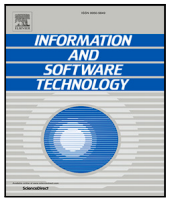
Information and Software Technology 145 (2022) 106840

Contents lists available at ScienceDirect

Information and Software Technology

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



Short communication: Evolution of secondary studies in software engineering

David Budgen [a,*], Pearl Brereton [b]

[a] Durham University, Department of Computer Science, Durham DH1 3LE, United Kingdom of Great Britain and Northern Ireland
[b] Keele University, School of Computing & Maths, Staffordshire ST5 5BG, United Kingdom of Great Britain and Northern Ireland

ARTICLE INFO

Keywords:
Systematic review
Mapping study
Qualitative study
Experience of authors

ABSTRACT

Context: Other disciplines commonly employ secondary studies to address the needs of practitioners and policy-makers. Since being adopted by software engineering in 2004, many have been undertaken by researchers.
Objective: To assess how the role of secondary studies in software engineering has evolved.
Methods: We examined a sample of 131 secondary studies published in a set of five major software engineering journals for the years 2010, 2015 and 2020. These were categorised by their *type* (e.g. mapping study), their *research focus* (quantitative/qualitative and practice/methodological), as well as the experience of the first authors.
Results: Secondary studies are now a well-established research tool. They are predominantly qualitative and there is extensive use of mapping studies to profile research in particular areas. A significant number are clearly produced as part of postgraduate study, although experienced researchers also conduct many secondary studies. They are sometimes also used as part of a multi-method study.
Conclusion: Existing guidelines largely focus upon quantitative systematic reviews. Based on our findings, we suggest that more guidance is needed on how to conduct, analyse, and report qualitative secondary studies.

1. Introduction

When secondary studies in the form of systematic reviews were first identified as a potentially useful tool for strengthening empirical knowledge about software engineering practices, ideas about their use were largely based upon the ways that they are employed in other disciplines. In particular, disciplines such as clinical medicine and education make extensive use of hypothesis-testing *quantitative* reviews, with synthesis involving aggregation of field studies in the forms of randomised controlled trials (RCTs) and experiments. Hence, although the potential usefulness of qualitative secondary studies such as *mapping studies* as preparation for PhD study was recognised, guidelines on performing secondary studies tended to focus upon quantitative forms of synthesis, such as meta-analysis [1].

Several *tertiary studies* have been performed to examine different aspects of the secondary studies being published in software engineering. We ourselves have examined the extent to which published findings contain material that can be employed in teaching about software engineering (and hence, implicitly, about practising it) [2]. And for those secondary studies relevant to teaching and practice, we noted that many of the primary studies that were conducted as field studies employed some form of qualitative case study [3]. In conjunction with this, we also noted that relatively weak forms of synthesis were being used to aggregate the knowledge from the primary studies.

Anecdotally, there was an impression that many secondary studies were led by relatively inexperienced researchers.

We have investigated how the use of secondary studies in software engineering has evolved, in terms of both the degree of experience of those who are conducting them, and also the forms of study used. To do so, we examined a set of reviews taken from both our previous study and also an ongoing tertiary study, published in five journals that are major sources of papers on secondary studies (see Section 3). Our aim was to use secondary studies published in the years 2010, 2015 and 2020 to seek insight into the following questions.

1. Has the proportion of inexperienced researchers as the first (leading) authors of secondary studies changed in the last decade?
2. What forms of secondary study have been used?
3. Are the forms of secondary study used different for lead authors who have different levels of experience?

* Corresponding author.
E-mail addresses: david.budgen@durham.ac.uk (D. Budgen), o.p.brereton@keele.ac.uk (P. Brereton).

https://doi.org/10.1016/j.infsof.2022.106840
Received 28 July 2021; Received in revised form 6 January 2022; Accepted 7 January 2022
Available online 22 January 2022

In terms of their form, studies were classified by:

- *type:* systematic review, mapping study, tertiary study, multivocal review (a form seeking to include various forms of grey literature [4]);
- *research focus:* practice (quantitative), practice (qualitative), methodological.

With regard to *research focus*, we used a two-dimensional categorisation of papers. On one axis we classified them as investigating *practice* or *methodology* (addressing some aspect of the conduct of systematic reviews). On the other, we classified them as being quantitative or qualitative. However, as we were primarily interested in papers addressing practice, we simply counted the total number of methodology papers.

Our indexing of secondary studies for use in tertiary studies has indicated that the number of secondary studies published in any one year does not deviate significantly from a gradually rising curve. So, by choosing the years 2010, 2015 and 2020, we aimed to ensure that our findings reflected current practice (2020), and also spanned a period when ideas about conducting secondary studies were well established (2010–2020). The year 2015 provides an intermediate point in the evolutionary process.

2. Method

Since our study was methodological in nature we did not seek to find all of the secondary studies published in this period, confining our search to five journals considered to be major sources of published secondary studies. These were: *Empirical Software Engineering* (EMSE); *IEEE Transactions on Software Engineering* (TSE); *Information & Software Technology* (IST); *Journal of Systems & Software* (JSS); and *Journal of Software: Evolution & Process* (JSEP). The secondary studies were identified through manual inspection, supported by an electronic search. For inclusion we sought "any form of secondary study as well as those studies using the EBSE methodology but not explicitly making reference to the EBSE literature".

Data extraction was performed by the two authors, working independently and then discussing any differences in order to resolve them.

To assess author experience, we examined the set of publications for each first author as listed in the DBLP database maintained by the University of Trier (www.uni-trier.de). For each paper we checked that the author and the paper were included in DBLP. We then counted the number of papers the first author had published prior to the secondary study, using the following counting rules:

- journal and conference entries up to a maximum of 10, excluding 'informal' [i] entries and excluding earlier conference versions of the secondary study;
- where there were more than ten papers, we recorded this as 'many'.

For the purpose of assessing the experience of the first author, we categorised the author as:

- *inexperienced*, if there were no prior papers;
- having *limited experience* where the first author had up to (and including) five other papers;
- being *experienced* where an author had published more than five papers.

(These were chosen on the basis that we felt that most postgraduate students were unlikely to have more than five previous papers.) We also sought information about author experience from other sources such as the 'author bio' included in some journals.

In addition, we categorised the studies themselves. Where possible this was done using the abstract, although we often also had to consult the full paper.

Table 1
Studies by journal and year.

Journal	2010	2015	2020	Total
IST	12	30	23	65
JSS	1	12	23	36
EMSE	1	6	5	12
JSEP	0	6	7	13
TSE	1	1	3	5
All secondary studies	15	55	61	131

Table 2
Experience of first authors.

Experience category	2010	2015	2020	Total
Inexperienced	0%	9%	13%	10%
	0	5	8	13
Limited experience	33%	35%	30%	32%
	5	19	18	42
Experienced	67%	56%	57%	58%
	10	31	35	76
All	15	55	61	131

3. Conduct of study selection and data extraction

We identified 131 papers by a combined manual and electronic search [dataset] [5]. This included a small number that described the process of a secondary study as their research method, but didn't actually use relevant terms or refer to the evidence-based literature.

All of the papers (and authors) had entries in DBLP, giving confidence in its comprehensive coverage of publications. In most cases it was relatively straightforward to count and categorise an author's prior publications. In a few cases, DBLP was unable to disambiguate between many authors with the same name. Where this occurred, we counted only those papers which had one or more of the same co-authors as the secondary study.

We also noted that categorisation of a study by the original authors was generally unreliable. Many studies that were described as systematic reviews were very clearly mapping studies.

We consider that *risk of bias* (threat to validity) is most likely to arise from our assessment of a lead author's degree of experience. Our measure of lead author experience relates only to the volume of their published work, not to its relevance to the topic of the secondary study. And we only assess the experience of the lead author, which may not correctly reflect the overall degree of experience and expertise available in a review team.

4. Findings from our analysis

The continuing growth in the number of published secondary studies makes it inappropriate to compare actual counts across the three years. So, while we do report these, we make most of our comparisons using percentages, where these are computed on a 'per year' basis. The profiles for the secondary studies across years and journals are summarised in Table 1.

To answer our first question about the perception that the leading authors were more likely to be inexperienced researchers, we looked at the number of previous papers published by the leading authors. Our findings are shown in Table 2.

What this shows is that while our expectation of there being an increase in the number of secondary studies led by first authors who are relatively inexperienced is correct, the degree of the change is limited. Many secondary studies are led by experienced authors.

Our second question asks about the nature of the secondary studies being conducted in software engineering. Table 3 provides a breakdown of the secondary studies by *type*, while Table 4 gives the breakdown by the *research focus* of the study.

Table 3
Studies by type and year.

Type of study	2010	2015	2020	Total
Systematic review	86%	11%	3%	16%
	13	6	2	21
Mapping study	7%	87%	85%	77%
	1	48	52	101
Tertiary study	7%	2%	7%	5%
	1	1	4	6
Multi-vocal review	0%	0%	5%	2%
	0	0	3	3
All secondary studies	15	55	61	131

Table 4
Studies by research focus and year.

Research focus	2010	2015	2020	Total
Practice (Quantitative)	7%	5%	3%	5%
	1	3	2	6
Practice (Qualitative)	93%	95%	94%	94%
	14	52	57	123
Methodological	0%	0%	3%	1%
	0	0	2	2
All secondary studies	15	55	61	131

Looking at Table 3 it is noticeable that the number of secondary studies that can be classified as systematic reviews has declined, while the number of mapping studies has increased. (The number of tertiary studies has also increased slightly, probably because there are now more secondary studies available; and the number of multi-vocal reviews is small because these are a relatively new type of review.) Not surprisingly therefore, the number of qualitative reviews (Table 4) has increased, and the number of quantitative reviews has remained low. We can therefore conclude that secondary studies performed by software engineering researchers, experienced or inexperienced, are predominantly in the form of qualitative mapping studies.

While many of these mapping studies address practice-related research trends, some do address practice. (In [2] we did note that the findings from qualitative studies could be used to provide guidance based upon the experiences of others, particularly regarding such aspects as identifying barriers to adoption or a technology, or lists of issues to consider for adopting a new technology.)

Sixteen (12%) reports incorporated a secondary study as part of a multi-method research approach, combining a secondary study with some other empirical form, commonly a survey or a panel of experts. While only one such study was reported in 2010, there were 7 in 2015 and a further 8 in 2020.

Another interesting characteristic is the number of people involved in performing a given secondary study. In Fig. 1 we have charted the most common numbers of authors involved in conducting the studies. (Again, we have analysed all three years together as there seemed to be no evident trends in this with time.) The only two studies not included in the figure were outliers that had 9 and 12 authors respectively.

We were surprised to find five papers that had a single author, given that conducting a secondary study is usually a team activity. On closer inspection, the one from 2010 used an existing data set; the two from 2015 both acknowledged some un-named assistance, particularly for the purposes of checking; and the two from 2020 were simply performed by one person. All were performed by experienced researchers (scoring "many" for previous papers). There was only limited acknowledgement of the possible biassing effects of performing a study with one researcher.

The lack of biographical information made it difficult to determine how many studies were led by postgraduate students. Looking at the 27 studies with two authors, as being likely to be a combination of student and supervisor, only 8 were explicitly identified as having a postgraduate first author, although the count of previous papers

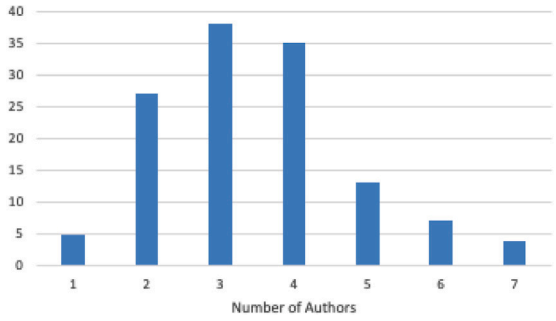


Fig. 1. Number of authors per study.

strongly suggested that at least 8 of the others were likely to be the same, while 8 were clearly performed by experienced researchers.

We concluded that we could not provide a useful answer to our third question about the types of study performed by authors with different levels of experience, given that so many were qualitative mapping studies.

5. Conclusions

As we observed in the introduction, early expectations for the emergence of useful empirical knowledge about software engineering from the adoption of secondary studies as a research tool were largely based upon the ways that these were used in other disciplines. In particular, it was expected that much of this knowledge would be derived from hypothesis-driven quantitative systematic reviews. However, our findings suggest that, while adopting secondary studies quite extensively, software engineering researchers have employed them differently to other disciplines, and perhaps in part, differently to expectation.

To begin with the most striking conclusion: most secondary studies are in the form of qualitative mapping studies. There is very little use of quantitative systematic reviews, despite extensive use of the term "systematic review" in the titles of papers. It can be argued that this is an appropriate profile for a "design science" discipline such as software engineering, where the goal of research is "to develop scientific knowledge to support the design of interventions or artefacts by professions" [6]. In our context, the scientific knowledge that is needed involves devising ways of codifying and conveying the experience of how to use software development techniques and strategies effectively in different situations. Hence studies that provide an analysis of factors and issues to be considered when using software engineering practices can provide a useful way to convey experiences to students and developers [2].

Our second conclusion relates to who it is that conducts the secondary studies. While the proportion led by relatively inexperienced authors has increased with time, probably reflecting a widespread recognition of the value of beginning postgraduate study by conducting a formal literature review, many secondary studies are led by quite experienced researchers. And obviously, such researchers are often involved as additional authors with inexperienced authors.

The emergence of the use of secondary studies as part of multi-method studies is also an interesting development which might point to an important future role.

What we can observe therefore, is that while all this may give us greater confidence regarding the extent to which evidence-based findings provided by secondary studies are based upon well-conducted research, there is a need to provide more guidance for both experienced and inexperienced researchers on effective ways to synthesise and report findings from qualitative studies, especially those that address topics relevant to teaching and practice. Doing so may then help to make them more useful for teaching and may assist with gaining greater acceptance by practitioners [7].

CRediT authorship contribution statement

David Budgen: Conceptualization, Investigation, Writing – original draft. **Pearl Brereton:** Investigation, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Our thanks to Professor Barbara Kitchenham for advice and observations, as well as for conducting independent electronic searching for publications that we might have missed. Our thanks also to the referees for their helpful observations.

References

[1] B. Kitchenham, D. Budgen, P. Brereton, Evidence-Based Software Engineering and Systematic Reviews, Chapman & Hall, 2015.
[2] D. Budgen, P. Brereton, N. Williams, S. Drummond, What support do systematic reviews provide for evidence-informed teaching about software engineering practice? E-Informatica 14 (1) (2020) 7–60.
[3] D. Budgen, P. Brereton, N. Williams, S. Drummond, The contribution that empirical studies performed in industry make to the findings of systematic reviews: A tertiary study, Inf. Softw. Technol. 94 (2018) 234–244.
[4] V. Garousi, M. Felderer, M. Mantylä, Guidelines for including grey literature and conducting multivocal literature reviews in software engineering, Inf. Softw. Technol. 106 (2019) 101–121.
[5] D. Budgen, P. Brereton, Evolution of secondary studies in SE, 2022, http://dx.doi.org/10.17632/gvt6y7sdjy.1, Mendeley Data, V1.
[6] J.E. van Aken, Management research based on the paradigm of the design sciences: The quest for field-tested and grounded technological rules, J. Manage. Stud. 41 (2) (2004) 219–246.
[7] P. Devambu, T. Zimmermann, C. Bird, Belief & evidence in empirical software engineering, in: Proc. 38th Int. Conf. On Software Eng. (ICSE), 2016, pp. 108–119.