A Citation Analysis of the SIGCSE 2007 Proceedings

Raymond Lister and Ilona Box
Faculty of Information Technology
University of Technology, Sydney
Broadway, NSW 2007, Australia
+61 (2) 9514 1850
raymond@it.uts.edu.au

ABSTRACT
This paper identifies the most commonly cited conferences, journals and books among the 1398 citations made in the 122 publications of the SIGCSE 2007 proceedings. The SIGCSE 2007 authors cited a very large array of conferences, journals and books, but the majority are only cited within a single paper. There are only a very small set of journals and conferences cited frequently. Most books cited are concerned with technical information or are textbooks. Only 2% of books are concerned with computer science education and 23% with education in general. The picture that emerges from this citation analysis is that the SIGCSE community does not have a substantial core set of educational literature. Also, the epistemology of the SIGCSE community is primarily objectivist, with a focus on content, rather than a constructivist, student-centered focus on learning.

Categories and Subject Descriptors
K.3 [**Computers & Education**]: Computer & Information Science Education - *Computer Science Education*.

General Terms
Measurement.

Keywords
Citation Analysis.

1. INTRODUCTION
The sources that computing education authors choose to cite are a reflection of our sub-discipline. While there are databases that index the citations of academic publications, such as the Science Citation Index ® [7], computer science journals and conferences are not comprehensively covered by such databases, and books are not covered at all by such databases. Furthermore, the research interest of the authors of this paper is not in finding how various computer science education conferences, journals and books rate compare to other computer science conferences, journals and books. Instead, our interest is in learning what conferences, journals and books are most popular within the SIGCSE community. Our way of investigating that issue is to analyze the citations in papers from the largest and best known computing education conference, the SIGCSE Technical Symposium on Computer Science Education. Consequently, we chose to conduct a citation analysis of the SIGCSE 2007 proceedings [5], the 38th and (at the time this paper was written) the most recent symposium in the series.

SIGCSE 2007 comprised 108 papers, which was 34% of all submissions received. Each paper received at least five reviews. The symposium also had seven panels (with a 47% acceptance rate) and 16 special sessions (with a 64% acceptance rate). There are 122 publications in the proceedings that contained citations, and we included all of those citations in our analysis. In total, we analyzed 1398 citations, an average of 11.5 per publication.

2. TYPES OF SOURCES CITED
Of the 1398 citations, 323 (23%) were to journal articles and 431 (31%) were to conference papers. Another 317 citations (23%) were to books, either a whole book or a chapter within a book. Table 1 shows a complete breakdown of the different types of sources cited in the proceedings. The source type "curricula" refers most commonly to the ACM Computing Curricula of 2001 [1] but also to a small number of similar documents. Among the "other" sources were unpublished dissertations, personal communications, newspaper articles, grant applications, films, and transcripts. This paper concentrates on the three most common types of sources cited: journals, conferences and books.

Table 1: The types of sources cited in SIGCSE 2007

Type of Source	Number of Citations	Percentage of all Citations
Conference	431	31%
Journal	323	23%
Book	317	23%
Web Page	256	18%
Curricula	38	3%
Technical Report	12	1%
Other	21	1%
Total	1398	100%

3. CONFERENCE PAPERS

The SIGCSE 2007 authors cited papers from 104 different conferences. However, most of those conferences received very few citations. Table 2 shows that 78 of those conferences (75%) received exactly one citation in the entire proceedings, and 92% of conferences received three or less citations.

For many of the less frequently cited conferences, the bulk of the citations occurred in one–to–three SIGCSE 2007 publications. Table 3 shows the number of different papers that cite a particular conference. Of the 104 conferences that were cited, 82 of those conferences (almost 80%) were cited in only one paper, and 94% were cited in three or less papers. As shown in Table 3, six conferences were cited in six or more papers:

- **SIGCSE:** Technical Symposium on Computer Science Education
- **ITiCSE:** Annual Conference on Innovation and Technology in Computer Science Education
- **FIE:** Frontiers in Education
- **CHI:** Int. Conf. on Human Factors in Computing Systems
- **SIGGRAPH:** Int. Conf. on Computer Graphics and Interactive Techniques
- **ACE:** Australasian Conference on Computing Education

The International Computing Education Research Workshop (ICER), which commenced in 2005, was cited in three SIGCSE 2007 publications.

Among the more cited conferences, allowance should be made for the differing sizes of those conferences – a large conference might have more papers cited than a small conference simply because the larger conference has more papers. In the three years of 2003, 2004 and 2005, the SIGCSE conference published 279 papers, ITiCSE published 158 papers, and ACE published 115 papers – an average of 93, 53 and 38 papers per year respectively. Table 4 shows the same data as for Tables 2 and 3, but only for SIGCSE, ACE, ITiCSE, with the data normalized to the average number of papers in each conference in the three years 2003-2005. With the data so normalized, SIGCSE remains the most popular source of citations for SIGCSE 2007 authors. We omitted FIE, CHI and SIGRAPH from this analysis because computing education papers are only a subset of the papers presented at those conferences, and we would be making arbitrary decisions if we attempted to count how many papers in those conferences were relevant to SIGCSE 2007 authors. We also omit ICER from this analysis as that workshop series only commenced in 2005.

3.1 Discussion of Conference Paper Results

It is not surprising that the most popular conference cited by authors of the 38th conference in the SIGCSE series (i.e. SIGCSE 2007) is the SIGCSE conference series. It is a little surprising that ITiCSE is less cited than SIGCSE, even when the citation counts are normalized, given that (1) both conferences are run by the same organization, with the same pool of reviewers, and (2) SIGCSE members automatically receive both the SIGCSE and ITiCSE proceedings even if they do not attend the conferences.

Table 2: The frequency distribution of all conferences cited in the SIGCSE 2007 proceedings.

No. of Times Conference Cited	No. of Conferences	Cumulative Percentage
1	78	75%
2	11	86%
3	7	92%
4	1 – Prog. Viz. Work.	93%
5	1 – ICER	94%
7	1 – ACE	95%
10	2 – CHI, SIGGRAPH	97%
13	1 – FIE	98%
40	1 – ITiCSE	99%
218	1– SIGCSE	100%
Total	**104**	

Table 3: The frequency distribution of papers citing a given conference in the SIGCSE 2007 proceedings.

No. of Papers in which a Conference is Cited	No. of Conferences	Cumulative Percentage
1	82	79%
2	10	88%
3	6 – incl. ICER	94%
6	1 – ACE	95%
7	1 – SIGGRAPH	96%
8	1 – CHI	97%
12	1 – FIE	98%
25	1 – ITiCSE	99%
77	1– SIGCSE	100%
Total	**104**	

Table 4: Citation and paper counts, normalized to the average number of papers in each conference, in 2003-5.

Conference	Average Papers 2003-05	Cites	Papers
SIGCSE	93	2.3	0.8
ITiCSE	53	0.8	0.5
ACE	38	0.2	0.2

4. JOURNAL PAPERS

The SIGCSE 2007 authors cited papers from 135 different journals. However, as with conference papers, most of the journals cited received very few citations. Table 5 shows that 101 of those journals (75%) received exactly one citation in the entire proceedings, and 90% of journals received three or less citations.

Table 5: The frequency distribution of all journals cited in the SIGCSE 2007 proceedings.

Times Journal Cited	No. of Journals	Cum. %age
1	101	75%
2	15	86%
3	5	90%
4	4	93%
5	1 – J. of Ed. Psy	93%
6	3 – IEEE Tran Ed, JVLC, JERIC	96%
12	1 – IEEE Computer	96%
13	1 – Comp. Sci. Education	97%
18	1 – Computing Research News	98%
31	1 – SIGCSE Bulletin	99%
34	1 – J. of Comp. Sci. in Colleges	99%
35	1 – Comm. of the ACM (CACM)	100%
Total	**135**	

Table 6: The frequency distribution of papers citing a given journal in the SIGCSE 2007 proceedings.

No. Papers which cite a journal	No. of Journals	Cum. %age
1	104	77%
2	16	89%
3	5	93%
4	2 – J. of Ed. Psy, IBM Sys. J.	94%
6	2 – JVLC, JERIC	96%
7	1 – IEEE Computer	96%
10	1 – Comp. Research News	97%
12	1 – Comp. Sci. Education	98%
21	2 – CACM, SIGCSE Bulletin	99%
24	1 – J. Comp. Sci. in Colleges	100%
Total	**135**	

Table 7: Normalized citation and paper counts of popular computer science education journals.

Conference	Av. Papers	Cites	Papers
J. of Comp. Sci. in Colleges	272	0.1	0.1
SIGCSE Bulletin (all)	69	0.4	0.3
SIGCSE Bulletin (refereed)	43	0.7	0.5
Computer Sci. Education	15	0.9	0.8

For some of the less frequently cited journals, the bulk of the citations occurred in one–or–two SIGCSE 2007 publications. Table 6 shows the number of different papers that cite a particular journal. Of the 135 journals cited, 104 of those journals (77%) were cited in only one paper, and almost 90% were cited in one–or–two papers.

The three most cited computer science education journals listed in Table 6 are the *Journal of Computer Science in Colleges*, *SIGCSE Bulletin* and *Computer Science Education* (i.e. CACM and Computing Research News are not specifically computer education journals). However, as with the conference citations, an allowance should be made for the differing number of papers published in these three journals. In the three years of 2003, 2004 and 2005, *Computer Science Education* published an average of 15 papers per year. The *SIGCSE Bulletin* appears four times a year, but two of those issues contain the SIGCSE and ITiCSE conference proceedings. We ignored those two conference proceedings issues in this analysis. Determining the average number of papers published in the remaining two yearly issues of the *SIGCSE Bulletin* is not straightforward, as it publishes a mix of refereed papers and small invited columns. Therefore, for *SIGCSE Bulletin,* we have calculated two yearly publishing averages. One of the averages is for all papers published. That yearly average is 69. The other average is just for the refereed papers (including the working group papers). That average is 43.

Despite its name, the *Journal of Computing Sciences in Colleges* contains nothing but the conference proceedings for each of the regional conferences sponsored by the Consortium for Computing Sciences in Colleges [3]. It is therefore not clear whether the analysis of citations to it for the SIGCSE 2007 proceedings should be included in our journal analysis, or should instead be included in our conference analysis. Given that it is explicitly named as a journal, we have elected to include it here in the journal analysis, but the figures provided here can be compared directly to the figures for conferences given earlier in this paper. In the three years of 2003, 2004 and 2005, the *Journal of Computing Sciences in Colleges* published an average of 272 papers per year.

Table 7 shows the same citation and paper counts as in Tables 5 and 6, but only for the three most popular computer science education journals, with the data normalized to the average number of papers in each journal in the three years 2003-2005. With the data in this normalized form, the high profile of *Computer Science Education* is more apparent. (N.B. the two rows for the SIGCSE Bulletin are calculated from the same total citation and paper counts across all SIGCSE Bulletin papers.)

4.1 Discussion of Journal Paper Results

We were initially surprised by the prominence of CACM. On inspection of the actual papers cited, we found that many of the CACM articles cited were shot opinion pieces, often written by prominent members of the computing education, for example: Dijkstra*, On the Cruelty of Really Teaching Computer Science*; Wing, *Computational thinking*; Bruce et al., *Why Math*?; Denning & McGettrick, *Recentering computer science*; Denning, *Great principles of computing*; Camp, *The incredible shrinking pipeline*. It may be that when SIGCSE authors look for a citation that either explains or justifies their philosophical position, they look to these opinion pieces in CACM (and not, as counter example, to a non-computing paper written by a specialist in education).

5. BOOKS

In this citation analysis, we include as "books" both citations to complete books and citations to chapters within edited volumes. The SIGCSE 2007 authors made 317 citations to 268 different books. However, as with conference and journal papers, most of the books cited received very few citations. Table 7 shows that 236 of those books (88%) received exactly one citation in the entire SIGCCSE 2007 proceedings, and 96% of all books journals received two or less citations.

In Table 7, after each book listed, there appears a categorization of each book, within square brackets. For example, *Unlocking the Clubhouse* is categorized as "CSEd", which is short for "computer science education". We categorized books into one of eight categories:

- **Education:** books that discuss teaching and learning issues in a non-disciplinary specific fashion. Table 7 lists four such books.
- **Computing Content:** Table 7 lists three such books. Some of these books might have been cited because they were being used as a textbook, but this was not clear.
- **Text Book:** Books that we recognized as a textbook.
- **Psychology:** usually educational psychology.
- **Gender:** Books not concerned specifically with gender issues in computer science education. Table 7 lists one such book.
- **CSEd:** books specific to education issues within the computing discipline. Table 7 lists two such books.
- **Research Methods:** For example, books on statistics, or qualitative research.
- **Other**

Table 8 summarizes our categorization of all 268 books into one of the eight categories. A majority of books cited (57%) are concerned with computing content or are textbooks. Only a quarter of all books are concerned with education, with only 2% of books concerned specifically with computer science education.

6. AGE OF CITATIONS

Figure 1 shows the number of citations in the SIGCSE 2007 proceedings, by year of publication, drawn from sources published in the previous 10 years. For publications that were more than four years old when SIGCSE 2007 was held, the citation rate drops quickly. Inspection of data from even earlier years shows a steady trickle of citations going back decades.

For each year, the citations shown in Figure 1 are categorized as to whether the citations come from books, journals or conferences. The drop in total citations for publications more than four years old is primarily due to a drop in citations of conference papers. The drop in citations of journal articles is a little more gradual than the drop in conference papers, but the drop in journal article citations is substantial for articles that are seven years or more old. Book citations only decrease slowly, and the inspection of data going back decades shows that the steady trickle of citations are primarily to books. The four oldest books cited were published in 1956 (Bloom *et al*., *Taxonomy of Educational Objectives*), 1955 (a translation of Rousseau's *Émile*, the original French version is about 200 years old), 1938 (Dewey, *Experience*

Table 8: The frequency distribution of all books cited in the SIGCSE 2007 proceedings.

Times Book Cited	No. of Books	Cum. %age
1	236	88%
2	22	96%
3	7 – Beck (1999) *Extreme Programming Explained* [content] Rosser (1990) *Female-friendly science: Applying women's studies methods and theories …* [gender] Bransford, Brown, and Cocking (2000) *How People Learn* [education] Tanenbaum & Woodhull, (1997) *Operating Systems Design and Implementation* [content] Bloom et al. (1956) *Taxonomy of Educational Objectives*. [education] Hutchings & Shulman (1998) *The Course Portfolio: How Faculty Can Examine Their Teaching to …* [education] Cohoon & Aspray (2006) *Women and Information Technology* [CSEd]	99%
4	2 – Gamma, Helm, Johnson & Vlissides (1995) *Design patterns*. [content] Seymour & Hewitt (2000) *Talking About Leaving.* [education]	99%
7	1– Margolis & Fisher (2002) *Unlocking the Clubhouse* [CSEd]	100%
Total 317	**268 Different Books**	

Table 9: The frequency distribution of different types of books cited in the SIGCSE 2007 proceedings.

Type of Book	Percentage of all Books
Computing Content	41%
Education	23%
Text Book	16%
Psychology	6%
Gender	5%
Research Methods	3%
CSEd	2%
Other	4%

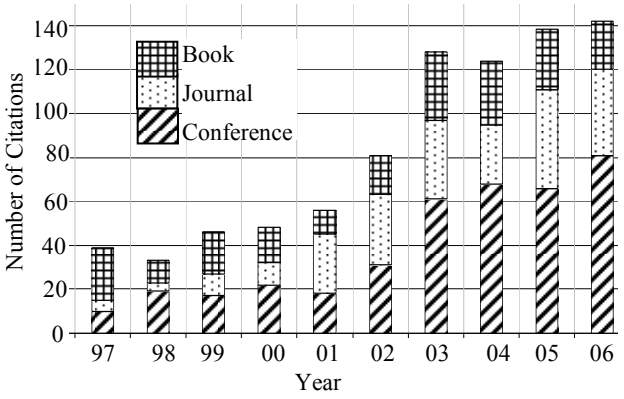

Figure 1: The frequency of citations to books, journals and conferences in the period 1997-2006.

& Education) and 1926 (Kandinsky, *Punkt und Linie zu Fläche*). A total of 91 citations (i.e. 29% of all book citations) are to books published in the 30 intervening year's between Bloom's (1956) book and 1997, the earliest year shown in Figure 1.

7. DISCUSSION AND CONCLUSION

The most striking statistic common to conference, journal and book citations is the variety of sources cited – consequently the majority of sources are only cited in one SIGCSE 2007 paper. What inference, if any, should we draw? In their discussion of Hyland's (1999) study of citation practices, Becher & Trowler [2, [p. 115].] asserted that:

> *... in the less highly structured and strongly specialized disciplines ... there was a general expectation that people will read widely. In history and modern languages, for example, interesting and potentially relevant ideas may be gleaned from a variety of sources.*

In that same discussion of Hyland's work, Becher & Trowler also wrote:

> *In some fast moving areas of biology, it was held to be very difficult to keep up with developments outside one's own specialism, not only because of time pressures but also because of the incomprehensibility of language and logical structure ...* [whereas] *... A characteristic view in slower-moving areas within a discipline, however, held that it was important to keep in touch with developments in other fields ... what you learn about other specialisms helps you to change your attitude towards your own – and you can sometimes pick up useful concepts and techniques which other groups have developed.*

Perhaps then, computer science education (at least how it is practiced by SIGCSE 2007 authors) is a less highly structured, less specialized and slower moving sub-discipline than other aspects of computing. The content we teach in our classrooms changes quickly, is highly structured and highly specialized, but perhaps the pedagogy we bring to our classrooms is less structured and less specialized.

It is apparent in the citations, most obviously in the book citations, that SIGCSE 2007 authors place most emphasis on content literature rather than education literature. It may be that there is within the SIGCSE community an educational epistemology so well agreed upon that there is no need to cite it, but it is more likely that there is not an agreed upon educational epistemology.

Several past phenomenographic studies have studied academics' conception of their teaching [4, 6, 8]. These studies have identified several qualitatively different conceptions of teaching that teachers bring to their classroom. At one extreme, teachers focus on the content of their course, seeing teaching as the act of transmitting knowledge and concepts to the student. At the other extreme, teachers focus upon the student, seeing teaching as the act of helping students to develop or change their own conceptions. The SIGCSE 2007 citations suggest that the educational epistemology of the SIGCSE community is primarily objectivist, with the focus on course content, rather than a constructivist, student-centered focus on learning.

8. REFERENCES

[1] ACM and IEEE (2001) *Computing Curricula 2001* http://www.sigcse.org/cc2001/ [Accessed September 2007]

[2] Becher, T. and Trowler, P. (2001) Academic Tribes and Territories. Open University Press; 2nd edition. ISBN: 033520628X

[3] Consortium for Computing Sciences in Colleges (2007) *Journal of Computing Sciences in Colleges.* http://www.ccsc.org/publications/pubsJournal.htm [Accessed September 2007]

[4] Prosser, M., Trigwell, K., Taylor, P. (1994) A Phenomenographic Study of Academics' Conceptions of Science Learning and Teaching. *Learning and Instruction*, 4, 217–231.

[5] Russell, I., Haller, S. and Dougherty, J. and Rodger, S. (2007) *SIGCSE '07: Proceedings of the 38th SIGCSE technical symposium on Computer science education*. Covington, Kentucky, USA. March 7-10. ACM Press, New York, NY, USA. ISBN 1-59593-361-1.

[6] Samuelowicz, K. and Bain, J. (1992) Conceptions of teaching held by academic teachers. *Higher Education*, 24, 93-111.

[7] Thomson Scientific (2007) The *Science Citation Index* http://scientific.thomson.com/products/sci/ [Accessed September 2007]

[8] Trigwell, K, Prosser, M, Taylor, P (1994): Qualitative differences in approaches to teaching first year university science. *Higher Education* **27**(1):75-84