

# A Classification of Author Co-citations: Definitions and Search Strategies

**Ronald Rousseau**

*KHBO, Dept. of Industrial Sciences and Technology, B-8400 Oostende, Belgium; University of Antwerp, IBW, B-2610 Wilrijk, Belgium. E-mail: ronald.rousseau@khbo.be*

**Alesia Zuccala**

*Faculty of Information Studies, University of Toronto, 140 St. George Street, Toronto, Ontario M5S 3G6, Canada; School of Information Systems, University of New South Wales, Technology and Management, Sydney NSW 2052, Australia. E-mail: zuccala@fis.utoronto.ca; a.zuccala@unsw.edu.au*

**The term author co-citation is defined and classified according to four distinct forms: the *pure first-author co-citation*, the *pure author co-citation*, the *general author co-citation*, and the *special co-author/co-citation*. Each form can be used to obtain one count in an author co-citation study, based on a binary counting rule, which either recognizes the co-citedness of two authors in a given reference list (1) or does not (0). Most studies using author co-citations have relied solely on *first-author co-citation* counts as evidence of an author's *oeuvre* or body of work contributed to a research field. In this article, we argue that an author's contribution to a selected field of study should not be limited, but should be based on his/her complete list of publications, regardless of author ranking. We discuss the implications associated with using each co-citation form and show where simple first-author co-citations fit within our classification scheme. Examples are given to substantiate each author co-citation form defined in our classification, including a set of sample Dialog™ searches using references extracted from the SciSearch database.**

## Introduction

Citation analysis may be described as the sub-field of bibliometrics where one analyzes the patterns and frequencies of citations given as well as received. This is done on the level of authors, journals, scientific disciplines, and so on. Citation analysis also studies relations between cited and citing units (documents, authors, countries, etc.). This definition of the term "citation analysis" is adapted from the one given by Spinak (1996).

Citation studies are mainly undertaken for faculty and institutional evaluation purposes (e.g., Moed, Burger, Frankfort, & Van Raan, 1985), to study the structure and development of a scientific field (e.g., Small, 1999; Small & Griffith, 1974), and to study citer motivation (e.g., Brooks, 1985; Case & Higgins, 2000). Such studies may focus on different actors, including authors, journals, institutes, countries, scientific journals, or combinations thereof. In this article, we focus on the "co-citedness" of *authors* and the structure of a research field based on the research interests of its researchers.

Author Co-citation Analysis (ACA) is a widely recognized research technique, which has received a significant amount of attention in past years (e.g., Ahlgren, Jarneving, & Rousseau, 2003; McCain, 1990b; Persson, 2001; White, 1986, 1990; White & Griffith, 1981a,b, 1982). White and Griffith (1981a,b) first introduced ACA in their study of authors from the field of judgment and decision research. Since the publication of this article, the standard references for researchers interested in applying the technique have been White's (1986) report on co-cited *author* retrieval and McCain's (1990a) technical overview. ACA has been used specifically to trace changes in a field over time (McCain, 1984, McCain, 1985), to test for the possibility of convergence among research traditions (Borgman & Rice, 1992), to test the "branching" model of scientific growth (Perry & Rice, 1998), and to understand how scholars seek and use information in the creation of new knowledge (Sandstrom, 1998).

Despite its widespread use, one of the interesting problems associated with ACA is that the term used to substantiate this bibliometric technique—*author co-citation*—has not been fully defined or classified according to its various forms. Although there are different ways to *collect* and *observe an author co-citation*, the research literature has

---

Received July 1, 2003; revised October 14, 2003; accepted October 14, 2003

© 2004 Wiley Periodicals, Inc.

mainly been using one definition, which refers to the co-citationist's use of oeuvres or "body of writings by the same author (or first author in collaborations)" (White & Griffith, 1982, p. 257). Expressed otherwise, it is stated that "two authors are co-cited when at least one document in each other's oeuvre occurs in the same reference list" (McCain, 1990b, p. 195). The term *author co-citation*, therefore, needs further clarification. To provide this clarification, we propose a new classification scheme, one that defines the various forms of author co-citedness and describes the techniques that might be used to retrieve or collect co-citation data for an ACA. This description is inspired and guided in part by Christensen and Ingwersen's articles (1996; 1997), which concentrate more on the use of RANK, MAP, and TARGET provided by Dialog™ and the duplicate removal technique. In the following, a list of author co-citation forms is presented and for each form we provide a fictitious example.

This article is an expanded version of a talk presented during the 9th International Conference on Scientometrics and Informetrics, and published in the Conference Proceedings (Rousseau & Zuccala, 2003).

### Author Co-citation Forms: A Classification

The classification scheme that we propose is based on what "being co-cited" can tell us about the field's structure and the relationships among authors who contribute to this structure. Citing two reports with no authorship overlap, as shown in Example a, *might* contain a clue about the intellectual relationship between the authors. Citing a co-authored report, as shown in Example b, tells us *nothing new* about the relationship (scientific interest) between joint authors.

#### Example a

Smith, K. (1990). *A splendid result*. *The Alpha Journal*.  
Thomas, T. (1991). *The use of this-and-that: Part I*. *The Beta Journal*.

#### Example b

Smith, K., & Thomas, T. (1992). *The k-procedure*. *The Gamma Journal*.  
Thomas, T., & Smith, K. (1992). *More details about the k-procedure*. *The Gamma Journal*.

The first form, shown in Example a is considered to be the purest (and most interesting) form. Example b by comparison is "less pure" because there is already a given relationship or "intellectual similarity" between the authors, which does not result from "being co-cited" (namely co-authorship).

Consequently, the first principle of our classification scheme is to give preference to author co-citations that do not involve the type of co-authorship shown in Example b. The second classifying principle is that if the order of the authors in a co-authorship pair is significant with respect to the claims or novelties brought forward by each collaborator, then prefer-

ence may be given to the co-citation of first-authored articles over the co-citation of secondary-authored articles. This approach, however, is not studied in detail in this contribution. Note also that it is clearly not always the case that ranking of authors in the byline reflects importance (the authors of this article are ordered alphabetically).

And, finally, we consider co-citation counting as a form of binary counting (0-1 counting): one reference list produces at most one count for the total co-citation score of two authors. Once we have introduced the details of this full binary classification, we will suggest some further generalizations, including some non-binary methods.

### Pure First-Author Co-citations

The term *pure first-author co-citation* refers to the situation where at least one publication with A as first or sole author and one publication with B as first or sole author co-occur in the reference list of an article. Articles, however, with A and B as co-authors are not taken into account. In practice, this means that when we examine a reference list for a particular pair of authors, all articles with these two authors as co-authors are not included in the calculation of their pure first-author co-citation frequency.

Let us consider examples a and c. Example a yields one Smith-Thomas *pure first-author co-citation*. Equally, the next example c yields exactly one *pure first-author co-citation* for the Smith-Thomas pair even though there are two cited papers with Thomas as first author in the reference list (due to the binary method of counting).

#### Example a

Smith, K. (1990). *A splendid result*. *The Alpha Journal*.  
Thomas, T. (1991). *The use of this-and-that: Part I*. *The Beta Journal*.

#### Example c

Smith, K. (1990). *A splendid result*. *The Alpha Journal*.  
Thomas, T. (1991). *The use of this-and-that: Part I*. *The Beta Journal*.  
Thomas, T., & Zhang, W. (1992). *The use of this-and-that: Part II*. *The Beta Journal*.

Examples b and d, however, do not yield a pure first-author co-citation for the pair Smith-Thomas. With example b, there is a 0 co-citation count because of the co-authorship link between the pair of authors, and with example d, the pure first author form also yields a 0 count for the authors in question because both are ranked secondary.

#### Example b

Smith, K., & Thomas, T. (1992). *The k-procedure*. *The Gamma Journal*.  
Thomas, T., & Smith, K. (1992). *More details about the k-procedure*. *The Gamma Journal*.

#### Example d

Janssens, J., & Smith, K. (1996). *The breakthrough*. New World: Delta Publishers.

Peters, P., Zhang, W., & Thomas, T. (1999). *The breakthrough revisited*. The Alpha Journal.

#### Pure Co-citations

The term *pure co-citation* refers to a situation where at least one publication with A as a co-author (regardless of rank order) and one publication with B as a co-author (regardless of order) co-occur in the reference list of an article. Articles with A and B as co-authors are, once again, not taken into account.

Every pure first-author co-citation is also a pure co-citation. Example d gives one pure co-citation for the Smith-Thomas pair. Examples b and e below, however, do not yield a pure co-citation for this pair.

#### Example b

Smith, K., & Thomas, T. (1992). *The k-procedure*. The Gamma Journal.

Thomas, T., & Smith, K. (1992). *More details about the k-procedure*. The Gamma Journal.

#### Example d

Janssens, J., & Smith, K. (1996). *The breakthrough*. New World: Delta Publishers.

Peters, P., Zhang, W., & Thomas, T. (1999). *The breakthrough revisited*. The Alpha Journal.

#### Example e

Smith, K., & Thomas, T. (1992). *The k-procedure*. The Gamma Journal.

Thomas, T., & Smith, K. (1992). *More details about the k-procedure*. The Gamma Journal.

Rao, S., & Thomas, T. (1994). *Problems with the k-procedure*. The Gamma Journal.

Finally, example f yields one pure co-citation for Smith-Thomas, in that the first two articles are not taken into account, but the last two are.

#### Example f

Smith, K., & Thomas, T. (1992). *The k-procedure*. The Gamma Journal.

Thomas, T., & Smith, K. (1992). *More details about the k-procedure*. The Gamma Journal.

Rao, S., & Thomas, T. (1994). *Problems with the k-procedure*. The Gamma Journal.

Janssens, J., & Smith, K. (1996). *The breakthrough*. New World: Delta Publishers.

#### General Co-citations

The term *general co-citation* refers to a situation where at least one publication with A as a co-author, and one (additional) publication with B as a co-author co-occur in the same reference list. All pure co-citations, and

consequently also all pure first-author co-citations, are general co-citations, but the general co-citation is distinct because this time the articles co-authored by A and B are taken into account. If A and B *only* occur in a reference list as co-authors, then the co-citation of their co-authored articles does not count as a *general co-citation*. This definition implies that example e shown above includes a Smith-Thomas general co-citation. Examples b, g and h, however, are not Smith-Thomas general co-citations.

#### Example g

Smith, K., & Thomas, T. (1992). *The k-procedure*. The Gamma Journal.

Zhang, W. (1999). *A local k-procedure*. The Phi Journal.

Rao, S., & Janssens, J. (1998). *A note on the m-procedure*. The Iota Journal.

#### Example h

Smith, K., & Thomas, T. (1992). *The k-procedure*. The Gamma Journal.

Thomas, T., & Smith, K. (1992). *More details about the k-procedure*. The Gamma Journal.

Thomas, T., & Smith, K. (1999). *A review of the k-procedure*. The Gamma Journal.

Zhang, W. (1999). *A local k-procedure*. The Phi Journal.

Rao, S., & Janssens, J. (1998). *A note on the m-procedure*. The Iota Journal.

#### Co-author/Co-citation Scores

The term *co-author/co-citation* score is used to suggest that in addition to the first three co-citation forms, a publication with A and B as co-authors can be used to count a special form of co-citation, which recognizes the intellectual link between the authors due to their collaborative work together. In all previous examples, including g and h, we, therefore, have at least one Smith-Thomas *co-author/co-citation*. A *co-author/co-citation* occurs when authors A and B, in this case, Smith and Thomas, co-occur in the reference list (as a whole). This may happen because they are co-cited, either as first author or secondary author, OR because they are listed as co-authors.

The full classification scheme yields a hierarchy, beginning with: (1) *pure first-author co-citations*, (2) *pure co-citations*, (3) *general co-citations*, and (4) the “special” *co-author/co-citation score*. Figure 1 illustrates this hierarchy.

#### Weighted Counting

In addition to the “normal” (binary) form of counting, one may also use *weighted counting* to take into consideration the number of co-citations in the same document. For instance, example c (also shown in Pure First Author Co-citations) would, under a weighted scheme, yield two co-citation counts for the Smith-Thomas pair.

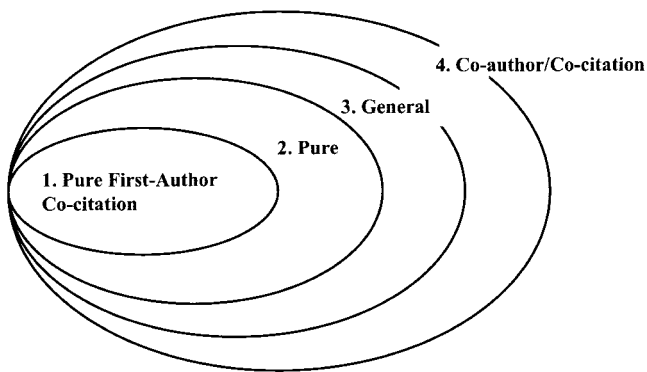


FIG. 1. Classification hierarchy of author co-citation forms.

Example c

Smith, K. (1990). *A splendid result. The Alpha Journal.*  
 Thomas, T. (1991). *The use of this-and-that: Part I. The Beta Journal.*  
 Thomas, T., & Zhang, W. (1992). *The use of this-and-that: Part II. The Beta Journal.*

White (2003) recently developed a Pathfinder Network (PFNET) technique for weighted counting in ACA, where nodes are used to represent authors on a co-citation map and links are drawn in between the author nodes to represent significantly high (weighted) co-citation counts. With the inclusion of such weighted links, the network is reduced to show “only the most salient relationships,” for example, “those with many links to other authors [representing] a high degree centrality“ (p. 423).

First-author co-citations, under another *weighted counting* scheme may also be weighted more heavily than secondary author co-citations. For example, in the reference list shown in example i the first two references might yield Smith and Thomas a co-citation count of 2 (each), and the last two references, with Smith and Thomas listed as secondary authors, might be given only a standard co-citation count of 1. Note that this is just an example. Many other weighting schemes are feasible. In the case of mega-authorship, it might even be advisable, if only on practical grounds, to give some of the co-authors a weighted count equal to zero.

Example i

Smith, K. (1990). *A splendid result. The Alpha Journal.*  
 Thomas, T., & Zhang, W. (1992). *The use of this-and-that: Part II. The Beta Journal.*  
 Rao, S., & Thomas, T. (1994). *Problems with the k-procedure. The Gamma Journal.*  
 Janssens, J., & Smith, K. (1996). *The breakthrough. New World: Delta Publishers.*

Same-Author Co-citations

Finally, the A-A co-citation, or *same-author co-citation*, deserves special mention, particularly in cases where A is not a co-author with B or with any other authors in question

for an ACA. Example c, for instance (also shown in Pure First-Author co-citations) yields one Thomas-Thomas co-citation.

Example c

Smith, K. (1990). *A splendid result. The Alpha Journal.*  
 Thomas, T. (1991). *The use of this-and-that: Part I. The Beta Journal.*  
 Thomas, T., & Zhang, W. (1992). *The use of this-and-that: Part II. The Beta Journal.*

The *same-author co-citation* is relevant to the current diagonal problem in ACA, which focuses technically on what type of data should be added to the diagonal cells of an ACA matrix: missing, scaled, or statistically complete data (Ahlgren et al., 2003). In ACA, the key to observing intellectual similarities between author pairs rests on the collection of inter-author co-citation counts. For this reason, a researcher may omit retrieving same-author co-citation counts (because authors are already intellectually similar to themselves) and leave the cells along the diagonal empty. On the other hand, data based on the number of times an author has been co-cited with him/herself (excluding self-citations) yields a more mathematically complete co-citation matrix, which is statistically easier to study (Ahlgren et al., 2003). A new retrieval method dedicated to this form of co-cited author count might, therefore, be of use in solving the diagonal problem.

Importance of These Different Forms

Clearly, every article that results from the collaborative work of different authors, and that is cited at least once in the dataset under investigation is captured when collecting *co-authorship co-citations*. Hence, the resulting co-citation network includes an underlying co-authorship network. *General co-citation* counts take co-authorship into account, but *pure co-citation* counts are not intended to include co-authorships. A pure co-citation network might be totally different from the underlying co-authorship network; thus, both have the potential of leading to a complementary view of a research field under investigation. Table 1 includes brief definitions of the four co-citation forms again, and describes why a researcher may use one of the four forms for a particular ACA.

First-Author Co-citations as Used in Most ACA Studies

In a *first-author co-citation*, as used in most ACA studies, at least one publication written by A as the first or sole author and one publication written by B as the first or sole author co-occurs in the reference list of an article. This co-occurrence of the surnames in one article gives only one *first-author co-citation* (a binary procedure) for the pair (A, B) and the data collection technique, which uses the authors’ names as input to Dialog™ (e.g., CA = Smith K? and

TABLE 1. Author co-citation forms and implications for their use.

Co-citation form	Implications for use
<p><b>Pure first-author co-citation</b></p> <ul style="list-style-type: none"> <li>• One publication with A as first or sole author and one publication with B as first or sole author co-occur in the reference list of an article</li> </ul>	<ul style="list-style-type: none"> <li>• Suitable for developing a “thematic” picture of a research area or subject specialty</li> <li>• Relationships among first/sole authors usually make sense (high citation counts are necessary) and are described via clusters based on intellectual similarities</li> </ul>
<p><b>Pure author co-citation</b></p> <ul style="list-style-type: none"> <li>• One publication with A as co-author (regardless of rank order) and one publication with B as co-author (regardless of order) co-occur in the reference list of an article</li> </ul>	<ul style="list-style-type: none"> <li>• Supports the view that authors, regardless of their overall authorship ranking, can contribute substantially to the development of a research area</li> <li>• Presents a more accurate portrayal of an individual author’s contribution to a research area where high rates of co-authorship are prevalent (e.g., natural sciences, physics, chemistry)</li> <li>• Useful in the evaluation of an individual’s research performance over time (e.g., for faculty promotion)</li> </ul>
<p><b>General author co-citation</b></p> <ul style="list-style-type: none"> <li>• One publication with A (any rank) as co-author and one publication with B (any rank) as a co-author co-occur in reference list, including articles co-authored by A and B</li> </ul>	<ul style="list-style-type: none"> <li>• Again, presents a more accurate portrayal of an individual author’s contribution to a research field where high rates of co-authorship are prevalent</li> </ul>
<p><b>Co-author/co-citation</b></p> <ul style="list-style-type: none"> <li>• Same definition of general co-citation (above) but with special count for a co-authored (A &amp; B) paper, which recognizes the intellectual link between authors due to their collaborative work</li> </ul>	<ul style="list-style-type: none"> <li>• Creates a structure that examines more precisely the intellectual similarities among authors because intellectual similarity is recognized also as a function of co-authorship</li> <li>• Special co-authorship co-citation counts may be included in the ACA matrix of all co-citation counts or may be counted separately and mapped on top of a general co-citation network to show “connector lines” between authors and demonstrate the strength of various co-authorship ties</li> </ul>

CA = Thomas T?), forms the basis of what is known as a “traditional ACA” (see White’s 1986 article “Co-cited Author Retrieval”).

A simple first-author search carried out in a traditional ACA yields all *pure first-author co-citations*, but this search may also yield some *general co-citations* and some *co-author/co-citations*. Examples f and h below illustrates such cases.

Example h

Smith, K., & Thomas, T. (1992). *The k-procedure. The Gamma Journal.*

Thomas, T., & Smith, K. (1992). *More details about the k-procedure. The Gamma Journal.*  
 Thomas, T., & Smith, K. (1999). *A review of the k-procedure. The Gamma Journal.*  
 Zhang, W. (1999). *A local k-procedure. The Phi Journal.*  
 Rao, S., & Janssens, J. (1998). *A note on the m-procedure. The Iota Journal.*

Example f

Smith, K., & Thomas, T. (1992). *The k-procedure. The Gamma Journal.*  
 Thomas, T., & Smith, K. (1992). *More details about the k-procedure. The Gamma Journal.*  
 Rao, S., & Thomas, T. (1994). *Problems with the k-procedure. The Gamma Journal.*  
 Janssens, J., & Smith, K. (1996). *The breakthrough. New World: Delta Publishers.*

Example d below demonstrates that with the simple first-author retrieval technique (i.e., CA = Smith K? and CA = Thomas T?), a co-citation count with Smith and Thomas ranked in secondary positions (the *pure co-citation form*) would never be retrieved (see Fig. 2).

Example d

Janssens, J., & Smith, K. (1996). *The breakthrough. New World: Delta Publishers.*  
 Peters, P., Zhang, W., & Thomas, T. (1999). *The breakthrough revisited. The Alpha Journal.*

Given that a traditional search for first-author co-citation counts contributes only partially to a pure or general ACA, an alternative retrieval method is needed so that all author co-citations may be counted regardless of an author’s rank. Harsanyi (1993, p. 329) explains that “complete counts for a given author require obtaining a complete bibliography of that individual’s work and adding citations found under various first authors to the count for which the individual is the first or only author. At best this is tedious, and at worst (if a bibliography is not available) it is impossible.” Persson (2001) also states that “a full scale test [of co-citedness] is unrealistic since it means tracing all authors of several thousand cited documents” (p. 339). Among those who have

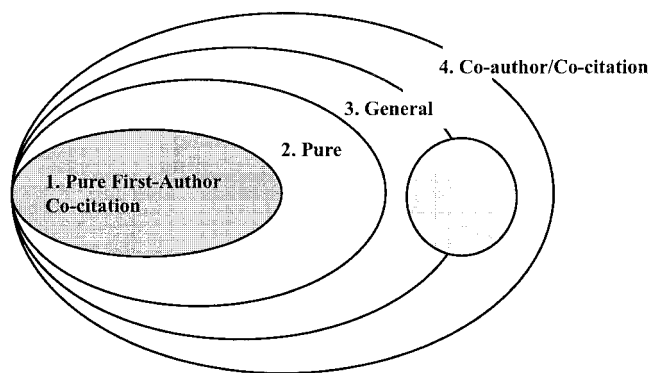


FIG. 2. Retrieval relationship (shown in gray) between simple first-author searches and other co-citation searches in the hierarchy.

TABLE 2. Complete publication lists of authors A and B.

A	B
A X Y, xxxxx	B, xxxxx
X A Z H, xxxxx	B Z, xxxxx
T Y H A, xxxxx	T A S B, xxxxx
T A S B, xxxxx	U B, xxxxx

carried out a traditional ACA, the first-author co-citation form is considered adequate, since the main purpose “is not to rank authors but rather to identify research themes” (p. 343). The first author co-citation form is most useful then for illustrating “overall trends or dimensions in scholars’ approach toward research” and providing a “general historical view of the intellectual structure of a research area” (McCain, 1990b, p. 213).

Nevertheless, if an ACA is designed to determine the historical development and thematic structure of a field, our view is that it does not make sense conceptually to omit certain authors (i.e., secondary authors), particularly if they have been participants all along in the development of that structure. Some authors who make significant contributions to a research field collaborate frequently with other members of their field and this is a fact that should not be ignored. McCain (1988) previously recognized this problem “in the natural sciences, where multiple authorship is common,” especially in fields (e.g., fly genetics) where “multiple authorship is correlated with perceived quality and greater utility” (pp. 428–429).

Persson (2001) recently compared the structure of a first-author co-citation map to that of an all-author co-citation map of information studies, using a special search key based on a list of source items from the Social Sciences Citation Index CDE editions 1986–1996. This search key “included the following subparts: first author’s last name, publication year, volume number, and starting page” (p. 340) (e.g., AU-Leggate P; Dyer H and JN-Electronic Library, 1986, V4, N3, P152–165). What he found as a result of his two-map comparison was that the change from first author citations to all author citations allowed new names to enter the map, while others had to leave. Essentially, it is a “risky business to rank the most influential researchers by first author citations only” (p. 342). New authors who were added to the information science structure probably should have been there in the first place and were not included because they simply had been ranked for non-essential (e.g., alphabetical) reasons. Eom and Farris’ (1996) earlier study of the decision support systems (DDS) field demonstrates another successful all-author co-citation study. The distinguishing aspect of their approach is that “a Fox-based matrix generation system was developed to . . . give access to cited co-authors as well as first authors.” This specialized system was able to “compute a co-citation frequency between any pair of authors” (p. 943). Not all researchers, however, have access to or experience in developing such systems. In cases where first author counts are used, then it

is clear that the counts may not necessarily reflect selected authors’ level of contribution. An ACA must, therefore, be open to the inclusion of all authors, regardless of their co-authorship rank, and must be open to different retrieval and counting methods other than the traditional first-author binary count that is currently used.

### How to Determine a Co-citation Matrix with Existing Dialog™ Software Capability?

Assume we want to determine the number of times authors A and B are co-cited in the ISI-database(s) (e.g., *Science Citation Index*, *Social Sciences Citation Index*, *Arts & Humanities Citation Index*, *Web of Knowledge*). First, we have to determine which kind of co-citation counts we want to consider: traditional first-author co-citations, or pure first-author ones, or pure and/or general co-citations? Since the traditional first-author case is well known, we will not examine it, but focus instead on the other co-citation forms of counting.

In addition to choosing the kind of co-citation (according to our classification scheme), we may also want to introduce other constraints. Perhaps, we are only interested in publications dealing with a certain topic, or articles published in journals covered by ISI, or articles published during a certain time span, or articles cited during a fixed citation window. All these constraints are likely to complicate the retrieval process considerably.

The first step in an ACA exercise would be to compile a complete publication list for A, and for B. Complete here means taking into account the restrictions noted above, for instance, all publications in ISI journals published during

TABLE 3. Example publication lists for authors P. S. ASPINWALL and D. R. MORRISON.<sup>a</sup>

A: P. S. ASPINWALL	B: D. R. MORRISON
1. <b>Aspinwall, P.S.</b> Enhanced gauge symmetries and $K3$ surfaces. <i>Phys. Lett. B</i> 357 (1995), no. 3, 329–334.	5. <b>Morrison, D.R.;</b> Vafa, C. Compactifications of $F$ -theory on Calabi-Yau threefolds. I. <i>Nuclear Phys. B</i> 473 (1996), no.1–2, 74–92.
2. <b>Aspinwall, P.S.,</b> Greene, B.R. On the geometric interpretation of $N = 2$ superconformal theories. <i>Nuclear Phys. B</i> 437 (1995), no. 1, 205–227.	6. <b>Morrison, D.R.,</b> Vafa, C. Compactifications of $F$ -theory on Calabi-Yau threefolds. II. <i>Nuclear Phys. B</i> 476 (1996), no. 3, 437–469.
3. <b>Aspinwall, P.S., Morrison, D.R.</b> $S^1$ -duality and integral structures. <i>Phys. Lett. B</i> 355 (1995), no. 1–2, 141–149.	7. <b>Aspinwall, P.S., Morrison, D.R.</b> $S^1$ -duality and integral structures. <i>Phys. Lett. B</i> 355 (1995), no. 1–2, 141–149.
4. <b>Aspinwall, P.S.</b> Point-like instantons and the $(\text{Spin}(32)/Z_2)$ heterotic string. <i>Nuclear Phys. B</i> 496 (1997), no. 1–2, 149–176.	8. <b>Distler, J., Greene, B.R., Morrison, D.R.</b> Resolving singularities in $(0,2)$ models. <i>Nuclear Phys. B</i> 481 (1996), no. 1–2, 289–312.

<sup>a</sup>Note that due to the mathematical content of the MathSci database, all of the AMS-LaTeX typesetting symbols (e.g.,  $K3$ ) have been retained.

TABLE 4. Search Sets for T<sub>A</sub> = ASPINWALL PS and T<sub>B</sub> = MORRISON DR.

File 34: *SciSearch* (R) Cited Ref Sci 1990-2003/Jan W4  
(c) 2003 Inst for Sci Info  
File 434: *SciSearch* (R) Cited Ref Sci 1974-1989/Dec  
(c) 1998 Inst for Sci Info

Set	Items	Description
? e	CR=ASPINWALL PS, 1995, V437, P205, NUCL PHYS B	
Ref	Items	Index-term
E1	3	CR=ASPINWALL PS, 1995, V369, P233, PHYS LETT B
E2	1	CR=ASPINWALL PS, 1995, V37, P329, PHYS LETT B
E3	21	*CR=ASPINWALL PS, 1995, V437, P205, NUCL PHYS B
E4	1	CR=ASPINWALL PS, 1995, V443, P329, PHYS LETT B
Enter P or PAGE for more		
? s e3		
S1	21	CR='ASPINWALL PS, 1995, V437, P205, NUCL PHYS B'
? e	CR=ASPINWALL PS, 1995, V357, P329, PHYS LETT B	
Ref	Items	Index-term
E1	2	CR=ASPINWALL PS, 1995, V357, P141, PHYS LETT B
E2	1	CR=ASPINWALL PS, 1995, V357, P329, PHYS LETT
E3	98	*CR=ASPINWALL PS, 1995, V357, P329, PHYS LETT B
E4	3	CR=ASPINWALL PS, 1995, V369, P233, PHYS LETT B
Enter P or PAGE for more		
? s e3		
S2	98	CR='ASPINWALL PS, 1995, V357, P329, PHYS LETT B'
? e	CR=ASPINWALL PS, 1995, V355, P141, PHYS LETT B	
Ref	Items	Index-term
E1	1	CR=ASPINWALL PS, 1995, P352, 1994 SUMM SCH HIGH E
E2	1	CR=ASPINWALL PS, 1995, V141, P141, PHYS LETT B
E3	22	*CR=ASPINWALL PS, 1995, V355, P141, PHYS LETT B
E4	2	CR=ASPINWALL PS, 1995, V357, P141, PHYS LETT B
Enter P or PAGE for more		
? s e3		
S3	22	CR='ASPINWALL PS, 1995, V355, P141, PHYS LETT B'
? e	CR=ASPINWALL PS, 1997, V496, P149, NUCL PHYS B	
Ref	Items	Index-term
E1	1	CR=ASPINWALL PS, 1997, V1, P355, STUDIES ADV MATH
E2	1	CR=ASPINWALL PS, 1997, V2, P703, MIRROR SYMMETRY
E3	21	*CR=ASPINWALL PS, 1997, V496, P149, NUCL PHYS B
E4	1	CR=ASPINWALL PS, 1997, V503, P33, NUCL PHYS B
Enter P or PAGE for more		
? s e3		
S4	21	CR='ASPINWALL PS, 1997, V496, P149, NUCL PHYS B'
? e	CR=MORRISON DR, 1996, V473, P74, NUCL PHYS B	

TABLE 4. (continued)

Ref	Items	Index-term
E1	1	CR=MORRISON DR, 1996, V46, P177, NUCL PHYS S
E2	1	CR=MORRISON DR, 1996, V473, P122, NUCL PHYS B
E3	144	*CR=MORRISON DR, 1996, V473, P74, NUCL PHYS B
E4	1	CR=MORRISON DR, 1996, V476, P4237, NUCL PHYS B
Enter P or PAGE for more		
? s e3		
S5	144	CR='MORRISON DR, 1996, V473, P74, NUCL PHYS B'
? e CR='MORRISON DR, 1996, V476, P437, NUCL PHYS B'		
Ref	Items	Index-term
E1	144	CR=MORRISON DR, 1996, V473, P74, NUCL PHYS B
E2	1	CR=MORRISON DR, 1996, V476, P4237, NUCL PHYS B <sup>a</sup>
E3	142	*CR=MORRISON DR, 1996, V476, P437, NUCL PHYS B
E4	3	CR=MORRISON DR, 1996, V483, P229, NUCL PHYS B
Enter P or PAGE for more		
? s e3		
S6	142	CR='MORRISON DR, 1996, V476, P437, NUCL PHYS B'
? e CR=ASPINWALL PS, 1995, V355, P141, PHYS LETT B <i>[Morrison is ranked as the second author of this paper]</i>		
Ref	Items	Index-term
E1	1	CR=ASPINWALL PS, 1995, P352, 1994 SUMM SCH HIGH E
E2	1	CR=ASPINWALL PS, 1995, V141, P141, PHYS LETT B
E3	22	*CR=ASPINWALL PS, 1995, V355, P141, PHYS LETT B
E4	2	CR=ASPINWALL PS, 1995, V357, P141, PHYS LETT B
Enter P or PAGE for more		
? s e3		
S7	22	CR='ASPINWALL PS, 1995, V355, P141, PHYS LETT B'
? e CR=DISTLER J, 1996, V481, P289, NUCL PHYS B		
Ref	Items	Index-term
E1	17	CR=DISTLER J, 1995, V442, P64, NUCL PHYS B
E2	1	CR=DISTLER J, 1995, 1994 SUMM SCH HIGH E
E3	14	*CR=DISTLER J, 1996, V481, P289, NUCL PHYS B
E4	1	CR=DISTLER J, 1997, V15, P8041, NUCLEIC ACIDS RES
Enter P or PAGE for more		
? s e3		
S8	14	CR='DISTLER J, 1996, V481, P289, NUCL PHYS B'
? ds		
Set	Items	Description
S1	21	CR='ASPINWALL PS, 1995, V437, P205, NUCL PHYS B'
S2	98	CR='ASPINWALL PS, 1995, V357, P329, PHYS LETT B'
S3	22	CR='ASPINWALL PS, 1995, V355, P141, PHYS LETT B'
S4	21	CR='ASPINWALL PS, 1997, V496, P149, NUCL PHYS B'
S5	144	CR='MORRISON DR, 1996, V473, P74, NUCL PHYS B'
S6	142	CR='MORRISON DR, 1996, V476, P437, NUCL PHYS B'
S7	22	CR='ASPINWALL PS, 1995, V355, P141, PHYS LETT B'
S8	14	CR='DISTLER J, 1996, V481, P289, NUCL PHYS B'



TABLE 4. (continued)

? s s1 or s2 or s3 or s4		
	21	S1
	98	S2
	22	S3
	21	S4
S9	149	S1 OR S2 OR S3 OR S4
? s s5 or s6 or s7 or s8		
	144	S5
	142	S6
	22	S7
	14	S8
S10	196	S5 OR S6 OR S7 OR S8
$T_A (\text{ASPINWALL}) = S1 \cup S2 \cup S3 \cup S4$ $\#(T_A (\text{ASPINWALL})) = 149$		
$T_B (\text{MORRISON}) = S5 \cup S6 \cup S7 \cup S8$ $\#(T_B (\text{MORRISON})) = 196$		

<sup>a</sup>Reference E2 has been underlined to indicate that P4237 is an indexing error.

the period 1995–2000, or perhaps “all publications ever written.” This leads to the construction of lists similar to those shown in Table 2. In Table 2, A, B, X, Y, etc., symbolize authors, while xxxxxx represents other types of bibliographical data (i.e., title, journal name, volume number, page number, etc.)

With the lists composed, we can then search for each publication as a Cited Reference (CR), using complete bibliographic data, in Dialog<sup>TM</sup> format. Ideally, the search should be carried out with no errors (e.g., mistyped volume) or omissions (e.g., volume is not mentioned) in the database, but this is often not the case. In addition to volume-type errors, Rice, Borgman, Bednarski, & Hart (1989, p. 267) point to an indexing error that can occur with journal titles: they are “sometimes abbreviated in unique ways that inhibit accurate identification of the journal name.” For instance, the authors show how the title *Inform Processi* may represent *Information Processing*, a non-core journal in the fields of Communication, Information Science, and Library Science or *Information Processing & Management*, which is a core journal.

Once the CR search is performed, it leads to a set S1 consisting of all articles citing document *AXY,xxxx*, and further sets S2, S3, and S4 consisting of all articles citing documents *XAZH,xxxxxxx*, *TYHA,xxxxx*, and *TASB,xxxxx*. Under author B, S5 consists of all articles citing document *B,xxxx* and so on (for all A and all B publications).

To illustrate the procedure further, we present Table 3 with a “real” example of a short (incomplete) publication list of two authors: P. S. ASPINWALL and D. R. MORRISON. Both authors are theoretical physicists and the bibliographic data for their lists has been extracted from Dialog<sup>TM</sup> MathSci.

To begin the co-citation procedure, we use the EXPAND command provided by Dialog<sup>TM</sup> to search for the cited

reference (CR) of each listed article. We then use the CR data to form the set  $T_A = S1 \cup S2 \cup S3 \cup S4$ , consisting of all articles that cite at least one of the A-publications (ASPINWALL). Similarly, we form the set  $T_B = S5 \cup S6 \cup S7 \cup S8$  consisting of all articles that cite at least one of the B-publications (MORRISON). The results of the search procedure in Dialog<sup>TM</sup> for both ASPINWALL and MORRISON are shown in Table 4. Note from Table 4 that indexing inconsistencies within Dialog<sup>TM</sup> can result in missed reference counts, unless one takes a meticulous approach to incorporating them. For the list of publications leading to Set 2 (S2), an indexing error at E2 has been found and underlined to show that the letter “B” in PHYS LETT B has been omitted from CR = ASPINWALL PS, 1995, V357, P329, PHYS LETT. We allow for this small error in the search procedure, but draw attention to it for tutorial purposes.

Next, we form  $U = T_A \cap T_B$ , consisting of all articles that co-cite (at least) one A-publication (ASPINWALL) and (at least) one B-publication (MORRISON).

Generally, if the same article appears in the first list of A-publications, and occurs also in the second list of B-publications, then two of the S-sets are identical. In this case, we know from Table 3 that ASPINWALL and MORRISON share one co-authored article in their respective publication sets: S3 is equal to S7, therefore, it is a part of the intersection (the U-set).

Note from the Dialog<sup>TM</sup> search shown in Table 5, that there is a total of 43 articles in *SciSearch* that co-cite at least one ASPINWALL publication with at least one MORRISON publication. Among these 43 articles, one bibliographic record, specifically number 6 (“11/TIAU,CR/6”) is an example of a co-author co-citation score resulting from a co-authorship, and not from a “real” co-citation of these two authors. This co-authored article may be given one count

TABLE 5. Search results for  $U = T_A \text{ (ASPINWALL)} \cap T_B \text{ (MORRISON)}$ .

File 34: *SciSearch* (R) Cited Ref Sci 1990-2003/Jan W4  
 File 434: *SciSearch* (R) Cited Ref Sci 1974-1989/Dec (c) 1998 Inst  
 for Sci Info

Set	Items	Description
? s	CR=ASPINWALL PS, 1995, V437, P205, NUCL PHYS B	
S1	21	CR=ASPINWALL PS, 1995, V437, P205, NUCL PHYS B
? s	CR=ASPINWALL PS, 1995, V357, P329, PHYS LETT B	
S2	98	CR=ASPINWALL PS, 1995, V357, P329, PHYS LETT B
? s	CR=ASPINWALL PS, 1995, V355, P141, PHYS LETT B	
S3	22	CR=ASPINWALL PS, 1995, V355, P141, PHYS LETT B
? s	CR=ASPINWALL PS, 1997, V496, P149, NUCL PHYS B	
S4	21	CR=ASPINWALL PS, 1997, V496, P149, NUCL PHYS B
? s	CR=MORRISON DR, 1996, V473, P74, NUCL PHYS B	
S5	144	CR=MORRISON DR, 1996, V473, P74, NUCL PHYS B
? s	CR=MORRISON DR, 1996, V476, P437, NUCL PHYS B	
S6	142	CR=MORRISON DR, 1996, V476, P437, NUCL PHYS B
? s	CR=ASPINWALL PS, 1995, V355, P141, PHYS LETT B	
S7	22	CR=ASPINWALL PS, 1995, V355, P141, PHYS LETT B
? s	CR=DISTLER J, 1996, V481, P289, NUCL PHYS B	
S8	14	CR=DISTLER J, 1996, V481, P289, NUCL PHYS B
? s	s1 or s2 or s3 or s4	
	21	S1
	98	S2
	22	S3
	21	S4
S9	149	S1 OR S2 OR S3 OR S4 [S9=T <sub>A</sub> (ASPINWALL)]
? s	s5 or s6 or s7 or s8	
	144	S5
	142	S6
	22	S7
	14	S8
S10	196	S5 OR S6 OR S7 OR S8 [S10=T <sub>B</sub> (MORRISON)]
? s	S9 and S10	
	149	S9
	196	S10
S11	43	S9 AND S10 [S11=U]

? t s11/ti, au, cr/all [view all 43 co-citations]<sup>a</sup>

11/TI, AU, CR/1 (Item 1 from file: 34)

DIALOG (R) File 34: (c) 2003 Inst for Sci Info. All rts. reserv.

Title: Codimension-three bundle singularities in F-theory - art. no.  
 014

Author(s): Candelas P (REPRINT); Diaconescu DE; Florea B; Morrison  
 DR; Rajesh G

Cited References:

- ASPINWALL PS, 1997, V496, P149, NUCL PHYS B
- DISTLER J, 1996, V481, P289, NUCL PHYS B
- MORRISON DR, 1996, V476, P437, NUCL PHYS B

11/TI, AU, CR/2 (Item 2 from file: 34)

DIALOG (R) File 34: (c) 2003 Inst for Sci Info. All rts. reserv.

Title: Reflexive polyhedra, weights and toric Calabi-Yau fibrations

Author(s): Kreuzer M (REPRINT); Skarke H

Cited References:

- ASPINWALL PS, 1995, V437, P205, NUCL PHYS B
- MORRISON DR, 1996, V476, P437, NUCL PHYS B
- MORRISON DR, 1996, V473, P74, NUCL PHYS B

TABLE 5. (continued)

11/TI, AU, CR/3 (Item 3 from file: 34)

DIALOG (R) File 34: (c) 2003 Inst for Sci Info. All rts. reserv.

Title: Type I on (generalized) Voisin-Borcea orbifolds and nonperturbative orientifolds

Author(s): Kakushadze Z (REPRINT)

Cited References:

ASPINWALL PS, 1997, V496, P149, NUCL PHYS B

ASPINWALL PS, 1995, V357, P329, PHYS LETT B

MORRISON DR, 1996, V473, P74, NUCL PHYS B

11/TI, AU, CR/4 (Item 4 from file: 34)

DIALOG (R) File 34: (c) 2003 Inst for Sci Info. All rts. reserv.

Title: K3 surfaces and string duality

Author(s): Aspinwall PS (REPRINT)

Cited References:

ASPINWALL PS, 1995, V355, P141, PHYS LETT B

ASPINWALL PS, 1995, V357, P329, PHYS LETT

MORRISON DR, 1996, V473, P74, NUCL PHYS B

MORRISON DR, 1996, V476, P437, NUCL PHYS B

11/TI, AU, CR/5 (Item 5 from file: 34)

DIALOG (R) File 34: (c) 2003 Inst for Sci Info. All rts. reserv.

Title: On elevating free-fermion  $Z(2) \times Z(2)$  orbifolds models to compactifications of F theory

Author(s): Berglund P (REPRINT); Ellis J; Faraggi AE; Nanopoulos DV; Qiu Z

Cited References:

ASPINWALL PS, 1997, V496, P149, NUCL PHYS B

MORRISON DR, 1996, V473, P74, NUCL PHYS B

MORRISON DR, 1996, V476, P437, NUCL PHYS B

11/TI, AU, CR/6 (Item 6 from file: 34)

DIALOG (R) File 34: (c) 2003 Inst for Sci Info. All rts. reserv.

Title: T-duality can fail

Author(s): Aspinwall PS (REPRINT); Plesser MR

Cited References:

ASPINWALL PS, 1995, V355, P141, PHYS LETT B

*(1 co-author/co-citation for the following co-authored paper:**Aspinwall, P.S., Morrison, D.R.  $S^1$ -duality and integral structures. Phys. Lett. B 355 (1995), no. 1-2, 141-149.)*

$$\#(U) = \#(T_A \text{ (ASPINWALL)} \cap T_B \text{ (MORRISON)}) = 43$$

---

 \*Items 7-43 have been excluded.

based on our definition of the co-author co-citation, which recognizes the intellectual link between the authors due to their collaborative work.

### How to Find Pure First-Author, Pure, or General Co-citations?

Suppose we are not interested in counting co-author co-citations and want only the number of *pure first-author*, *pure co-citations*, or *general co-citations*? To obtain the *pure first-author co-citations*, we remove all articles from the two lists of publications (the A-list and the B-list) where

A and B are identified as co-authors. Also, as shown in Table 6, we remove all articles where ASPINWALL or MORRISON is not the first author (see Table 6).

This restriction gives T-sets consisting of all articles that cite at least one first-authored article by A and one first-authored article by B, so that  $T_A = S1 \cup S2 \cup S4$ , and  $T_B = S5 \cup S6$ . The cited A article, moreover, is certainly not co-authored by B and vice versa. An article in the corresponding U-set is thus one that co-cites at least one A first-authored article, and one B first-authored article, and does not include any co-authored publications by A and B.

TABLE 6. Restricted publication lists for authors P. S. ASPINWALL and D. R. MORRISON, used for the calculation of pure first-author co-citations.

A: P. S. ASPINWALL	B: D. R. MORRISON
1. <b>Aspinwall, P.S.</b> , Enhanced gauge symmetries and $K3$ surfaces. <i>Phys. Lett. B</i> 357 (1995), no. 3, 329–334.	5. <b>Morrison, D.R.</b> , Vafa, C. Compactifications of $F$ theory on Calabi-Yau threefolds. I. <i>Nuclear Phys. B</i> 473 (1996), no. 1–2, 74–92.
2. <b>Aspinwall, P.S.</b> , Greene, B.R. On the geometric interpretation of $N = 2$ superconformal theories. <i>Nuclear Phys. B</i> 437 (1995), no. 1, 205–227.	6. <b>Morrison, D.R.</b> , Vafa, C., Compactifications of $F$ -theory on Calabi-Yau threefolds. II. <i>Nuclear Phys. B</i> 476 (1996), no. 3, 437–469.
3. <b>Aspinwall, P.S.</b> , Morrison, D.R., $S$ -duality and integral structures. <i>Phys. Lett. B</i> 355 (1995), no. 1–2, 141–149.	7. <b>Aspinwall, P.S.</b> , Morrison, D. R., $S$ -duality and integral structures. <i>Phys. Lett. B</i> 355 (1995), no. 1–2, 141–149.
4. <b>Aspinwall, P.S.</b> , Point-like instantons and the $\text{Spin}(32)/\mathbb{Z}_2$ heterotic string. <i>Nuclear Phys. B</i> 496 (1997), no. 1–2, 149–176.	8. Distler, J., Greene, B.R., Morrison, D.R., Resolving singularities in $(0,2)$ models. <i>Nuclear Phys. B</i> 481 (1996), no. 1–2, 289–312.

Note. Numbering is kept the same as in Table 3 and removed publications are highlighted.

Table 7 presents the *pure first-author co-citation* data collection procedure, which was carried out in Dialog™ *SciSearch*. With the intersection of the T-sets, we arrive at a total of 20 co-citation counts. This means that there are 20 pure first-author co-citations for ASPINWALL and MORRISON resulting from the small set of publications.

For the coupling of all *pure author co-citations*, all A and B co-authored articles are removed again from the A and B publication lists. However, if A or B are ranked as secondary authors in joint papers with “other” authors (C, D, E, etc), the restriction on the T-sets no longer applies. This gives T-sets consisting of all articles that cite at least one first or co-authored article, for instance, by A. This cited A article, however, is not co-authored by B (by the removal of co-authored articles). An article in the corresponding U-set is thus one that co-cites at least one A-authored article, and one B-authored article, but does not include any articles co-authored by A and B.

In Table 8, the common co-author set for ASPINWALL and MORRISON has been highlighted again to indicate that it has been excluded. The publication for MORRISON, who is listed as a secondary author to DISTLER and GREENE, is now included in the B publication set. Note that ASPINWALL is always the first-named author; therefore, we do not show an example where he is secondary author.

Table 9 presents the *pure author co-citation* data collection procedure, which was carried out in Dialog™ *SciSearch*. With the intersection of the T-sets, we arrive at a total of 22 pure author co-citation counts. Records 1 and 2 provide examples of pure first-author co-citations. Yet,

record 1 is at the same time an example of a *pure author co-citation* where at least one article written by ASPINWALL and one written by MORRISON are co-cited together *regardless of their ranked author order*. In the Cited References portion of this record, we see that ASPINWALL’s paper no. 4 was co-cited with the DISTLER, GREEN, and MORRISON paper no. 8 as well as with another first-author MORRISON paper no. 6. The resulting co-citation count, however, for this record is only one, based on our binary counting rule. This count may be associated with either the co-citation of articles 4 and 6 or articles 4 and 8.

Finally, obtaining *general co-citations* is the most elaborate procedure. First, one obtains the T sets  $T_A$  and  $T_B$  as in the pure co-citation case. Then, one determines S sets for all co-authored publications:  $S_{AB1}$ ,  $S_{AB2}$ ,  $S_{AB3}$ , . . . etc. These are the sets of articles citing co-authored article  $AB_j$ ,  $j = 1, 2, 3 . . .$  etc. (note that in the case of ASPINWALL and MORRISON, there is only one co-authored article  $S_{AB1}$  common to the two authors’ publication lists). In addition, we form the union of all these S sets, and denote this new set by  $T_{AB}$ . Thus,  $T_{AB}$  consists of all articles that cite, among other ones, at least one A-B co-authored article (see Table 9).

The next step, shown in Table 10, is to form the sets  $U_1 = T_A \cap T_B$ ,  $U_2 = T_A \cap T_{AB}$  and  $U_3 = T_B \cap T_{AB}$ . The set  $U_1$  is the U-set obtained for the pure co-citation case. The set  $U_2$  consists of all articles that co-cite at least one article co-authored by A (but certainly not by B) and an article co-authored by A and B. Similarly,  $U_3$  consists of all articles that co-cite at least one article co-authored by B (but certainly not by A) and an article co-authored by A and B. For the final result, a search is carried out for  $U = U_1 \cup U_2 \cup U_3$  and the number of elements in this U-set is equal to the number of A-B general co-citations (see Table 9). In our example, we find a total of 31 general co-citations.

## Conclusion

The term author co-citation has been defined more precisely than was previously done, and classified according to four distinct forms: the *pure first-author co-citation*, the *pure author co-citation*, the *general author co-citation*, and the *special co-author/co-citation*. This distinction removes the ambiguity of a simple first-author co-citation search. We argue that an author’s contribution to a selected field of study should not be limited to first authorship, but should be based on his/her complete list of publications, regardless of author ranking. An all-author co-citation study based on the *pure or general co-citation forms* supports a more accurate portrayal of an individual author’s contribution to a research area where high rates of co-authorship are prevalent (e.g., natural sciences, physics, chemistry). Such studies are also useful in evaluating more fairly an individual’s research performance over time (e.g., for faculty tenure). When special co-author co-citations are included in the thematic structure of a research field, the intellectual similarities

TABLE 7. Pure first-author co-citation search results for  $U = T_A (\text{ASPINWALL}) \cap T_B (\text{MORRISON})$ .

File 34: *SciSearch*(R) Cited Ref Sci 1990-2003/Jan W4 (c) 2003 Inst for Sci Info  
 File 434: *SciSearch*(R) Cited Ref Sci 1974-1989/Dec (c) 1998 Inst for Sci Info

Set	Items	Description
? s	CR=ASPINWALL PS, 1995, V357, P329, PHYS LETT B	
S1	98	CR=ASPINWALL PS, 1995, V357, P329, PHYS LETT B
? s	CR=ASPINWALL PS, 1995, V437, P205, NUCL PHYS B	
S2	21	CR=ASPINWALL PS, 1995, V437, P205, NUCL PHYS B
? s	CR=ASPINWALL PS, 1997, V496, P149, NUCL PHYS B	
S3	21	CR=ASPINWALL PS, 1997, V496, P149, NUCL PHYS B
? s	CR=MORRISON DR, 1996, V473, P74, NUCL PHYS B	
S4	144	CR=MORRISON DR, 1996, V473, P74, NUCL PHYS B
? s	CR=MORRISON DR, 1996, V476, P437, NUCL PHYS B	
S5	142	CR=MORRISON DR, 1996, V476, P437, NUCL PHYS B
? s	S1 or S2 or S3	
	98	S1
	21	S2
	21	S3
S6	132	S1 OR S2 OR S3 [S6= $T_A$ (ASPINWALL)]
? s	S4 or S5	
	144	S4
	142	S5
S7	170	S4 OR S5 [S7= $T_B$ (MORRISON)]
? s	s6 and s7	
	132	S6
	170	S7
S8	20	S6 AND S7 [S8=U]
$\#(U) = \#(T_A (\text{ASPINWALL}) \cap T_B (\text{MORRISON})) = 20$		

between paired authors are recognized further because similarity is now seen as a function of co-authorship.

Examples were given for each author co-citation form defined in our classification scheme, including a final set of sample Dialog searches using full references extracted from the *SciSearch* database. Note that, as expected, the totals obtained in the four sample searches form an ascending sequence corresponding to the concentric levels illustrated in Figure 1:  $20 < 22 < 31 < 43$ .

The Dialog<sup>TM</sup> search procedures described in the previous section were relatively simple in that small, manageable sample publication lists were used for ASPINWALL and MORRISON to generate precise results. With the inclusion of additional authors' names and the development of complete publication lists, the Dialog<sup>TM</sup> search procedure should remain effective, but may require more focus on the part of the researcher at certain steps. For instance, when creating a set of authors' complete publication lists, all the preferred (if not exact) forms of the individual cited references (CR) must be determined first, using the "expand cited reference" command (e CR) in Dialog<sup>TM</sup>. In some cases, the researcher may find that one article has not only been cited many times over the years, but also in alternative cited reference (CR) formats. Goodrum, McCain, Lawrence, & Giles (2001)

TABLE 8. Example publication lists for authors P. S. ASPINWALL and D. R. MORRISON, with the removal (highlighted print) of all co-authored publications common to both lists and the inclusion of secondary-ranked author papers with "other" authors.

A: P. S. ASPINWALL	B: D. R. MORRISON
1. <b>Aspinwall, P.S.</b> Enhanced gauge symmetries and $K3$ surfaces. <i>Phys. Lett. B</i> 357 (1995), no. 3, 329-334.	5. <b>Morrison, D.R., Vafa, C.</b> Compactifications of $F$ -theory on Calabi-Yau threefolds. I. <i>Nuclear Phys. B</i> 473 (1996), no.1-2, 74-92.
2. <b>Aspinwall, P.S., Greene, B.R.</b> On the geometric interpretation of $N = 2$ superconformal theories. <i>Nuclear Phys. B</i> 437 (1995), no. 1, 205-227.	6. <b>Morrison, D.R., Vafa, C.</b> Compactifications of $F$ -theory on Calabi-Yau threefolds. II. <i>Nuclear Phys. B</i> 476 (1996), no. 3, 437-469.
3. <b>Aspinwall, P.S., Morrison, D.R.,</b> $S$ -duality and integral structures. <i>Phys. Lett. B</i> 355 (1995), no. 1-2, 141-149.	7. <b>Aspinwall, P.S., Morrison, D.R.</b> $S$ -duality and integral structures. <i>Phys. Lett. B</i> 355 (1995), no. 1-2, 141-149.
4. <b>Aspinwall, P.S.</b> Point-like instantons and the $\text{Spin}(32)/Z_2$ heterotic string. <i>Nuclear Phys. B</i> 496 (1997), no. 1-2, 149-176.	8. <b>Distler, J., Greene, B.R., Morrison, D.R.</b> Resolving singularities in $(0,2)$ models. <i>Nuclear Phys. B</i> 481 (1996), no. 1-2, 289-312.

TABLE 9. Pure co-cited author search results for  $U = T_{A(\text{ASPINWALL})} \cap T_{B(\text{MORRISON})}$ .

File 34: SciSearch (R) Cited Ref Sci 1990-2003/Jan W4  
(c) 2003 Inst for Sci Info  
File 434: SciSearch (R) Cited Ref Sci 1974-1989/Dec  
(c) 1998 Inst for Sci Info

Set	Items	Description
? s	CR=ASPINWALL PS, 1995, V437, P205, NUCL PHYS B	
S1	21	CR=ASPINWALL PS, 1995, V437, P205, NUCL PHYS B
? s	CR=ASPINWALL PS, 1995, V357, P329, PHYS LETT B	
S2	98	CR=ASPINWALL PS, 1995, V357, P329, PHYS LETT B
? s	CR=ASPINWALL PS, 1997, V496, P149, NUCL PHYS B	
S3	21	CR=ASPINWALL PS, 1997, V496, P149, NUCL PHYS B
? s	CR=MORRISON DR, 1996, V473, P74, NUCL PHYS B	
S4	144	CR=MORRISON DR, 1996, V473, P74, NUCL PHYS B
? s	CR=MORRISON DR, 1996, V476, P437, NUCL PHYS B	
S5	142	CR=MORRISON DR, 1996, V476, P437, NUCL PHYS B
? s	CR = DISTLER J, 1996, V481, P289, NUCL PHYS B	
S6	14	CR = DISTLER J, 1996, V481, P289, NUCL PHYS B
? s	s1 or s2 or s3	
	21	S1
	98	S2
	21	S3
S7	132	S1 OR S2 OR S3 [S7=T <sub>A</sub> (ASPINWALL)]
? s	s4 or s5 or s6	
	144	S4
	142	S5
	14	S6
S8	180	S4 OR S5 OR S6 [S8=T <sub>B</sub> (MORRISON)]
? s	s7 and s8	
	132	S7
	180	S8
S9	22	S7 AND S8 [S9=U]

? t s9/ti, au,cr/all [view all 22 co-citations]

9/TI, AU, CR/1 (Item 1 from file: 34)

DIALOG (R) File 34: (c) 2002 Inst for Sci Info. All rts. reserv.

Title: Codimension-three bundle singularities in F-theory-art.  
no. 014

Author(s): Candelas P (REPRINT); Diaconescu DE; Florea B;  
Morrison DR; Rajesh G

Cited References:

- ASPINWALL PS, 1997, V496, P149, NUCL PHYS B
- DISTLER J, 1996, c481, P289, NUCL PHYS B
- MORRISON DR, 1996, V476, P437, NUCL PHYS B

9/TI, AU, CR/2 (Item 2 from file: 34)

DIALOG (R) File 34: (c) 2002 Inst for Sci Info. All rts. reserv.

Title: Reflexive polyhedra, weights and toric Calabi-Yau fibrations

Author(s): Kreuzer M (REPRINT); Skarke H

Cited References:

- ASPINWALL PS, 1997, V496, P149, NUCL PHYS B
- ASPINWALL PS, 1995, V357, P329, PHYS LETT B
- MORRISON DR, 1996, V473, P74, NUCL PHYS B

$$\#(U) = \#(T_{A(\text{ASPINWALL})} \cap T_{B(\text{MORRISON})}) = 22$$

Note. Items 3-22 have been excluded.

TABLE 10. General author co-citation search results for  $U = [T_A (ASPINWALL) \cap T_B (MORRISON)] \cup [T_A (ASPINWALL) \cap T_{AB}(ASPINWALL-MORRISON)] \cup (T_B(MORRISON) \cap T_{AB}(ASPINWALL-MORRISON))$ .

File 34: SciSearch (R) Cited Ref Sci 1990-2003/Jan W4 (c) 2003 Inst for Sci Info

File 434: SciSearch (R) Cited Ref Sci 1974-1989/Dec (c) 1998 Inst for Sci Info

Set	Items	Description
S1	21	CR=ASPINWALL PS, 1995, V437, P205, NUCL PHYS B
S2	98	CR=ASPINWALL PS, 1995, V357, P329, PHYS LETT B
S3	21	CR=ASPINWALL PS, 1997, V496, P149, NUCL PHYS B
S4	132	S1 OR S2 OR S3

S4=TS4=T<sub>A</sub> (ASPINWALL)

Set	Items	Description
S5	144	CR=MORRISON DR, 1996, V473, P74, NUCL PHYS B
S6	142	CR=MORRISON DR, 1996, V476, P437, NUCL PHYS B
S7	14	CR=DISTLER J, 1996, V481, P289, NUCL PHYS B
S8	180	S5 OR S6 OR S7

S8=T<sub>B</sub> (MORRISON)

S9 22 CR=ASPINWALL PS, 1995, V355, P141, PHYS LETT B

S9=T<sub>AB</sub> (ASPINWALL-MORRISON)

? s s4 and s8

	132	S4
	179	S8
S10	22	S4 and S8 [S10 = U <sub>1</sub> = T <sub>A</sub> ∩ T <sub>B</sub> ]

? s s4 and s9

	132	S4
	22	S9
S11	5	S4 and S9 [S11 = U <sub>2</sub> = T <sub>A</sub> ∩ T <sub>AB</sub> ]

? s s8 and s9

	179	S8
	22	S9
S12	6	S8 and S9 [S12 = U <sub>3</sub> = T <sub>B</sub> ∩ T <sub>AB</sub> ]

? s s10 or s11 or s12

	22	S10
	5	S11
	6	S12
S18	31	S10 OR S11 OR S12 [S18 = U = U <sub>1</sub> ∪ U <sub>2</sub> ∪ U <sub>3</sub> ]

t s18/ti, au, cr/all[view all 31 co-citations]

18/TI, AU, CR/1 (Item 1 from file: 34)

DIALOG (R) File 34: (c) 2002 Inst for Sci Info. All rts. reserv.

Title: Codimension-three bundle singularities in F-theory - art. no. 014

Author(s): Candelas P (REPRINT); Diaconescu DE; Florea B; Morrison DR; Rajesh G

Cited References:

ASPINWALL PS, 1997, V496, P149, NUCL PHYS B  
 DISTLER J, 1996, V481, P289, NUCL PHYS B  
 MORRISON DR, 1996, V476, P437, NUCL PHYS B

18/TI, AU, CR/2 (Item 2 from file: 34)

DIALOG (R) File 34: (c) 2002 Inst for Sci Info. All rts. reserv.

TABLE 10. (continued)

Title: Reflexive polyhedra, weights and toxic Calabi-Yau fibrations

Author(s): Kreuzer M (REPRINT); Skarke H

Cited References:

ASPINWALL PS, 1995, V437, P205, NUCL PHYS B

MORRISON DR, 1996, V476, P437, NUCL PHYS B

MORRISON DR, 1996, V473, P74, NUCL PHYS B

18/TI, AU, CR/3 (Item 3 from file: 34)

DIALOG (R) File 34: (c) 2002 Inst for Sci Info. All rts. reserv.

Title: Type I on (generalized) Voisin-Borcea orbifolds and nonperturbative orientifolds

Author(s): Kakushadze Z (REPRINT)

Cited References:

ASPINWALL PS, 1997, V496, P149, NUCL PHYS B

ASPINWALL PS, 1995, V357, P329, PHYS LETT B

MORRISON DR, 1996, V473, P74, NUCL PHYS B

$$\#(U) = \#([T_{A(ASPINWALL)} \cap T_{B(MORRISON)}] \cup [T_{A(ASPINWALL)} \cap T_{(ASPINWALL-MORRISON)}] \cup [T_{B(MORRISON)} \cap T_{AB(ASPINWALL-MORRISON)}]) = 31$$

<sup>a</sup>Items 4–31 in this printout have been excluded.

point specifically to the cited reference (CR) format in *SciSearch*, which is based on an 80-character string, and state that “there is no standard format beyond these data elements and no attempt to regularize citations at the time of data entry” (p. 665). The result is “a number of subtly or radically different CR strings, which Howard White has called *allonyms* (White, 2001)” (Goodrum et al., 2001, p. 665). To correct this problem, the researchers relied on the CiteSeer database (for computer science publications). The CiteSeer normalization capability or algorithm basically extracts publication titles correctly by using multiple hypotheses to identify the first author. Varied journal titles and different editions of the same work are grouped together so that it is possible to track versions of the same document over time (Goodrum et al., 2001, p. 665).

In research fields where multi-authored or even mega-authored papers are common (e.g., high energy physics) and the alphabetical ranking of authors is not a standard procedure (e.g., medical trial reports), an author co-citation study is naturally more complex. Ideally, the author co-citation procedures proposed in this article should be automated.

## Acknowledgments

The authors acknowledge the support of the John Metcalfe Visitor’s Grant award to Ronald Rousseau in 2002. Collaborative work on this article began as a result of his visit to the School of Information Systems, Technology of Management, University of New South Wales.

## References

Ahlgren, P., Jarneving, B., & Rousseau, R. (2003). Requirements for a cocitation similarity measure, with special reference to Pearson’s corre-

lation coefficient. *Journal of the American Society for Information Science & Technology*, 54, 550–560.

Borgman, C.L., & Rice, R.E. (1992). The convergence of information science and communication: a bibliometric analysis. *Journal of the American Society for Information Science*, 43, 397–411.

Brooks, T. (1985). Private acts and public objects: An investigation of citer motivations. *Journal of the American Society for Information Science*, 34, 132–135.

Case, D.O., & Higgins, G.M. (2000). How can we investigate citation behaviour? A study of reasons for citing literature in communication. *Journal of the American Society for Information Science*, 51, 635–645.

Christensen, F.H., & Ingwersen, P. (1996). Online citation analysis. A methodological approach. *Scientometrics*, 37, 39–62.

Eom, S.B., & Farris, R.S. (1996). The contributions of organizational science to the development of decision support systems research specialties. *Journal of the American Society for Information Science*, 47(12), 941–952.

Goodrum, A.A., McCain, K.W., Lawrence, S., & Giles, C.L. (2001). Scholarly publishing in the Internet age: A citation analysis of computer science literature. *Information Processing and Management*, 37, 661–675.

Harsanyi, M.A. (1993). Multiple authors, multiple problems, bibliometrics and the study of scholarly collaboration: A literature review. *Library and Information Science Research*, 15, 325–254.

Ingwersen, P., & Christensen, F.H. (1997). Data set isolation for bibliometric online analyses of research publications: Fundamental methodological issues. *Journal of the American Society for Information Science*, 48, 205–217.

McCain, K. (1984). Longitudinal author co-citation mapping: The changing structure of macroeconomics. *Journal of the American Society for Information Science*, 35, 351–359.

McCain, K. (1985). Longitudinal author co-citation mapping: A test of congruence in two scientific literatures. Unpublished Ph.D. Dissertation, Drexel University, Philadelphia, PA.

McCain, K. (1988). Evaluating cocited author search performance in a collaborative specialty. *Journal of the American Society for Information Science*, 39, 428–431.

McCain, K.W. (1990a). Mapping authors in intellectual space: A technical overview. *Journal of the American Society for Information Science*, 41, 433–443.



- McCain, K. (1990b). Mapping authors in intellectual space: Population genetics in the 1980s. In C.L. Borgman (Ed.), *Scholarly communication and bibliometrics* (pp. 194–216). Newbury Park, CA: Sage.
- Moed, H.F., Burger, W.J.M., Frankfort, J.G., & Van Raan, A.F.J. (1985). The use of bibliometric data for the measurement of university research performance. *Research Policy*, 14, 131–149.
- Perry, C.A., & Rice, R.E. (1998). Scholarly communication in developmental dyslexia: Influence of network structure on change in a hybrid problem area. *Journal of the American Society for Information Science*, 49, 151–168.
- Persson, O. (2001). All author citations versus first author citations. *Scientometrics*, 50, 339–344.
- Rice, R.E., Borgman, C.L., Bednarski, D., & Hart, P.J. (1989). *Scientometrics*, 15(3–4), 257–282.
- Rousseau, R., & Zuccala, A. (2003). A classification of author-co-citations. In G. Jiang, R. Rousseau, & Y. Wu (Eds.), *Proceedings of the 9th International Conference on Scientometrics and Informetrics* (pp. 261–270). Dalian (China): Dalian University of Technology Press.
- Sandstrom, P.E. (1998). *Information foraging among anthropologists in the invisible college of human behavioral ecology: An author cocitation analysis*. Unpublished doctoral dissertation, Indiana University, Bloomington, Indiana.
- Small, H. (1999). Visualizing science by citation mapping. *Journal of the American Society for Information Science*, 50, 799–813.
- Small, H., & Griffith, B. (1974). The structure of scientific literatures I: identifying and graphing specialties. *Science Studies*, 4, 17–40.
- Spinak, E. (1996). *Diccionario Enciclopédico de Bibliometría e Informetría*. Caracas: Unesco.
- White, H.D. (1986). Co-cited author retrieval. *Information Technology and Libraries*, 5, 93–99.
- White, H.D. (1990). Author cocitation analysis: Overview and defense. In C.L. Borgman (Ed.), *Scholarly communication and bibliometrics* (pp. 84–106). Newbury Park, CA: Sage.
- White, H.D. (2001). Authors as citers over time. *Journal of the American Society for Information Science*, 52(2), 87–108.
- White, H.D. (2003). Pathfinder networks and author cocitation analysis: A remapping of paradigmatic information scientists. *Journal of the American Society for Information Science and Technology*, 54(5), 423–434.
- White, H.D., & Griffith, B.C. (1981a). Author cocitation: A literature measure of intellectual structure. *Journal of the American Society for Information Science*, 32, 163–172.
- White, H.D., & Griffith, B.C. (1981b). A cocitation map of authors in judgement and decision research. In B.F. Anderson (Eds.), *Concepts in judgement and decision research: Definitions, sources, interrelationships, comments* (pp. 261–271). New York: Praeger.
- White, H.D., & Griffith, B.C. (1982). Authors and markers of intellectual space: Co-citation in studies of science, technology and society. *Journal of Documentation*, 38, 255–272.