate's contributions to these papers constitutes the modern equivalent of a traditional monograph thesis (Wilson 2002).

However, institutional guidelines for modern dissertations are not yet standardized (Jackson 2013), and pertinent questions persist, such as, how many papers it is necessary to include in a thesis, and how large a portion of each paper the candidate is expected to produce (Wilson 1996). A tentative answer to these questions was provided by a recent study of two Scandinavian universities, which showed that most PhD theses consisted of 4 coauthored papers, and suggested that between 40% and 50% of the authorship credit in the submitted papers was attributable to the candidate (Hagen 2010). The authorship credit of a typical dissertation from these universities corresponded to approximately 1.6 undivided papers, thus providing a *de facto* baseline for comparison with other publication based PhD programs.

This presentation outlines and exemplifies the bibliometric procedure for quantifying and dividing publication credit between the candidate and the collaborators appearing as coauthors on papers included in the dissertation.

Methodology

Doctoral candidates, like most coauthors of scientific papers (Waltman 2012), generally do not contribute equally, and neither do they specify numerically the relative size of their contribution. Therefore, the candidate's share of a multi-authored dissertation cannot be accurately estimated by currently employed bibliometric methods, which rely on the false assumption that all coauthor contributions are approximately equal. To accurately estimate the candidate's share requires an evidence-based approach capable of including information implicit in the byline structure of a multi-authored paper, as well as all relevant information explicitly provided in the associated footnotes.

The harmonic formula, first proposed by Hodge and Greenberg (1981), provides a standardized quantification scheme for the characteristic hierarchical byline structure of a multi-authored paper with unequal contributors. Furthermore, harmonic estimates are easily modified to accommodate additional byline information about the specific size of one or more contributions, the equality of some (or all) contributions, or the presence of a senior author. The senior author is a corresponding last author (e.g. a supervisor), whose contribution is conventionally considered as being equivalent to the first author's.

The harmonic formula captures the essential features of a hierarchical byline structure by adhering to 3 simple ethical criteria (Hagen 2008, Hagen 2010):

1. One publication credit is shared among all coauthors.
2. The first author gets the most credit, and in general the i^{th} author receives more credit than the $(i + 1)^{th}$ author.
3. The greater the number of authors, the less credit per author.

Other proposed formulas do not fully satisfy the criteria, and tend to be more complicated. They also tend to be less accurate than the harmonic formula when validated against empirical data (Hagen 2010, Hagen 2013).

Harmonic coauthor credit

For a paper with 2 coauthors and an unadorned hierarchical byline, the harmonic credit is 66.67% for the first author and 33.33% for the second author. If the last (second) author is a senior (corresponding) author, or if it is explicitly indicated that both authors have contributed equally, then each receives 50% of the credit.

For a paper with 3 coauthors the respective harmonic credit estimates are 54.55%, 27.27%, and 18.18%. If there are 2 equal first authors they each get 40.91% credit, and the credit of the last author is unaffected. But if the last author is a senior author then the first and the senior get 40.91% and the second author is demoted and gets 18.18%. If the paper has 2 senior authors, or 2 equal first authors and a senior author, then all authors are considered as equal contributors and each receives 33.33% credit.
Similarly, for a paper with 4 coauthors the respective harmonic credit estimates are 48%, 24%, 16%, and 12%. It follows that the only way that a doctoral candidate can safely assume responsibility for more than 50% of a paper with more than 3 coauthors, is to indicate specifically that the sum of the other coauthors' contributions is less than 50% of the total credit.
For 5 coauthors the estimates are 43.8%, 21.9%, 14.6%, 10.95%, and 8.76%.
And for 6 coauthors 40.82%, 20.41%, 13.61%, 10.2%, 8.16%, and 6.8%.

Implications
When authorship credit is shared among several coauthors the credit for each author is reduced (cf. ethical criterion 3 above). Thus a baseline of approximately 40% credit retention by the doctoral candidate, as appears to be the current baseline in Scandinavia (Hagen 2010), corresponds to first authorship on papers with no more than 6 coauthors. However, if the first authorship is shared, or the paper has a senior author, there can be no more than 3 coauthors if the candidate's (first author's) share is going to remain above 40%.
Alternatively, a baseline value corresponding to approximately 1.6 undivided papers (Hagen 2010), could be accomplished with an average retention value of less than 40% if the dissertation includes more than 4 papers. The implication is that candidates who are members of large research teams and have many coauthors would have to submit more papers than candidates who are members of small teams, in order to achieve comparable authorship credit for their dissertations.
Note that a well-documented trend towards an increasing number of coauthors (Wuchty, Jones et al. 2007), combined with a less well-documented trend towards fewer dissertation papers, may erode the candidate's share of credit (Hagen 2010). It is important to document whether this "sliding baseline syndrome" affects publication based doctoral programs, by standardizing quantification of authorship credit per dissertation over appropriate time periods.

It is also important to compare the amount of credit per dissertation in different disciplines and different institutions. Improved knowledge about the actual amount of authorship credit in completed doctoral dissertations may facilitate informed debate, and presumably provide a useful reference for individual stakeholders in research teams which includes PhD candidates (cf. Gross, Alhusen et al. 2012).

References

Gross, D., J. Alhusen and B. M. Jennings (2012). "Authorship ethics with the dissertation manuscript option." Research in Nursing & Health **35**(5): 431-434.

Hagen, N. T. (2008). "Harmonic allocation of authorship credit: source-level correction of bibliometric bias assures accurate publication and citation analysis." PLoS ONE **3**(12): e4021:4021-4027.

Hagen, N. T. (2010). "Deconstructing doctoral dissertations: How many papers does it take to make a PhD?" Scientometrics **85**(2): 567-579.

Hagen, N. T. (2010). "Harmonic publication and citation counting: sharing authorship credit equitably – not equally, geometrically or arithmetically." Scientometrics **84**(3): 785-793.

Hagen, N. T. (2013). "Harmonic coauthor credit: A parsimonious quantification of the byline hierarchy." Journal of Informetrics **7**(4): 784–791.

Hodge, S. E. and D. A. Greenberg (1981). "Publication credit." Science **213**(4511): 950.

Jackson, D. (2013). "Completing a PhD by publication: a review of Australian policy and implications for practice." Higher Education Research & Development **32**(3): 355-368.

Waltman, L. (2012). "An empirical analysis of the use of alphabetical authorship in scientific publishing." Journal of Informetrics **6**(4): 700-711.

Wilson, K. (1996). The award of the degree of PhD on the basis of published work in the UK, UK Council for Graduate Education.

Wilson, K. (2002). "Quality assurance issues for a PhD by published work: A case study." Quality Assurance in Education **10**(2): 71-78.

Wuchty, S., B. F. Jones and B. Uzzi (2007). "The increasing dominance of teams in production of knowledge." Science **316**(5827): 1036-1039.