Using Systematic Literature Reviews to Enhance Student Learning

Prof. Branimir Pejcinovic, Portland State University

Branimir Pejcinovic received his Ph.D. degree from University of Massachusetts, Amherst. He is a Professor and former Associate Chair for Undergraduate Education at Portland State University, Electrical and Computer Engineering department. In this role he has led department-wide changes in curriculum with emphasis on project- and lab-based instruction and learning. His research interests are in the areas of engineering education, semiconductor device characterization, design and simulation, signal integrity and THz sensors. He is a member of IEEE and ASEE.
Using Systematic Literature Reviews to Enhance Student Learning

I. Introduction and motivation

Literature review is a skill assumed to be in the arsenal of all graduate students pursuing thesis options at the MS or PhD level. There are many resources on writing literature reviews, from campus writing centers to books such as Machi and McEvoy.¹ One would also assume that this is among the very first tasks that research-oriented students would undertake. However, our brief and preliminary survey of graduate students in our electrical and computer engineering department showed that they have very little to no experience in performing literature reviews, and discussions with other faculty confirmed this observation. Unlike some other fields, such as social sciences, it also seems that engineering education programs do not place as much emphasis on the development of this skill. The most obvious use of training graduate students in literature reviews is in helping them satisfy dissertation or thesis requirements. Literature reviews, however, have other uses, such as starting a new research area by identifying holes in the existing literature or summarizing one’s own research area. Recently, it has been argued that a variant of literature review, so-called “systematic literature review” (SLR) can help students publish their first original work and transition them from novice to knowledgeable.²,³ Finally, systematic literature reviews have become a research area by themselves, although they are less common in engineering than in areas like medicine, psychology or education.

For all these reasons it is appropriate to intentionally train and educate students in performing literature reviews in general and SLR in particular. One possible approach, taken by many departments, is to design a research methods course that also covers literature review topics. Experience with other so-called soft-skills, such as technical writing, suggests that it is very important to provide a specific disciplinary context for learning technical writing “… so that students appreciate it as part of their professional engineering skills, not a skill separate from them.”⁴ Therefore, learning how to do literature reviews and SLR can best be accomplished by their incorporation in various courses across the engineering curriculum. For now, however, we will concentrate at the course-level implementation.

Recent literature in the area of software engineering has advocated using SLR as a more generic educational tool, potentially suitable even for undergraduate student use.⁵ A variation of SLR called interactive SLR (iSLR) has been published very recently⁶ arguing for some flexibility in the protocol design so that iterations on some key components can be accomplished. Satisfactory results with undergraduates were reported. Originally, the development of the iSLR protocol took several iterations and was done with small groups of graduate students.

In this report we will present a pilot study demonstrating that iSLR is a useful and practical educational tool that can and should be done in the context of a specific course problem and not
as a generalized approach, as is usually done in research methods courses. We propose making iSLR part of a research-like project on a specific subject matter covered in a course. When set up in this fashion, we believe that educational benefits will include improved critical thinking and writing, increased motivation, improved life-long learning skills, better topic coverage, and increased depth of coverage. Ideally, iSLR would be introduced into the wider curriculum and would address student skills and abilities that are difficult to attain in regular coursework.

The rest of the paper is organized as follows: Section II gives an overview of uses of SLR in other disciplines, especially in medicine and software engineering. Section III discusses uses of iSLR as pedagogical tool in engineering and includes implementation details. Section IV is devoted to assessment methods and results. Finally, section V discusses lessons learned, offers some conclusions, and points to future work.

II. Systematic literature reviews in different disciplines

The topic of writing literature reviews is not new and there are general and field-specific books that cover the process. Typically, these books are aimed at graduate students preparing their theses or dissertation proposals, but they do not discuss SLR- or iSLR-based approaches. For our discussion of benefits of iSLR as pedagogical tool, it is important to properly distinguish it from other forms of review. In that context it is also important to understand more broadly where SLR comes from, its history, and how it is used.

One discipline using SLR extensively is medicine where SLR is treated as a research discipline by itself. Typically, the purpose of SLR is not to just summarize the state-of-the-art at a given point in time, but also to provide meta-analysis of available data, which then leads to some conclusions and policy decisions. Given the potential impact and importance of such studies, SLR practitioners and other constituents felt there was a need to provide specific guidance with respect to how such studies should be performed and reported. This has resulted in two statements: QUORUM (Moher et al.⁷) and PRISMA (Moher et al.⁸). PRISMA statement defines SLR as:

\[
\text{A systematic review is a review of a clearly formulated question that uses systematic and explicit methods to identify, select, and critically appraise relevant research, and to collect and analyze data from the studies that are included in the review. Statistical methods (meta-analysis) may or may not be used to analyze and summarize the results}^{8}\.
\]

The PRISMA statement provides guidelines on seven areas that SLR studies should address: Title, Abstract, Introduction, Methods, Results, Discussion, and Funding. There is a total of 27 items in a checklist format. For example, it is required that an SLR study:

- Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.
• State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis). 

A similar attempt has been made to codify SLR in the software engineering area where the following definition is used:

*At systemic review is a means of evaluating and interpreting all available research relevant to a particular research question, topic area, or phenomenon of interest. Systematic reviews aim to present a fair evaluation of a research topic by using a trustworthy, rigorous, and auditable methodology.*

While their emphasis and wording is different, both definitions are attempting to explain what “systematic” means and implicitly distinguish such studies from other, less well defined approaches to literature review. Since 2004, procedures and guidelines have been available on how to conduct SLR in software engineering. This has even resulted in one systematic literature review of systematic literature reviews in software engineering.

Most engineers and engineering educators are more familiar with a different kind of literature review: narrative review. Narrative review is meant to provide an overview of a given field and is written by a recognized expert in that field. Compared to a systematic literature review, the main differences lie in the areas of problem definition and methodology. Table 1 below is adapted from the field of evidence-based medicine and it summarizes the main differences between the two review approaches.

**Table 1. Summary of main differences between systematic literature reviews and narrative reviews.**

<table>
<thead>
<tr>
<th>Systematic Reviews</th>
<th>Narrative Reviews</th>
</tr>
</thead>
<tbody>
<tr>
<td>Investigate a clearly defined topic or question.</td>
<td>Intended to provide an overview of an area.</td>
</tr>
<tr>
<td>Literature is gathered using explicit search protocols.</td>
<td>Explicit, systematic literature search protocol not used.</td>
</tr>
<tr>
<td>Studies selected using a protocol that specifies inclusion, exclusion criteria.</td>
<td>Studies used are not selected according to an explicit, predetermined protocol.</td>
</tr>
<tr>
<td>Data from primary study may be synthesized in a meta-analysis. Evidence &quot;grades&quot; may be applied to individual studies.</td>
<td>Evidence &quot;grades&quot; may be applied to individual studies.</td>
</tr>
<tr>
<td>When evidence is lacking, the authors usually recommend further research.</td>
<td>When evidence is lacking, the authors make recommendations based on their opinions and experience.</td>
</tr>
</tbody>
</table>

Recently, Borego et al. have argued that SLRs are a very useful tool that should be applied more broadly in the field of engineering education research. They also provide a very useful and detailed explanation of methodology for proper application of SLR in engineering education.
research and point out a major difference between SLR and narrative reviews as “... narrative reviews differ from systematic reviews in that the identification and selection criteria for sources are usually implicit; narrative reviews typically do not include methods sections” (Borego et al. 13 66).

Use of SLR in engineering education is a more recent development and is discussed next.

III. SLR as pedagogical tool in engineering

To the best of our knowledge, SLR as a pedagogical tool has only been used in the software engineering area. The most recent report discussed development of iterative SLR (iSLR) and its educational benefits, while an earlier study described successfully teaching undergraduate students some software engineering skills and concepts. We decided to implement iSLR due to its flexible approach, which is suitable for novices in a given area of study.

Several assumptions underlie our implementation of iSLR:

1. iSLR can be taught to and performed by subject area novices.
2. iSLR can be a useful learning tool in a given subject area.
3. Students learn best in specific context (e.g., course on a defined EE topic)
4. Proficiency is attained through repeated performance, which means that iSLR should be implemented across the curriculum.

Items 1 and 2 have been demonstrated for software-engineering. Item 3 is based on best practices for similar skills, e.g., technical writing, but we have no direct evidence yet that would support it for iSLR implementation. Item 4 is also based on best curricular practices but collecting evidence for it would require a much broader scope than this pilot study.

In this pilot study we followed procedures discussed in Lavalle et al. with a few modifications, as explained later. There are eight stages in the iSLR process:

1. Review planning: Plan the review effort and training activities.
2. Question formulation: Define the research questions.
3. Search strategy: Define the review scope and search strings.
4. Selection process: Define inclusion and exclusion criteria.
5. Strength of the evidence: Define what makes a high quality paper.
6. Analysis: Extract the evidence from the selected papers.
7. Synthesis: Structure the evidence in order to draw conclusions.
8. Process monitoring: Ensure the process is repeatable and complete.

Furthermore, Lavallee et al. propose modifications to the usual SLR practice by allowing iterations between different stages, e.g., between stages 2 and 3. This is especially important for pedagogical purposes because student understanding of the process itself and of the topic improves over the course of the project, and students should be allowed to revisit their initial
decisions. For example, finding too many references during the initial search may indicate that the question was defined too broadly and it needs to be modified. One modification that we introduced deals with the Search Strategy stage. Instead of letting students come up with search strings right away, we provide them with one seed article that they use for forward and backward snowballing, i.e., looking up references cited in that article and looking up papers citing that article. Our intent is to provide some scaffolding to students so that they are not immediately overwhelmed and to help them formulate the initial question.

III.a Implementing iSLR

Initial implementation of iSLR was done in a Solid-State Electronics I graduate course, which is taken by MS and PhD electrical engineering students. The course covers many common solid-state physics topics such as band theory of semiconductors, conduction in metals and semiconductors, and carrier transport in classical and semi-classical approaches. Within this course, students undertake characterization of very thin metal films using THz Time-Domain-Spectroscopy (TDS) methods as a research project. This naturally led to an iSLR project related to literature on the topic of “TDS characterization of thin metal films.” A total of seven students took the course in the Fall 2014 quarter, and they were divided into three teams (2+2+3). Each team was given a different starting paper.

Figure 1 Example of a well-organized set of references obtained during iSLR project.

Each team was also set up as an online group in Zotero\textsuperscript{14} so that students could share papers they found and do the sorting using directories and annotation features provided by Zotero. This made collaboration on paper search and selection very easy and transparent. An example of a well-
organized submission on Zotero is given in Figure 1. The final product was a summary review report along with annotated bibliography.

Specific tasks for the instructor in our implementation included:

1. Defining a general research-like topic which will contain iSLR as part of it
2. Defining teams
3. Setting up collaborative tools (Zotero)
4. Selection and distribution of initial paper(s)
5. Engaging engineering area librarian
6. Following weekly program of tasks and deliverables (i.e. our protocol)
7. Weekly meetings with students
8. Assessment of final reports based on rubric

In order to define a protocol in item 6 above, each stage in iSLR is broken down into a more detailed list of specific tasks, e.g., for items 3. Search Strategy and 4. Selection Process we have:

a) Perform snowballing search from the starting paper and deliver
   a. Raw list of references, (this should be exported from Zotero in some electronic format for future inclusion in written documents)
   b. Selection criteria for eliminating / keeping papers from that list
   c. List of references after selection; each eliminated paper should have a comment or code explaining why it was eliminated.
   d. Suggestions for possible refinement of research question

b) Perform database literature search based on keywords and deliver:
   a. Raw list of all papers
   b. Selection criteria for eliminating / keeping papers from that list (can be the same as the one used for snowballing)
   c. List of references after selection; each eliminated paper should have a comment or code explaining why it was eliminated.
   d. Suggestions for possible refinement of research question

c) Combine references from a) and b) into a single list

Underlined tasks indicate opportunities for iterative improvement of the research question – the “i” in iSLR. This is implemented in a weekly protocol, which specifies tasks and deliverables. For example, for the fifth week we have:

- Finalize the problem statement (last chance to refine it)
- Divide the list of papers from the database search among group members
- Perform selection (use titles, keywords and abstracts)
- Annotate and code papers as selection is done (tag)
- Assignment for next week:
  - Report on the total number of papers found and number of eliminated ones
  - Produce a diagram explaining the core idea or concept and how it is divided into sub-concepts.
  - List themes that you observed, if applicable.
Report on how you are doing coding, i.e., which tags are used.

We attempted to locate a rubric to help with scoring the final report and the best we found so far is by Green and Bowser\textsuperscript{15}. It was developed as a general tool but primarily for social science and education. We modified it only slightly to make it more applicable to engineering.

**IV. Initial assessment results**

Pre-course survey was done at the beginning of the course to establish students’ familiarity with any type of literature review process and their self-efficacy with respect to performing literature reviews of any type. The first two questions asked about previous experience:

1. How many times have you done a literature review or something similar to it? Include reviews that you may have done outside of technical coursework or technical thesis work, e.g., in science classes (physics, sociology etc.). Exclude the current one. (Scale: Never, Once, Twice, 3 or more.) Six out of seven students responded Once or Never.

2. How many times have you done literature review in technical areas alone? (excluding current one). Scale: Never, Once, Twice, 3 or more. Six out of seven students responded Once or Never.

This result indicates a very poor familiarity with literature review. Conversations with other faculty provide anecdotal support for this observation, i.e., that students are generally unprepared to perform literature reviews. In the future we plan to survey our entire graduate student population. The next set of pre-course questions, shown in Table 2, dealt with self-efficacy and the same questions were asked in a post-course survey.

**Table 2. Results of pre-assignment student survey assessing their preparedness and confidence in their ability to perform literature reviews.**

<table>
<thead>
<tr>
<th>Question</th>
<th>Strongly agree</th>
<th>Agree</th>
<th>Neutral</th>
<th>Disagree</th>
<th>Strongly disagree</th>
<th>Not applicable</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I am familiar with literature review process</td>
<td>3</td>
<td>1</td>
<td></td>
<td></td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>2. I can explain various stages in doing literature review</td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3. I am confident that I can do a literature review on my own</td>
<td>2</td>
<td>2</td>
<td></td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>4. Learning how to do literature review will be valuable in my studies</td>
<td>4</td>
<td>2</td>
<td></td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>5. Learning how to do literature review will be valuable in my current workplace (if not currently employed, select NA)</td>
<td>3</td>
<td>3</td>
<td></td>
<td>0</td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>
Students expressed some familiarity with the literature review process, but given that they claimed to have used it at most once, it seems likely that they were complete novices in the area of literature reviews. This was confirmed by discussions with students where some expressed outright fear of this project. Six out of seven students worked in outside jobs in some capacity, and they believed that literature review would be helpful in their work as well as their studies.

For the post-course survey, we added questions related to perceived effectiveness of various iSLR stages, as given in Table 3, and usefulness of the electronic reference management system (Zotero). For the former, students were asked to select their top two choices so that we can get a somewhat wider data set. The top two selections are indicated in Table 3, where the number indicates the number of times a given column was selected either as the first or second choice.

Table 3. Student assessment of difficulties encountered during iSLR and teaching effectiveness.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>7. Which stage did you find the most time consuming?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Which stage did you find the most challenging or confusing?</td>
<td>2</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Which stage was explained the best during class?</td>
<td></td>
<td></td>
<td>6</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Which stage was explained the worst during class?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td></td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

Due to the small sample size, it is impossible to draw definitive conclusions, but these results are good enough to inform improvements in the process. It seems that the Selection Process was the most time consuming, so we plan to introduce some exercises to familiarize students with effective and efficient ways of doing it. The Search Strategy stage was explained by a research librarian, who has a lot of experience in explaining this topic to a variety of audiences. Clearly, collaboration with library staff is beneficial to student success. On the other hand, the Synthesis stage comes at the end, and it requires critical thinking and writing skills, which are difficult to explain in a traditional classroom setting. In prior work, it was found to be among the most mentally demanding for novices. Performance in this stage could be improved by smaller writing assignments along the way, which we plan to implement in the future.

From our surveys, it was possible to measure the shift in student self-efficacy in performing iSLR and its perceived usefulness, as shown in Table 4. Shift is measured by assigning unit
change if the answers changed between neighboring categories. For example, change from Neutral to Strongly agree gives a “+2” shift while change from Neutral to “Strongly Disagree” gives a “-2” shift.

Table 4. Shifts in student assessment of their preparedness and confidence in their ability to perform literature reviews, and perception of usefulness of iSLR.

<table>
<thead>
<tr>
<th>Gain (shift) in student selections relative to pre-course survey</th>
<th>-2</th>
<th>-1</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I am familiar with literature review process</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>2. I can explain various stages in doing literature review</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>3. I am confident that I can do a literature review on my own</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>4. Learning how to do literature review will be valuable in my studies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>5. Learning how to do literature review will be valuable in my current workplace (if not currently employed, select NA)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
</tbody>
</table>

We can make two observations based on Table 4. The first one is that student self-efficacy, as measured by the first three questions, has improved, which is a positive finding. The second one is that the perceived usefulness of iSLR declined, which was an unexpected and potentially negative finding. Without more data we can only speculate that this was due to very optimistic initial expectations, which got tempered by demands of actually performing the tasks required for iSLR. This speculation is supported by comments in surveys, which indicated that the iSLR project was useful but time consuming in an already fairly difficult course. However, some of this can also be attributed to less than perfect time management. Finally, students found learning the Zotero software to be valuable, and we will continue using it as part of the iSLR process. Our own observation is that as the quarter progressed, students became more comfortable with the idea of reading and analyzing papers in an unfamiliar area.

To help us with grading and to better and more consistently judge the quality of submitted reports, we initially adapted an existing rubric which is available online. To the best of our knowledge, there are no validated instruments available for this task, especially within the engineering education context. The main components of this rubric are grouped in the areas of Content, Presentation, and Writing/Format, as given in Table 5. Scoring scale is: Exemplary (5), Developed (4), Average (3), Undeveloped (2), and Deficient (1). Among the three reports, one was assessed to be between Developed and Exemplary, one was Developed and one in between Average and Developed. This was deemed to be a very good performance for a pilot study.
Table 5. Summary of criteria used in the assessment rubric.

<table>
<thead>
<tr>
<th>Criteria and qualities of iSLR reports</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Content</strong></td>
</tr>
<tr>
<td>• Historical and theoretical background. Seminal literature.</td>
</tr>
<tr>
<td>• Breadth of subtopics</td>
</tr>
<tr>
<td>• Quality of literature</td>
</tr>
<tr>
<td>• Relevance of published studies to current topic</td>
</tr>
<tr>
<td>• Relevance of published studies to each other</td>
</tr>
<tr>
<td><strong>Presentation</strong></td>
</tr>
<tr>
<td>• Organization</td>
</tr>
<tr>
<td>• Transitions</td>
</tr>
<tr>
<td>• Current study rationale and contribution</td>
</tr>
<tr>
<td><strong>Writing/Format</strong></td>
</tr>
<tr>
<td>• Clarity of writing and interpretation of literature</td>
</tr>
<tr>
<td>• Bibliographic format</td>
</tr>
</tbody>
</table>

However, this rubric was originally developed and used in non-engineering fields and for generic literature reviews, which in practice limits its usefulness in evaluating iSLR reports. Therefore, modifications will be needed to line it up with iSLR stages by adding, for example, explicit criteria related to Question Formulation, Search, Selection and Analysis stages. This work is in progress.

V. Lessons learned, conclusions and future work

This was a pilot study and we expected that some refinements would be needed, which would be guided by assessment results and observations. We learned several lessons that will inform future revisions and implementations:

- Students have to stay on task as outlined in the protocol. It is very easy for them to get lost in the minutia of finding proper search strings or trying to understand every detail in their papers.
- It is best to get Zotero set up right away and follow that up with a presentation from an area librarian. Students can learn software best through repeated use.
- Students should explicitly define their question as early as possible to avoid being sidetracked. For example, one group pursued superconductive thin films, which were not meant to be part of the assignment.
- Progress reports should be submitted periodically in written form or else the usual procrastination sets in.
- Due to a 10-week long quarter our protocol is compressed and less than ideal, but it would be very reasonable to do it within a 15-week semester. We are currently working on implementing iSLR in a two-quarter sequence, which should provide a convenient breakpoint at the Analysis stage so that the 2nd quarter is devoted to more critical
analysis, synthesis, and writing. This extra time may enable submission of early drafts and revisions, which is how writing exercises should be done.

- An iSLR project is not a simple task that can just be added on top of existing assignments and course materials. One should carefully plan how much time students are expected to devote to this project and communicate that to students. Our best estimate is that this project should take around 20% of the total student time. In the next iteration we will attempt to measure this more directly. Note that this does not necessarily result in less time to cover the usual material because the project is structured so that it covers material that would have to be learned through other means.

The overall conclusion from this pilot study is that iSLR is a very promising approach to teach students both the methodology of systematic literature reviews as well as material relevant to the course in question. Student self-efficacy for doing iSLR was improved, but their estimate of its usefulness declined. The Selection stage was found to be the most time consuming, while Question Formulation was the most challenging for students. More explanation will have to be devoted to the Synthesis stage. Students generally supported the introduction of the iSLR project but voiced concerns about additional time required. Implementation of iSLR in one 10-week quarter is possible, but the schedule must be followed exactly. A 15-week long semester implementation would allow some tasks to be expanded and iterated.

In our future work, we will fix items mentioned in the lessons learned section. The assessment rubric also needs improvement to align it with an engineering context and with the systematic literature review process. We will also implement iSLR in a two-quarter circuit design course, which will require a different approach in the initial stages. Finally, we need to either develop or implement some existing assessment instruments to find whether iSLR actually produces the anticipated goals of improved critical thinking, better writing, increased motivation, and improved life-long learning skills. We also need to evaluate consistency of the reports – did they focus on the same literature and did they formulate their questions in a consistent fashion? Since our primary goal was the educational effects of iSLR, we have not yet attempted to address this question of reliability and validity of studies produced by students, but will do so in the future.

References


