Educ Inf Technol (2009) 14:105–126
DOI 10.1007/s10639-008-9078-4

Categorising computer science education research

**Mike Joy · Jane Sinclair · Shanghua Sun ·
Jirarat Sitthiworachart · Javier López-González**

Published online: 24 December 2008

Abstract The interdisciplinary nature of Computer Science Education as a field of study is a complicating factor when authors are choosing appropriate journals and conferences for publishing research results. This paper reports a survey of 42 such publications in order to identify the types of paper frequently accepted for publication in each. We review existing taxonomies developed for the general fields of Computer Science and Education, and from this starting point, we develop a novel faceted taxonomy, specifically aimed to help new researchers in the field understand what types of papers are published and where they appear. Our results confirm previous studies indicating the predominance of practice-based, technology-driven reports. We also observe

M. Joy (✉) · J. Sinclair
Department of Computer Science, University of Warwick,
Coventry, CV4 7AL, UK
e-mail: m.s.joy@warwick.ac.uk

J. Sinclair
e-mail: j.e.sinclair@warwick.ac.uk

S. Sun
Division of Community Health Sciences, St. George's,
University of London, Cranmer Terrace, London, SW17 0RE, UK
e-mail: ssun@sgul.ac.uk

J. Sitthiworachart
School of Informatics, Walailak University, 222 Thaiburi,
Tha Sala, Nakhon Si Thammarat 80160, Thailand
e-mail: sjirarat@wu.ac.th

J. López-González
Department of Computer Science and Statistics,
University of Joensuu, PL 111, 80101 Joensuu, Finland
e-mail: jgonza@cs.joensuu.fi

certain differences in educational themes between those found in publications grounded in the Computer Science Education corpus and those emerging from more general Higher Education literature.

Keywords Categorization scheme · Taxonomy · Computer science education research

1 Introduction

A problem often faced by academics working on interdisciplinary topics is that the focus of a potential paper may not fit comfortably with established journals and conferences. This is particularly true where the subject area involves both pure science and social science. The field of Computer Science Education is an interesting case study, since the subject is rapidly changing, in part due to technological advances which are informing both the content of Computer Science courses and the means of course delivery. Many researchers coming into this field may well be used to carrying out Computer Science research and have a good understanding of the research methods used in their own particular field of expertise and a knowledge of relevant journals and conferences. However, the move to researching aspects of Education involves a completely different research landscape.

This paper reports a survey of conferences and journals, in the fields of both Education and Computing, which publish Computer Science Education material. In order to answer the "where to publish" question, we sought to distinguish between those journals and conferences by classifying the types of paper which have recently appeared in each, in order that an author might select a publication appropriate to his or her paper.

We present a novel categorisation scheme for Computer Science Education papers developed specifically for the purpose described above. The taxonomy was used as part of an analysis of all papers published in relevant journals and conference proceedings during a chosen 12 month period. This amounted to a total of 21 journals and 21 conferences, providing a collection of over 3,500 papers in all. Our approach to this study is presented here together with an overview of the results obtained.

The paper is organised as follows. Section 2 considers the use of classification schemes and taxonomies in general. The need to organise and classify material is fundamental to many areas of investigation, and the way in which such classifications are developed and assessed has become a subject of study in its own right. We introduce some of the issues relevant to development of a suitable, straightforward scheme to be applied to Computer Science Education papers. Various classifications schemes, which have already been used in the context of Computer Science and Education research, are considered in Section 3. Section 4 then describes how our taxonomy was developed and presents the final scheme. The next stage of the research was the classification exercise which is reported in Section 5. The findings of our work and issues relating to the categorisation of research are discussed in Section 6.

2 Classification schemes and taxonomies

Classification schemes have long been used to organise large collections of objects, such as books in libraries, and professionals in Information Sciences regard working with classification schemes and taxonomies as a key skill (Morgan and Bawden 2006). Bibliographic schemes, such as Dewey Decimal and the Library of Congress (Rowley and Farrow 2000, pp. 215–229), can be used to categorise accurately the subject matter of a document, but are complex and are normally only used in the context of a library. In a scientific context, the word *taxonomy* is also used (with essentially the same meaning), and the Linnæan Taxonomy of living things is a well-known example (Broughton 2004, p. 13).

The science of classification is of importance in many other fields. For example, the rise of eCommerce has fuelled an interest in taxonomies to support content management, and categorising information stored at the enterprise level (Corcoran 2002).

Gilchrist (2003) identifies four principle factors in the current information-rich environment which contribute to a renewed interest in creating taxonomies.

- Information overload—the large volume and complex structure of data held by organizations preclude the effective use of simple search mechanisms.
- Information literacy—most users do not know how to search for information effectively.
- Organizational terminology—published categorisation schemes currently in use may be inaccurate or inappropriate for a given organization.
- “Destructuring” of organizations—partnerships between organizations (including mergers) may lead to cultural inconsistencies.

In the context of the work of this paper, these are all pertinent points. There exists an increasing body of literature relating to educational topics in Computer Science, and knowing where and how to find relevant documents can be problematic. The publications involved are of a relatively diverse nature, ranging from subject-specific interests to more general education issues. The topic of research may in some cases be interpreted as “education for computer scientists” and in others the focus may be more on educational technology. All of these outlets are currently used by Computer Scientists with an interest in education and are therefore relevant to this study. The diversity of subjects that may be involved includes not just education and the computing disciplines themselves but may also involve other inter-disciplinary connections, for example with psychology, mathematics and social sciences.

This gives great scope for culture clashes and lack of common methods and theories to unite the strands.

A taxonomy aims to divide a population of entities into groups based on similarity in some key characteristics. The groups should partition the population and, preferably, there should be clear differences between distinct groups (Bailey 1994). This may be difficult to achieve in practice, and any nontrivial taxonomy is likely to be subjective to some extent.

Rowley and Farrow (2000, p. 180) identify two principal types of taxonomy: *enumerative* and *faceted* (or *analytico-synthetic*). An enumerative taxonomy is created in a manner which may appear unsystematic and ad-hoc. In practice, however, the categories are often derived by the use of *literary warrant*—"words and phrases drawn from the literature of the field should determine the formulation of descriptors" (NISO 1994). The Dewey Decimal and Library of Congress bibliographic classification schemes and the Linnæan Taxonomy are all essentially enumerative. Within an enumerative taxonomy simple hierarchical classifications are made, such as:

University → Warwick → Department → Computer Science.

However, in a faceted taxonomy, attributes and attribute values allow for much greater detail, for example:

Computer Science → Academic Staff 24
→ Undergraduates 600
→ Postgraduates 100

To create a faceted taxonomy, a *facet analysis* of the target objects is carried out in order to identify mutually exclusive and conceptually distinct categories. An early (though little-used today) example is Ranganathan's Colon Classification Scheme, first published in 1933, which uses a five-faceted approach for bibliographic classification (Broughton 2004, pp. 258–259). The creation of headings in a thesaurus, such as the Thesaurus of Psychological Index Terms (APA 2007), is another example of facet analysis.

A taxonomy—whether enumerative or faceted—may also be described as belonging to one of the following three distinct types (Côté 2005).

- Flat—there is a controlled set of possible categories, and there may not be meaningful relationships between the categorical values. A list of football players organized by team would be a simple example.
- Hierarchical—this is a traditional tree-structured categorisation, as in the enumerative example on the previous page. Examples of its use include the Linnæan Taxonomy and MeSH (the National Library of Medicine's controlled vocabulary thesaurus) (NLM 2007).
- Network—complex relationships exist between the categories, such as might exist in the Semantic Web.

This distinction is useful in helping to further refine and clarify the description of a taxonomy, and is one we will refer to again below.

3 Taxonomies in computing and education

Within the Computing disciplines, the ACM (Association for Computing Machinery) has a system (ACM 1998) for classifying computing papers which has been in use since 1964 and is regularly updated. It is an hierarchical enumerative system, which is designed to be regularly updated in order to keep pace with the evolving technologies. However, the ACM scheme caters only for scientific papers. It does have a category (K3) which addresses Computers and Education and this is divided into four sub-categories (ACM 1998):

- K.3.0 General
- K.3.1 Computer Uses in Education
  - Collaborative learning
  - Computer-assisted instruction
  - Computer-managed instruction
  - Distance learning
- K.3.2 Computer and Information Science Education
  - Accreditation
  - Computer science education
  - Curriculum
  - Information systems education
  - Literacy
  - Self-assessment
- K.3.m Miscellaneous

This approach brings out the distinction between computing education (K.3.2) and computing in education (K.3.1). Papers categorised using the scheme are likely to identify multiple relevant categories, possibly from diverse branches of the hierarchy. While providing very comprehensive coverage for scientific papers, this scheme is not concerned with the educational content of the article and has one broad category (Computer Science Education) which would cover a large range of possibilities.

Vessey et al. (2005) describe a classification system intended to provide a unified perspective for research in all computing disciplines (encompassing Computer Science, Software Engineering and Information Systems). This is mainly focused on identifying a comprehensive set of topics relevant to the three areas and contains just one catch-all education category.

The Computing Ontology Project (Cassel 2007) is a work in progress which includes amongst its aims: “to describe fully the various topics and subtopics of interest to educators and researchers in any of the disciplines concerned with computing and the management and processing of information”. This is an interesting area of development, but at the time of writing contained no coverage of “Computing Education”.

An approach specific to Computer Science education research has been proposed by Fincher and Petre (2003). The aim of their categorisation is to map out the topics of interest as currently found in this area of research. Although there is no explanation of how these categories were arrived at, the authors have identified the topics which they believe motivate researchers in this area. The ten suggested categories are:

1. student understanding;
2. animation, visualisation and simulation;
3. teaching methods;
4. assessment;
5. educational technology;
6. transferring professional practice into the classroom;
7. incorporating new developments and new technologies;
8. transferring from campus-based teaching to distance education;
9. recruitment and retention; and
10. construction of the discipline.

Fincher and Petre (2003, p. 2) make the point that "Computer Science education research" should indicate work which demonstrates high theoretical content and is also strongly grounded in empirical evidence. However, in practice much of what is currently published is lacking in theoretical perspective. This is discussed further in Section 6. Here we note that we wish to include in our survey the broad range of publications to reflect current practice, and we would like our categorisation to include the means to distinguish articles along the lines of theoretical content. Also, both technology and the educational agenda change rapidly, and although precisely focused, any content-based categorisation may require regular revision.

Some studies have devised a classification system and have also used that classification to analyse particular samples of the literature. The proceedings of the *SIGCSE Technical Symposium* contain Computer Science Education papers, and Valentine (2004) analysed the proceedings over a 20 year period. His approach was pragmatic—he selected those published papers whose content specifically related to first year undergraduate Computer Science students. Six categories are identified (the percentages in this list are the proportions of papers which fall into each category during the 10-year period 1994–2003):

- "Tools" (24.6%): descriptions of new hardware or software;
- "Marco Polo" (23.9%): "Been there, done that" reporting classroom experiences;
- "Experimental" (22.1%): scientific evaluations of techniques or technologies applied in the classroom;
- "Nifty" (17.8%): single ideas used in a specific context;
- "Philosophy" (9.1%): general (fundamental) issues to do with Computer Science Education; and
- "John Henry" (2.5%): difficult and unusual ideas.

This study considers a single publication over a 20-year time period. Our survey is limited to a single year, but views the position over all relevant publications and without restriction to a particular topic. The names used for the categories capture the spirit of the articles: "Marco Polo" papers relate to Fincher and Petre's (2003) "practice paper". However, because of the specific nature of this exercise, broader categories such as curriculum are absent.

Rossano et al. (2005) scrutinised 54 papers relating to Learning Objects in a single conference (ICALT 2004), and identified three (overlapping) types of paper classified according to the apparent skills of the authors. The categories used for this purpose were: *technological*—computer scientists or IT developers; *pedagogic*—writing from a theoretical educational perspective; and *pragmatic*—teachers or instructional designers whose principal interest is authorship of effective Learning Objects. These were then subdivided into further categories relating explicitly to each paper's focus on the topic of Learning Objects. Again, this work looks at a single publication only, and more from the perspective of the author than the nature of the research.

In the broader domain of general education research, further taxonomies and categorisation exercises may be found. Of particular relevance is work carried out by Tight (2003, 2004) in which a survey was conducted of academic journals publishing research on higher education. Tight identifies 17 relevant publications and analyses the 406 articles appearing in these journals in the year 2000. A number of different categories are used to analyse the papers. The one most relevant to our work is the "Themes and Issues" categorisation which identifies eight specific themes: teaching/learning, course design, student experience, quality, system policy, institutional management, academic work and knowledge. The themes were developed in a pragmatic way by choosing appropriate keywords to associate with each article. This resulted in more than 100 keywords which were than grouped by inspection.

> Similar and related keywords were then successively grouped together to arrive at a limited number of key themes or issues. Of course, the way in which the keywords were grouped together was particular to me, and others would doubtless do it at least somewhat differently, but I do feel that it has some logic and is functional. (Tight 2003, pp. 6–7)

The categories produced provide a useful basis for the analysis that follows. This is an example of category identification through empirical analysis of the sample population, resulting in a taxonomy which, while necessarily subjective, is demonstrably fit for purpose.

4 A new approach

The existing classifications are interesting in their own right, and all have features which are useful to note in relation to the current exercise. However, each is based on a slightly (or, in some cases, greatly) different subject area and purpose. Our aim was to look at all publications in which Computer Science education research (interpreted in a broad and inclusive way) is the main focus. We did not include the many publications in which one or two relevant articles may sometimes be found. We wished to provide researchers with a

Table 1 Numbers of documents used in this survey

	Number of publications	Number of papers
Journals	21	517
Conferences	21	3,081
Total	42	3,598

general picture of what is published where: information which may be of use in deciding where to submit a given paper for publication.

4.1 Selecting publications

Our intention was to include all relevant academic journals and conference proceedings viewed over a 1 year period. This includes both publications which focus on Education *in* Computer Science (for example, the journal "Computer Science Education"(CSE)) and also those whose main area is Computers *for* Education (such as the journal "Computers and Education"). A list of titles was obtained through literature review, web search and consultations with colleagues in the field. The decision to include a publication was generally clear cut. However, an issue of availability did emerge. Several journals which appeared relevant were not obtainable by any of the authors, either on loan (from national lending libraries) or from the publishers. These publications were excluded from the survey since they were considered to be ineffective vehicles for dissemination. For conferences, full papers only were included. A list of the publications surveyed is given in Appendix A. The numbers of publications and papers included are given in Table 1.

Since there is a very large number of journals and conferences relating to comput*ing* and education, we have excluded publications in which the majority of articles appear to be of little or no interest to an educator in Computer Science. For example, journals which focus on the deployment of ICT technologies in high schools were deemed to be outside the scope of this exercise.

It is also worth noting that several publications have been included which do not yet have a substantive international standing, yet are regarded by practitioners as meeting the high standards for peer review expected of an international journal or conference.

For the purposes of this categorization exercise, we would not expect these decisions to affect our results.

4.2 Starting point for a taxonomy

In developing the taxonomy, it seemed to us that a very pragmatic approach was appropriate, based on our own and our colleagues' experiences of scrutinising calls for papers and the mission statements of academic journals.

As a starting point, we noted that most of the Computer Science Education papers that we have seen can be viewed as having both technical and educational content, but with one being the principal focus of attention. For example, a mainly technical paper might report the design of a novel software tool incorporating a new technological innovation, whereas a principally education-focused paper might discuss a pedagogic analysis of a software tool applied in a classroom context. This suggested to us that classification based on a linear scheme might be appropriate:

$$\text{Technical perspective} \Longleftrightarrow \text{Education perspective}$$

4.3 Methodology

We chose to perform a simplified facet analysis (Broughton 2004, pp. 257–283) in order to arrive at a workable categorisation scheme. Our approach was iterative.

First of all, since most papers are short (conference papers are seldom allowed to be longer than six pages), we worked on the premise that each paper has a principal identified result which is being reported. The categorisation would then focus on that result (and the evidence supporting it).

We then took four descriptions of possible paper types, as follows. These represent four stages along our continuum.

1. *System*—the main focus of the paper is a software (or hardware) computer system, with educational functionality, and the focus is technical.
2. *Technology*—the paper discusses a learning technology in the absence of any implementation.
3. *Practical pedagogy*—the paper is educational in interest, but reports on, for example, an experiment or case study without substantial educational theory.
4. *Theoretical pedagogy*—the paper contains substantial material grounded in education theory.

We selected one set of proceedings for each conference and one volume of each journal during the years 2004 and 2005 (2 years were used since journal volumes do not always coincide with calendar years, and some conferences are biennial). We then commenced reading the papers in the journals and conferences we had identified, and attempted to map each paper to one of our categories. From time to time, we would encounter a paper which could not map to any category, in which case we would create a new category. Furthermore, the description of each category might be modified in order to increase the precision and ensure that the categories remained mutually exclusive. Whenever a new category was created, or a category modified, we would iterate over our previously categorised papers to ensure consistency.

In the early stages of the exercise, each paper was classified independently by at least two of the authors. After approximately 10% of the exercise had been completed, papers were normally read by one author only, but consistency was ensured by randomly double-assigning papers.

Any such categorisation is, by its nature, approximate, and there were papers for which the choice of category was ambiguous. However, the proportion of these was only about 5%. For the purposes of this exercise, the classification is sufficiently accurate to indicate the profile of each of the journals and conferences which we have surveyed.

4.4 The taxonomy

At the end of the exercise, we had identified eleven categories, as described in Table 2 below. These are obviously subjective groupings with an element of overlap and ambiguity. However, the researchers found them to be a reasonable reflection of the interests reported in the literature and a useful, yet simple, framework within which to work.

It was perhaps surprising that the "Other" category should be required. We did not initially expect to find papers which seemed to have no discernable educational content and no computing element either! In practice, a small number of such articles was encountered leading to the introduction of this category.

Table 2 Final categorisation scheme

System	Any Computer Aided Learning system, where there is a detailed and concrete system (or tool, or architecture, or algorithm) reported and/ or evaluated.
Technology	Any learning technology (such as artificial intelligence, multimedia, virtual reality, learning objects, virtual learning environments VLEs) where the technology (or framework, or model, or environment) is reported in the absence of a concrete implementation or proposal.
Resources	Learning and teaching resources are reported (for example, a repository of learning objects, a web site, or a portal).
Other technical	The paper has no clear educational interest; examples might include—a mixture of topics/foci with no clear theme—an evaluation of an IT system applied to a domain which is not educational.
Theoretical pedagogy	The focus of the paper is principally educational, and reports results grounded in education theory (i.e. explicitly references, discusses and applies pure education theory). This category is used for "Learning Psychology" where the "learning" is the predominant focus of theory, such as Bruner and Vygotsky.
Practical pedagogy	The focus of the paper is principally educational, and reports results supported by education theory (i.e. the theory may be implicit and may not be supported by pure educational references).
Curriculum	The focus of the paper is the curriculum (or syllabus) for a course. (or module)
Social factors	The paper reports educational-related results where the principal interest is the social context (such as age, gender, culture).
Psychology factors	The paper reports results which are grounded in psychological theory.
Other educational	The paper is educational but with no technical interest; this is a paper which is probably not of interest to a Computer Scientist.
Other	Anything with neither Computing nor Education content.

5 Results

The data were collected and presented via a web resource (Joy et al. 2007) which contains information about each publication, including basic factual data such as its stated mission and the URL resource where further information useful for authors may be found. In addition, the proportions of papers in each category in our taxonomy are presented graphically.

5.1 Presenting the information

To illustrate the output of our exercise, the following data were identified for the journal *Computer Science Education* and for the *ICALT* conference, as illustrated in Figs. 1 and 2 below.

These charts confirm the technical focus of the ICALT conference and the pedagogic flavour of the Computer Science Education journal. Similar information for each of the publications surveyed is available from the web repository (Joy et al. 2007).

5.2 Overall results

The following histogram (Fig. 3) displays the publications in order of technical content (which we take to be the proportion of papers in each publication which we have categorised as System, Technology, Resources or Other Technical). The labels refer to the journal/conference acronyms (a full list is given in the Appendix A).

It would be unwise to make generalisations about all conferences and journals, since the remits of each are different. However, it is clear from Fig. 3 that most of the conferences (such as ICALT and ITiCSE) are strongly

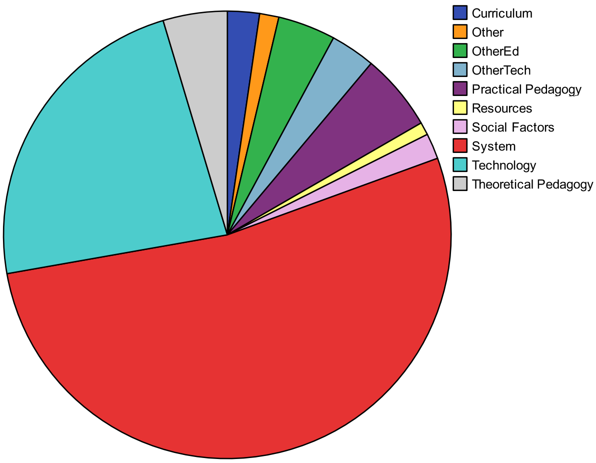
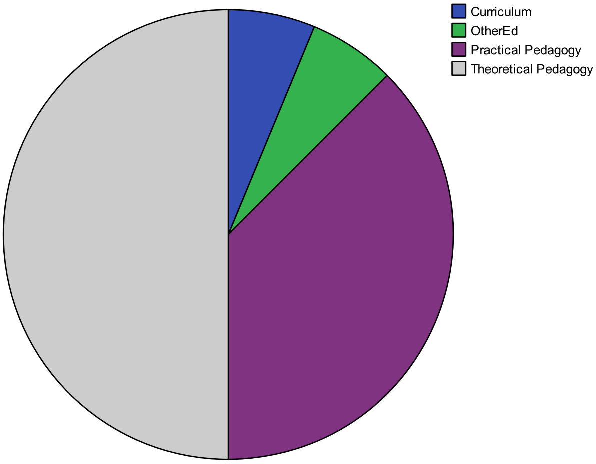


Fig. 1 Proportions of classes for the ICALT conference

Fig. 2 Proportions of classes for the CSE journal



technically focused, and that many journals (such as CSE or JCAL) take a more educational viewpoint.

6 Discussion

6.1 Journals and conferences

The results of our survey provide information about each individual publication, and about the body of literature as a whole. We observe that, in general, conferences tend to focus on reports of technical activities, whereas many journal articles offer a deeper pedagogical perspective. This is perhaps to be expected, since many conferences place severe limitations on the maximum length of papers. It is very difficult, for example, to provide an account of the supporting theory and to report experimental results in just four or five pages.

There appears to be an interesting difference between disciplines in the way that conferences are viewed and in the perceived value of conference papers as opposed to journal papers. In the field of Education, conference papers seem to be viewed as the lesser beast. Tight (2003, p. 16) dismisses them in a fairly cursory way from his sample of publications as "'unpublished' journal articles".

> My main reason for excluding them . . . is that many 'unpublished' articles, where they are of sufficient quality, later become published journal articles after some revision.

This contrasts with many Computer Science conferences which are regarded as prestigious in their own right. More generous length restrictions allow much greater expression, and publication may not necessarily be seen as a springboard to journal publication. For example, the Computing Research

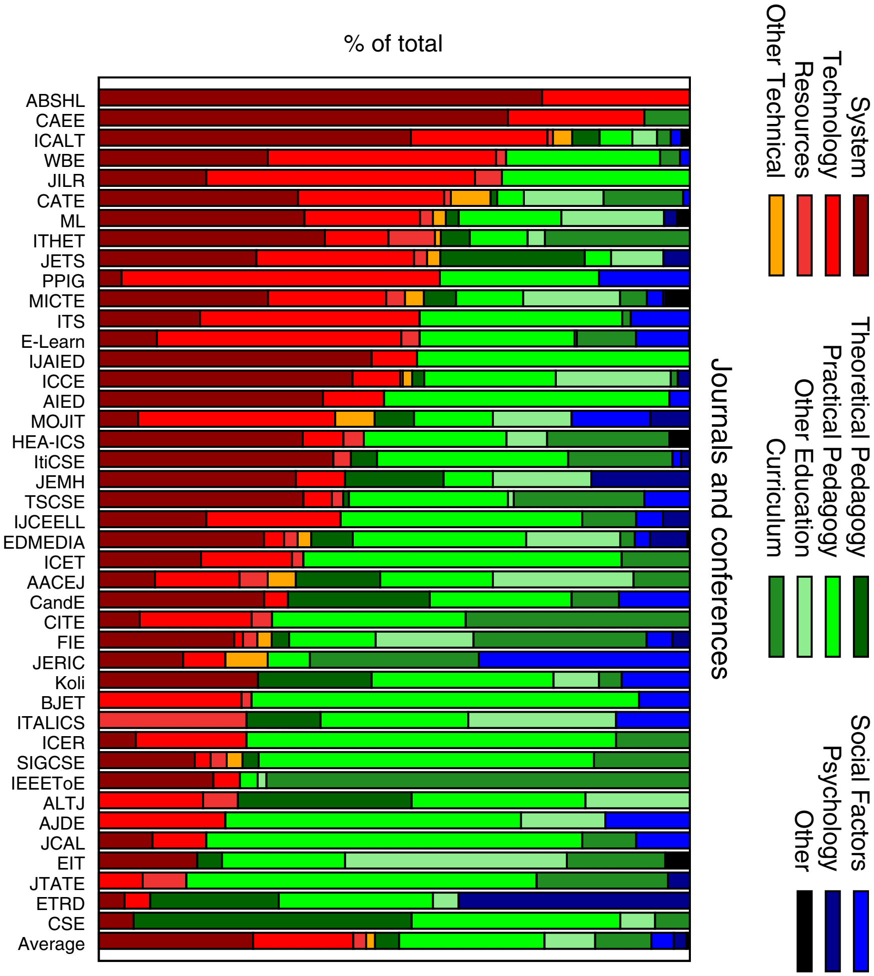


Fig. 3 Publications surveyed and proportions of classes

Association issued advice on assessing academic output which stressed the importance of conference publication (CRA 1999).

> For experimentalists conference publication is preferred to journal publication, and the premier conferences are generally more selective than the premier journals. …Conference publication is both rigorous and prestigious.

The Computer Science Education research conferences seem to follow the Education model in some respects, with their very limited page allowances and an emphasis on providing a fast route to presenting and discussing research often at an early stage. However, the topics of discussion appear mostly to be of a technical nature, veering more towards reporting areas such as educational tool development and examples of applying such technology. This is perhaps of interest in the context of a developing "community of practice" as discussed below.

6.2 Taxonomy limitations

The purpose of our classification is to help new researchers in the field understand what types of papers are published and where they appear. To that end, we have developed a simple, pragmatic taxonomy which covers the spectrum of current relevant publications. A basic linear scale between "technical" and "educational" perspectives was subdivided into categories suggested by empirical investigation of the papers surveyed. Providing such a classification helps to describe the area and to present an easily understandable, if simplified, picture of the current situation. Any taxonomy is necessarily subjective, and the test of a taxonomy has to be whether it is fit for purpose. In this case, a highly detailed and complex classification system would have been counterproductive, making it more difficult to provide a broad perspective. Having used the process of iteration described in Section 4.3 to reach a consensus on the eleven categories, the research team generally felt comfortable with the scheme and a high level of agreement was observed between researchers in the classification choices made. Although ours is by no means the only possible approach, we believe it has been useful in the context.

Any survey of this nature must make a number of assumptions and simplifications. One important factor here is that we chose to identify each paper with a single, main category. As noted in Section 4.3 this caused difficulty in only about 5% of cases. In most instances, there was no real challenge for the "main" classification: it was not the case, for example, that papers describing systems also had a *slightly* lower focus on educational theory—the difference in most cases was pronounced. Hence, we believe that the picture presented is a reasonably fair one. We note that Rossano et al. (2005) analysed a subset of the papers in ICALT 2004. The proportion of those 54 papers which were wholly or in part "technological" was 80%—this is very close to our analysis of the whole conference which identified that 79% of the papers had content which was principally technical, and supports the accuracy of our classification.

It is also important to note that our coverage is intended to survey the range and nature of publications in which Computer Science educators are currently likely to publish. This means that our use of the term "Computer Science Education research" is inclusive of the many practice-based papers that some might feel should not be considered under this label. Thus we have included a broad range of publications which accept mainly report-style papers of either educational technologies or classroom activities. Even amongst such publications, a few "theoretical" categorisations are observed to have been included. The inclusiveness of our work is therefore useful in attempting to identify where theory is addressed.

In general, the results of our survey appear in line with others such as Rossano et al. (2005). However, there are some contrasts with other work.

Tight (2003) found that the two most popular publication themes of Higher Education research were Course Design and System Policy. It is difficult to make direct comparison since the categories are defined in a different way. Tight's "Course Design" category includes (amongst other strands) both curriculum design and learning technologies. Given the breadth of this category it is not surprising that it scored so highly, and these issues were also common topics in our survey. However, the policy theme is not one which we encountered widely. In fact, so few such papers were observed that they were incorporated as part of our "Other Educational" category. It is perhaps to be expected that "general" publications will reflect greater interest in "general" issues while subject-focused ones will concentrate on issues specific to the subject. However, it is likely that many Computer Science educators believe issues such as national policy and funding to be highly relevant to Computer Science education. A number of the publications we surveyed explicitly encourage submissions on policy in their mission statements; many others have general statements which could include them. Yet there seems to be relatively little interest in this area directly related to Computer Science education at present.

The previous point also illustrates an obvious need for caution in how the results of our survey might be used. The fact that a particular journal has published a high proportion of papers on a particular topic suggests that it may well be a good place to send a new paper on the topic. However, an author should not exclude a journal or conference *only* on the basis of the past publication records, the editorial policy or call for papers should also be used for guidance.

The motivation behind our survey and the introductory material we collected (Joy et al. 2007) was to provide a resource for Computer Science educators who are interested in becoming active in education research. Such research, its dissemination and on-going critical appraisal are vital to provide a better understanding of the issues involved in this area and to ensure that new ideas are shared and analysed. Computer Science Education in practice gives rise to many new approaches and to the evolution of new technologies and educational systems. These need to be properly assessed using rigorous methodologies to determine what really provides educational benefits and what does not. As noted by Guzdial (cited in Almstrum et al. 2005):

> Too much of the research in computing education ignores the hundreds of years of education, cognitive science, and learning sciences research that have gone before us.

At the most basic level, an awareness of where to look in order to find relevant literature in the diversity of publications, and knowledge of where contributions might be placed, can help overcome the initial barrier to becoming grounded and active in this area of research. The classification system proposed here provides information that allows researchers to identify the best starting points both for finding out about particular themes they wish to research and for deciding on an appropriate home for papers they produce.

6.3 Theoretical aspects of computer science education

A further purpose of examining published work which we have not included in the present study is to classify (and assess the quality of) the methodology used. For example, in addition to their taxonomy of topics, Vessey et al. (2005) include a "Research Methods" classification (not education-focused). This encompasses both the methods commonly used in the disciplines considered (for example, case studies and literature reviews) together with less frequently encountered social science methods (such as action research and ethnography). Inclusion of a particular method appears to be based on its observation or likely use in practice. In the context of Computer Science education, Randolph et al. (2005) consider research papers appearing in the proceedings of ICALT 2004. The eight-category system they use is arrived at by focusing on the research carried out with human participants. All other activity (such as literature reviews and technical descriptions) are classed together as "other". With this approach, 61% of the papers surveyed fell into the "other" category. To gain an understanding of the methodologies used within this broad grouping, the classification scheme would need to be extended. The findings of this study indicate that there was much scope for improvement of methodology, both in terms of method chosen and in application of the chosen method.

Our study shows that the proportion of papers with a largely theoretical education focus is currently small overall (8.9% for journal papers and 3.3% for conference papers). We found a large proportion of the papers surveyed to be in the categories "System", "Technology" and "Practical pedagogy" (69.1% for conferences and 59.2% for journals). Our findings agree with those of Randolph et al. (2005) in that the papers in these categories do an excellent job of describing the system or classroom innovation undertaken, but generally devote less space to analysis and evaluation or to setting out an explicit evaluation or theoretical perspective. These observations are also in line with those of Tight's (2003, 2004) survey of higher education journals. In this study it was found that half the articles did not specifically identify a methodology, and roughly three quarters did not demonstrate explicit engagement with theory. Tight (2004, p. 409) identifies a possible danger in this:

> ...there is a need for more theoretical engagement so that the field (or community of practice) can develop further, and gain more credibility and respect.

In the specific area of Computer Science education, Fincher and Petre (2003) consider current research output on two dimensions: theory and empirical evidence. Our findings support their claim that most papers are low on the former scale and high on the latter. Fincher and Petre view this as a consequence of the comparative youth of the subject area which will in time mature, through cycles of inductive and deductive reasoning, to produce general theoretical frameworks. It would be interesting to see whether there is any evidence to show any spontaneous move in this direction, although our survey does not seem to indicate any recognisable shift of focus so far. Also, this view does not account for why there is little explicit engagement with the "theory-rich" area of education to which there is a strong link. Efforts have been made (Fincher and Tenenberg 2006) to jump-start the formation of a community of practice in order to establish a common language, working towards "normalisation of Computer Science Education as a distinct and rigorous research paradigm". Such initiatives are very encouraging, and it will be interesting to see if a distinct, commonly shared, body of theory does begin to emerge. The current profile of publications raises some interesting issues in relation to this. The popularity of, for example, some of the large practice-oriented conferences suggests that a community (or communities) of practice is already flourishing. The development of communities of practice is influenced both by the shared interests and aims of their members and by external issues, such as the need to find effective ways to teach Computer Science classes and pressure for academics to concentrate research efforts on the most highly-valued topics. For example, in British universities staff may be strongly encouraged to concentrate on research areas which are thought to be given greater recognition in the UK's Research Assessment Exercise. At the moment, a high preponderance of published papers provide a descriptive account of a single, isolated and unevaluated effort. Our work simply maps the current situation, reflecting the current position in a suitable taxonomy.

A further point in relation to this is the changing nature of Computer Science itself. The continuing development of technology and techniques means that it is a dynamic subject. Educators are constantly developing new material and working with new media and resources. On the one hand, it could be said that it is particularly important to evaluate properly the effect of such innovation to avoid changes which have no educational value. However, the time and effort required to do this (and to develop the enabling theoretical perspectives) may be daunting in a rapidly changing environment in which new technology is often regarded as progress regardless of the actual effect.

The developing nature of Computer Science education and the current move to understand and define Computer Science education research require sorting out the core concepts from the ephemeral trends and identifying the paradigm and theory particular to the subject. Any classification system appropriate now is likely to become less so as understanding progresses. If the prediction of theoretical development is correct, then a taxonomy that aimed to be of use to those entering this area of research would need a finer classification of theoretical aspects. However, indications currently are that the broad technical/educational spectrum we have worked with will remain relevant for reflecting what happens in practice for some time yet.

7 Conclusion

We have presented a new taxonomy which can be applied to articles in the domain of Computer Science Education—both conferences and journals—in order to assist an author in choosing a potential target publication for a paper. For this purpose, we have developed and used a simplified faceted classification scheme informed by a survey of 42 journals and conferences. Our results indicate that a large proportion of papers do not address issues of educational theory and that most articles are still based on reports of tool development or of use of technology in the classroom. Very often, there is little evaluation of impact or of educational benefit. This is much more pronounced in conference publications, which in this discipline tend to be very short. Our results are in broad agreement with previous single publication studies in Computer Science Education, and also with Tight's more general survey (Tight 2003) of Higher Education Research journals. However, it seems that researchers in Computer Science Education are less likely to write about themes of national and international policy and funding, or if they do, they do not publish these papers in the Computer Science Education publications.

Acknowledgements The work reported here was sponsored by the UK's Higher Education Academy Subject Network for Information and Computer Sciences.

Appendix A. Publications included in the survey

The following publications were reviewed in this research. All web references were accessed in October 2008.

Journals

AACEJ	Publication of the Association for the Advancement of Computing in Education www.aace.org
AJDE	American Journal of Distance Education www.ajde.com
ALTJ	ALT Journal www.alt.ac.uk/alt_j.html
BJET	British Journal of Educational Technology www.becta.org.uk
CAEE	Computer Applications in Engineering Education www.interscience.wiley.com/jpages/1061-3773
CandE	Computers and Education www.elsevier.com/wps/find/journaldescription.cws_home/347/description
CSE	Computer Science Education www.tandf.co.uk/journals/titles/08993408.asp
EIT	Education and Information Technologies www.springer.com/journal/10639

ETRD	Educational Technology Research and Development www.aect.org/Intranet/Publications
IEEETоE	IEEE Transactions on Education www.ewh.ieee.org/soc/es/esinfo.html
IJAIED	International Journal of Artificial Intelligence in Education aied.inf.ed.ac.uk
IJCEELL	International Journal of Continuing Engineering Education and Lifelong Learning www.inderscience.com/ijceell
ITALICS	Innovations in Teaching And Learning in ICS www.ics.heacademy.ac.uk/italics
JCAL	Journal of Computer Assisted Learning www.blackwellpublishing.com/journal.asp?ref=0266-4909
JEMH	Journal of Educational Multimedia and Hypermedia www.aace.org/pubs/jemh
JERIC	ACM Journal on Educational Resources in Computing www.acm.org/pubs/jeric
JETS	Journal of Educational Technology and Society www.ifets.info
JILR	Journal of Interactive Learning Research www.aace.org/pubs/jilr
JTATE	Journal of Technology and Teacher Education www.aace.org/pubs/jtate
MOJIT	Malaysian Online Journal of Instructional Technology pppjj.usm.my/mojit
SIGCSE	Bulletin of the ACM Special Interest Group on Computer Science Education www.acm.org/sigcse

Total number of journals: 21
Total number of journal papers: 517

Conferences

ABSHL	Agent-Based Systems for Human Learning Workshop agents.sci.brooklyn.cuny.edu/abshl05
AIED	International Conference on Artificial Intelligence in Education hcs.science.uva.nl/AIED2005
CATE	IASTED International Conference on Computers and Advanced Technology in Education www.iasted.org/conferences
CITE	ACM Conference on Information Technology Education portal.acm.org

E-Learn	World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education www.aace.org/conf/elearn
EDMEDIA	Educational Multimedia, Hypermedia and Telecommunications www.aace.org/conf/edmedia
FIE	IEEE Frontiers in Education Conference www.fie-conference.org
HEA-ICS	Higher Education Academy Subject Network for ICS Conference www.ics.heacademy.ac.uk/Events
ICALT	IEEE International Conference on Advanced Learning Technologies lttf.ieee.org
ICCE	International Conference on Computers in Education www.apsce.net
ICER	International Computing Education Research Workshop icer2005.cs.washington.edu
ICET	IASTED International Conference on Education and Technology www.iasted.org/conferences
ITHET	International Conference on IT Based HE and Training fie.engrng.pitt.edu
ITS	Intelligent Tutoring Systems www.itsconference.org
ITiCSE	Integrating Technology into Computer Science Education portal.acm.org
Koli	Baltic Sea Conference on Computing Education Research cs.joensuu.fi/kolistelut
MICTE	International Conference on Multimedia and ICTs in Education www.formatex.org/micte2005
ML	IADIS International Conference on Mobile Learning www.iadis.org
PPIG	Psychology of Programming Interest Group Annual Workshop www.ppig.org/workshops
TSCSE	ACM Technical Symposium on Computer Science Education portal.acm.org
WBE	IASTED International Conference on Web-based Education www.iasted.org/conferences

Total number of conferences: 21
Total number of conference papers: 3081

Publications too recent for inclusion

As this is a rapidly changing area of research, developments and additions are inevitable. Several more recent journals are not included as they did not exist at the time of our survey. These include:

IJET	International Journal of Emerging Technologies in Learning www.i-jet.org
IJMLO	International Journal of Mobile Learning and Organisation www.inderscience.com/ijmlo

References

ACM (1998). *Computing classification system.* [online] www.acm.org/class. Accessed 30 October 2008.

Almstrum, V., Guzdial, M., Hazzan, O., & Petre, M. (2005). Challenges to computer science education research. In W. Dann, T. Naps, P. Tymann & D. Baldwin (Eds.), *SIGCSE—Bulletin of the ACM special interest group on computer science education* (pp. 191–192). St. Louis, Missouri, February.

APA (2007). *Thesaurus of psychological index terms*. [online] www.apa.org/databases/thesaurus.html. Accessed 30 October 2008.

Bailey, K. D. (1994). *Typologies and taxonomies: An introduction to classification techniques.* Thousand Oaks, CA: Sage.

Broughton, V. (2004). *Essential classification*. London: Facet.

Cassel, L. (2007). *The computing ontology project*. [online] what.csc.villanova.edu/twiki/bin/view/Main/OntologyProject. Accessed 30 October 2008.

Corcoran, M. (2002). Taxonomies: hope or hype. *Online, 26*(5), 76–78.

Côté, J. A. (2005). Knowledge taxonomies: What's the role for information professionals? *Information Outlook*, *9*(6), 45–46, 48, 51–52.

CRA (1999). *Computing research association best practices memo evaluating computer scientists and engineers for promotion and tenure.* [online] www.cra.org/reports/tenure_review.html. Accessed 30 October 2008.

Fincher, S., & Petre, M. (Eds.) (2003). *Computer science education research.* London: Routledge Falmer.

Fincher, S., & Tenenberg, J. (2006). Using theory to inform capacity-building: bootstrapping communities of practice in computer science education research. *Journal of Engineering Education*, *95*(4), 265–278.

Gilchrist, A. (2003). Thesauri, taxonomies and ontologies–An etymological note. *Journal of Documentation*, *59*(1), 7–18.

Joy, M., Sun, S., Sitthiworachart, J., Sinclair, J., & López, J. (2007). *Getting started in computer science education research*. [online] www.ics.heacademy.ac.uk/resources/pedagogical/cs_research. Accessed 30 October 2008.

Morgan, J., & Bawden, D. (2006). Teaching knowledge organisation: educator, employer and professional association perspectives. *Journal of Information Science*, *32*(2), 108–15.

NISO (1994). *ANSI/NISO Z39.19-1993 guidelines for the construction, format and management of monolingual thesauri*. Bethesda, MD: NISO.

NLM (2007). *Medical subject headings*. [online] www.nlm.nih.gov/mesh/. Accessed 30 October 2008.

Randolph, J., Bednarik, R., Silander, P., Gonzalez, J., Myller, N., & Sutinen, E. (2005). A critical analysis of the research methodologies reported in the full papers of the proceedings of

ICALT 2004. In *Proceedings of the 5th IEEE conference on advanced learning technologies* (pp. 10–14). Taiwan, July.

Rossano, V., Joy, M., Roselli, T., & Sutinen, E. (2005). A taxonomy for definitions and applications of LOs: A metaanalysis of ICALT papers. *Educational Technology and Society*, *8*(4), 148–160.

Rowley, J., &Farrow, J. (2000). *Organizing knowledge: An introduction to managing access to information*, 3rd Edn. Aldershot, Hants: Gower.

Tight, M. (2003). *Researching higher education.* Milton Keynes, UK: Open University Press.

Tight, M. (2004). Research into higher education: An a-theoretical community of practice? *Higher Education Research and Development*, *23*(4), 395–411.

Valentine, D. W. (2004). CS educational research: A meta-analysis of SIGCSE technical symposium proceedings. In *Proceedings of the 35th SIGCSE technical symposium on computer science education (SIGCSE '04)* (pp. 255–259). Norfolk, VA.

Vessey, I., Ramesh, V., & Glass, R. (2005). A unified classification system for research in computing disciplines. *Information and Software Technology*, *47*, 245–255.