International Journal of Engineering & Technology, 7 (3.20) (2018) 35-39

SPC

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Theoretical Underpinnings of Learner Engagement in Software Visualization System: A Systematic Literature Review Protocol

Abdullah Al-Sakkaf, Mazni Omar and Mazida Ahmad

Human-Centered Computing Research Lab, School of Computing, Universiti Utara Malaysia.
Corresponding Author Email:abbdullah1@hotmail.com

Abstract

Having a solid theoretical foundation is essential for designing an effective software visualization (SV) tool. Despite the decades of developing different SV tools, there are still doubts about their effectiveness. Furthermore, learner engagement plays an important role in building a successful SV tool. In programming education, the problem of the high failure rates among students is still unresolved. Therefore, there is a need to understand the theories behind the exciting SV tools from the engagement perspective in order to have a road map for future tool construction. Yet the factors influencing learner engagement in SV tools are still unclear. This study set out to determine how to develop an SV design model to enhance student engagement in an introductory programming course. A systematic literature review (SLR) was used to obtain an overview of the current theoretical foundation used. The search identified a total of 432 papers between 2011 and 2017. This study examined 58 papers selected based on a well-defined selection process. In this paper, the contribution in constructing the protocol for SLR is presented as well as the preliminary results of the study. The researchers were in the process of data extraction phase to address the research questions. The expected outcomes of this review became the identification of a theoretical background used to construct and explain engagement in software visualization. The expected output of this study was a list of the factors that have a positive impact on learner engagement in SV tools.

Keywords: Engagement; Introductory programming education; Software visualization; Program visualization.

1. Introduction

Teaching programming is one of the widely-studied areas in computing education research (CER). Despite decades of research to decrease the failure and dropout rates in programming courses, where the failure rates are higher as compared to other courses, problems remain in existence [1]. In the same vein, many software visualization (SV) tools were developed in the last three decades to enhance the learning of novice students' programming languages, especially in higher education institutes. Software Visualization (SV) can be defined as "the use of the crafts of typography, graphic design, animation, and cinematography with modern human-computer interaction technology to facilitate both the human understanding and effective use of computer software"[2]. SV could be a valuable resource to help novice students to improve their programming skills. Generally, SV is classified into two types: Programming Visualization (PV), and Algorithm Visualization (AV). Although there are many reports in the literature on the efficiency of SV, there is still a lot of concern about the efficiency of SV and their impact on the learning outcome [3], [4].

There is an increasing attention by researchers on how to engage learners with the visualization system, although further work on engagement is still required [4]. Previous studies suggested that learner engagement is among the most important factors for designing and building an effective and successful SV tool [4], [5]. If the SV tool fails to engage the learners in active learning, it will fail to achieve its goals [6]. The term 'learner engagement' in the area has been used to refer to the learner involvement with the teaching system that includes the interaction with the system (tools), instructor, or other students within the system. Several taxonomies were proposed to explain the learner engagement level of interaction with the visualization system, which are: Engagement Taxonomy (ET) by [6], Extended Engagement Taxonomy (EET) by [7], and 2 Dimension Engagement Taxonomy (2DET) by [4]. These taxonomies hypothesized that using a higher level of engagement will improve learning outcome. Even though previous studies measured the effect of engagement level on learning outcomes, their results were not consistent [8]. Furthermore, these taxonomies still lack the description of the relationship between engagement levels and learning outcomes [9]. Despite the importance of engagement in the successful design of SV tools, the focus on how to improve engagement when constructing SV tools is very limited from the theoretical view. By applying the proper theories and techniques for learner engagement with SV within the learning process, instructors and designers of these tools can improve and increase student engagement which will affect learning outcome. In summary, little is known about engagement in SV and several essential questions remained unanswered in relation to the role of engagement in SV.

It is acknowledged that a solid theoretical foundation is essential for designing an effective educational technology [9]. Established theoretical frameworks have many benefits for researchers (e.g., they allow researchers to make sense of empirical findings). However, different theories may offer competing recommendations and it is important to balance multiple theoretical perspectives in technology design [10]. As computer science (CS) education is an interdisciplinary field, [11] identified that there is a need to link the research to a relevant theory for CS educational research. Likewise,

[3] observed in the PV studies that there is a lack of consideration for learning theories in the theoretical framework building. [12] advocated that the theoretical foundations for creating effective AVs appear to be steadily improving. Theories offer common conceptual frameworks that allow the structuring of knowledge in a concise and precise manner, thus facilitating the communication of ideas and knowledge [13]. Furthermore, there is no agreement on a clear definition of the term “theory” [14]. This paper will use the definition stated by [9], in which theory is defined as “a broad class of concepts that aims to provide a structure for conceptual explanation or established practice and use such terms as “theories”, “models”, and “frameworks” (TMF) to describe particular manifestations of the general concept of theory”. A theory provides explanations and understanding in terms of basic concepts and underlying mechanisms, which constitute an important counterpart to the knowledge of passing trends and their manifestations. It is used to understand why and how a phenomenon occurs [9], [13]. In summary, it is shown from the above review that there is a need for a theoretical understanding of the concept and process of teaching programming.

The overall aim of this review was to investigate how to develop an SV model to enhance student engagement in an introductory programming course; thus, it will help to increase SV educational effectiveness and enhance learning outcomes. This paper presents the protocol used to investigate the theories and factors applied in the SV engagement field.

2. Related Works

Literature reviews in CS education are common where the focus is usually on the areas that lack precision. To date, some literature reviews focused on the theoretical aspects in the CER literature in general such as [9], [15]. Other literature reviews focused only on teaching programming such as [1], [16]. However, there are studies [3], [4] that reviewed the SV literature from different aspects. In addition, several reviews [9], [13], [17] carried out the theoretical foundation in different areas (e.g., CER, software engineering). Table 1 summarizes the most relevant systematic literature reviews (SLR) compared in this study.

Table 1 Comparison of SLR with related works

Author	Period	Papers	Focus
[4]	1979-2012	n/a	Describe PV systems in the last 3 decades
[9]	2008-2011	308	Discover what theories, conceptual models and frameworks in CER literature
[3]	2013-2016	36	Examine learning principles on recent PV systems
Current Study	2011-2017	58	Discover what TMF, and factors to enhance engagement in SV tool

The SLR by [18] claimed that neither educational theories nor other theories from the different fields can be used to explain the process and challenges faced in the teaching and learning in the CER area. Furthermore, there is a lack in fundamental theories to be used in explaining how students learn computing [9]. As SV is part of CER, the same goes for it. The study observed that half of the papers were not build based on previous theoretical foundations. The study selected 308 papers for the period from 2005 to 2011.

[4] analyzed the literature from 1979 to 2012 and identified 46 PV tools developed in that period. Furthermore, the study compared the engagement level of the PV tools using 2DET. The study concluded that the results were unclear with respect to learner engagement. Additionally, this study summarized unstudied problems as the following: the interactions between program visualization tools, learners, learning environments, forms of engagement, and learning objectives. In the same way, a recent study by [3] was carried out on PV to cover PV tools from 2012 to 2015 as a continuation to [4] and identified 16 new PV tools. The study identified the principles that could contribute to the effectiveness of tools based on Vygotsky's learning theory. [3] concluded that very few systems used learning theory as a theoretical foundation.

Results from the earlier SLR studies demonstrated a strong need in the CER literature or more specifically in the SV literature for more empirical studies so as to understand the effectiveness of learner engagement within SV tools. These studies should be constructed or their result be explained based on a clear theoretical background. To date, no review has yet been conducted to underpin the theoretical background for engagement in SV. In conclusion, previous studies showed that the challenge in the theoretical foundation in CER, or SV, still exists.

3. Methods

SLR aims at providing a well-defined process for identifying, evaluating and interpreting all available evidence relevant to a particular research question or topic [19]. In order to achieve the objectives of this study, the protocol was built using the guidelines provided by [19]. The SLR processes used in this study are:

1. Planning the review
    a. Identification of the need for a review
    b. Development of a review protocol
2. Conducting the review
    a. Identification of research
    b. Selection of primary studies
    c. Study quality assessment
    d. Data extraction
    e. Data synthesis
3. Reporting the review

A. Research Questions

Beyond looking for evidence that theories are being used, the key research question of this study was “how to develop an SV model to enhance student engagement in introductory programming courses using SV tools?” To achieve this objective and ensure the collection of all relevant data, several research questions were formulated. These questions helped to ensure that the study was comprehensive in its nature, whilst providing an in-depth analysis into the past use of SV for teaching novice programming to improve student engagement. Specifically, the following questions were addressed:

RQ1. Which theories have been used to explain or construct learner engagement in SV?

RQ2. What are the engagement factors used to design an effective Software Visualization?

RQ3. How to measure the effectiveness of engagement in Software Visualization?

RQ4. To what extent have the SV tools been evaluated in terms of engagement effects?

B. Search Strategy

The search strategy was formulated where the main keywords were engagement and software visualization. Initially, the terms relating to SV were identified, such as program and algorithm visualizations. In the literature, sometimes animation was used instead of visualization (i.e., algorithm animation). The final search string was:

(engagement) AND ("software visualization" OR "software visualisation" OR "program visualization" OR "program visualisation" OR "algorithm visualization" OR "algorithm visualisation" OR "software animation" OR "program animation" OR "algorithm animation")

The search covered articles published between 2011 and the time of searching, i.e., March 2017. The following databases were used in the search to identify and collect relevant manuscripts: (1) IEEE Xplore; (2) ACM Digital Library; (3) Springer Link; (4) Science Direct; (5) Wiley Online; and (6) Scopus. The search strategy was validated by checking the five key papers.

C. Study Selection

The selection process was separated into three stages after removing duplicated titles. The first stage applied the exclusion criteria on articles by titles on the remaining paper set. The second stage involved screening the paper based on its abstract and conclusion. After this screening, papers must be looked at in more detail to finalize the screening process. The final stage was to select the paper based on the full text. The exclusion criteria were applied in the first two stages. In the final stage, the inclusion and exclusion criteria were applied together. Once a found source was included, it was not excluded anymore.

Table 2 Inclusion and exclusion criteria

Inclusion Criteria	1. The paper must be on the use of SV in teaching introductory to programming (CS1) course. 2. The paper must present SV tool(s) or discuss program visualization in general. 3. The paper must report empirical results.
Exclusion Criteria	1. The paper will be excluded if it is duplicated. 2. Books, letters, editorials, and position papers will all be excluded. 3. The paper will be excluded if it is not written in English. 4. The paper will be excluded if full text is not available. 5. The paper will be excluded if it is not a full paper (abstract only, poster, or workshop). 6. The paper will be excluded if it is not related to programming courses. 7. The paper will be excluded if it is not a primary study. 8. The paper does not include SV as defined in this paper. 9. The paper does not involve higher education students.

Table 2 presents the inclusion and exclusion criteria applied to all the retrieved studies. The selected papers must use SV tools in a CS1 course. Besides, the papers must report the empirical result to ensure the quality of articles. The paper is excluded if the system does not fit under the definition of SV as stated above. In addition, a secondary study, poster, an abstract, and workshop papers are excluded. If the article was not in the higher education domain (e.g., primary schools, k-12, ... etc.), it was also excluded.

Table 3 Data extraction form

Item	Description
1. Publication Details	
ID	Publication Identifier
Title	The title of the primary study
Publication year	The year when the primary study was published
Type	Determine the type of publication (e.g.: article, conference, etc.)
Source	The source where primary study found
2. SV Tool	
Tool name	List the SV tool(s) discussed in the article
Tool type	Determine the type of those tool(s) whether PV, or AV
3. Theory / Model / Framework (TMF)	
Name	The name given for the theory by the author if available
Theory overview	Short overview about the theory(s) used in this study
Terminology	Determine whether it is theory, model, or framework
Type of theory	Determine the type of theory
Theory disciplines	From where this theory was originated
Purpose of theory	To explain in more details the purpose of the theory in the study
4. Engagement factor	
Factor	Identify the factor that impact the engagement based in the study
5. Evaluation	
Engagement evaluation	How the engagement evaluated in the study
Learning outcome	What learning outcome factor used to measures the learning outcome

D. Data Extraction

To investigate the theoretical foundation in SV, the data extraction form was designed to collect all the information needed to address the review goals. The data extraction form was made up of five sections as shown in Table 3. The main outcome for this SLR was to identify the fundamental theories used in the area of SV, which is not an easy task. The first section in the form described the demographic data to be used in a descriptive analysis of the results. The second section listed the used SV tools in the papers; additionally, the type of system was determined based on the taxonomy by [2]. The third section described the TMF sections which items are discussed in detail below. The fourth section aimed to identify the factors that have a significant positive impact on learner engagement from the study. Finally, the last section summarized how the evaluation was conducted.

In the extraction form, the TMF section was used to extract the theories used in the study. Given that each paper could present different TMFs in discussing the related works, it is important to ensure that the selected TMFs were clearly used or developed in the paper. Thus, the theory selection focused on the following part of each paper:

1. If there was a dedicated theoretical section.
2. If it was stated in the abstract, introduction, or in the design of study description that the work was based on a specific theoretical background.
3. If the gained result was interpreted and explained using some theories or theoretical background in the discussion section.

Due to the limitation of this study, an explicit theoretical background mentioned in the paper was the focus for the search. There was a concern to find a large portion of papers that did not use or mention the theoretical background for their studies as revealed by [9] that this problem exists in CER research. To identify the theoretical foundation of each paper, the following process proposed by [13] was utilized:

1. Candidacy for theory if there was:
    a. The explicit mention of the terms, theory, model, or framework
    b. The identification of constructs and relationships in a body of conceptual argumentation delineated by diagrams, words, etc.
2. Or it was used to explain the cause-effect relationship in the experiment.

Table 4 A taxonomy of theory types [20]

	Type	Description
1	Analysis	The theory does not extend beyond analysis and description. No causal relationships among phenomena and no predictions are made.
2	Explanation	The theory provides explanations but does not aim to predict with any precision. There are no testable propositions.
3	Prediction	The theory provides predictions and has testable propositions but does not have well-developed justificatory causal explanations.
4	Explanation and prediction	Provides predictions and has both testable propositions and causal explanations.
5	Design and action	The theory gives explicit prescriptions (e.g., methods, techniques, principles of form and function) for constructing an artefact.

The theory has a different use or goal in a research. For example, it could be used to demonstrate how something should be used or developed in practice, or explain a phenomenon. Besides, theories are used to explain the relationships among constructs. [20] proposed a taxonomy to classify the structural nature of theories in information system based on their goals as shown in Table 4. After the theories selection, they were categorized according to where the theory originated (e.g., CER, Computing, Education, Psychology, etc.). This is important to understand how the area of SV was developed. This construct was identified based on the source of theory disciplines.

4. Results and Discussion

The full selection phases and number of papers identified at each stage can be seen in Figure 1. The initial search was conducted in April 2017 resulted in 432 papers. Among them, 67 papers were duplicates. The remaining 365 papers were included in the selection process. In Stage 1, the exclusion criteria were applied to the title of the paper. At the end of this stage, 66 articles were excluded. At Stage 2, 117 articles were excluded based on the abstract and conclusion section of the articles. At the final stage, the remaining 113 articles were screened based on a full-text criterion. In this stage, another 14 articles were excluded. By the end of the selection process, a total of 58 articles were selected.

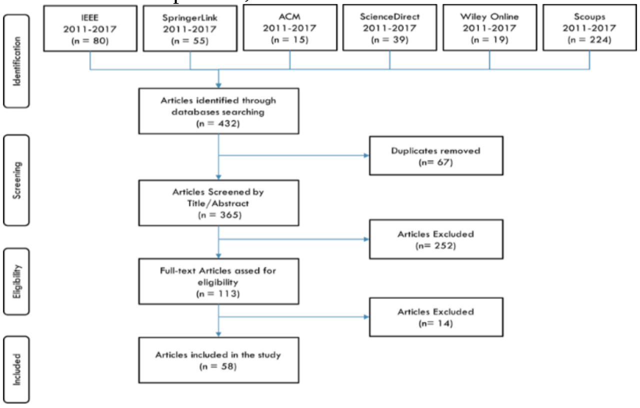

Figure 1: Stages of the study selection process

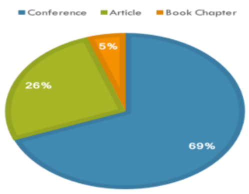

Figure 2: Selected papers by type of paper

A simple statistical analysis was used with the publication of demographic data (section 1 in the data extraction form). Among the selected papers, 69% were conference papers while 26% were articles of journals (see Figure 2). Table 5 shows an overview of the results based on the database for the initial search result and final selected papers. Given that the Scopus database returned was 38% of the selected papers (n = 22), the result from Scopus included papers from the other data sources; however, they did not appear in the original sources due to the way the database search works. It is apparent from this table that Scopus and IEEE were the main sources for the selected papers in this study. Brereton et al. (2007) reported the problem that software engineering is not designed to support SLR, where it interprets the search string in a different way. Looking at Figure 3, it is apparent that there is rising attention towards engagement in SV tools. This highlight just how important engagement in SV is. Finally, Table 6 displays the summary for the top ten of the selected papers from journals or conferences.

Table 5 Distribution of articles based on databases

Database	Initial Search		Selected Papers	
	n	%	n	%
IEEE Xplore	80	18.52	19	32.76
ACM	15	3.47	7	12.07
Springer Link	55	12.73	3	5.17
Science Direct	39	9.03	4	6.90
Wiley Online	19	4.40	3	5.17
Scopus	224	51.85	22	37.93
Total	432	100	58	100

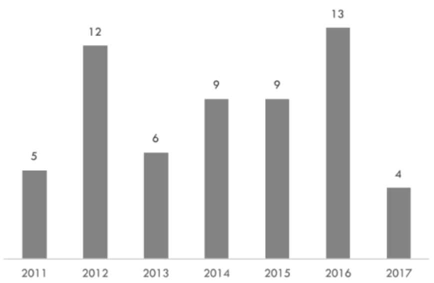

Figure 3: Selected papers by year of publication

Table 6 Distribution of selected articles based on top Journal/Conference

	Journal/Conference proceeding	Type	n	%
1	Proceedings of the ACM Technical Symposium on Computing Science Education	Proceedings	5	8.62
2	Proceedings of the Innovation and Technology in Computer Science Education Conference	Proceedings	4	6.90
3	International Symposium on Computers in Education	Proceedings	3	.17
4	Computer Applications in Engineering Education	Journal	3	.17
5	Proceedings of the Koli Calling International Conference on Computing Education Research & Proceedings	Proceedings	3	5.17
6	Frontiers in Education Conference (FIE)	Proceedings	2	3.45
7	IEEE International Conference on Advanced Learning Technologies	Proceedings	2	3.45
8	IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)	Proceedings	2	3.45
9	Computer Science Education	Journal	2	3.45
10	Computers & Education	Journal	2	3.45
Total			28	48.28

4. Conclusion

This paper was the researchers' first step towards enhancing understanding on how to increase learner engagement in SV tools. In this paper, the SLR protocol was presented for identifying the theoretical foundation used in software visualization. Furthermore, it aimed to present the recent progress in the SLR. This study found 432 articles from the initial search of the literature of which 58 were selected for the final extraction to answer the research questions. The researchers are currently in the process of data extraction. The theory and engagement factors are the expected outcomes of this study.

Acknowledgment

The authors wish to thank the Ministry of Education Malaysia for funding this study under Fundamental Research Grant Scheme (FRGS), S/O project code: 13581.

References

[1] A. Vihavainen, J. Airaksinen, and C. Watson, "A Systematic Review of Approaches for Teaching Introductory Programming and Their Influence on Success," in Proceedings of the Tenth Annual Conference on International Computing Education Research, 2014, pp. 19–26.

[2] B. A. Price, R. M. Baecker, and I. S. Small, "A Principled Taxonomy of Software Visualization," J. Vis. Lang. Comput., vol. 4, no. 3, pp. 211–266, 1993.
[3] J. Hidalgo-Céspedes, G. Marín-Raventós, and V. Lara-Villagrán, "Learning principles in program visualizations: A systematic literature review," in 2016 IEEE Frontiers in Education Conference (FIE), 2016, pp. 1–9.
[4] J. Sorva, V. Karavirta, and L. Malmi, "A Review of Generic Program Visualization Systems for Introductory Programming Education," ACM Trans. Comput. Educ., vol. 13, no. 4, p. 15.1-15.64, 2013.
[5] E. Isohanni and H.-M. Järvinen, "Are Visualization Tools Used in Programming Education?: By Whom, How, Why, and Why Not?," in Proceedings of the 14th Koli Calling International Conference on Computing Education Research, 2014, pp. 35–40.
[6] T. Naps, S. Rodger, Á. Velázquez-Iturbide, G. Rößling, V. Almstrum, W. Dann, R. Fleischer, C. Hundhausen, A. Korhonen, L. Malmi, and M. McNally, "Exploring the role of visualization and engagement in computer science education," SIGCSE Bull., vol. 35, no. 2, p. 131, Jun. 2003.
[7] N. Myller, R. Bednarik, E. Sutinen, and M. Ben-Ari, "Extending the Engagement Taxonomy: Software Visualization and Collaborative Learning," Trans. Comput. Educ., vol. 9, no. 1, p. 7:1-7:27, 2009.
[8] G. Banerjee, S. Murthy, and S. Iyer, "Effect of active learning using program visualization in technology-constrained college classrooms," Res. Pract. Technol. Enhanc. Learn., vol. 10, no. 1, p. 15, Dec. 2015.
[9] L. Malmi, J. Sheard, Simon, R. Bednarik, J. Helminen, P. Kinnunen, A. Korhonen, N. Myller, J. Sorva, A. Taherkhani, J. Sheard, R. Bednarik, J. Helminen, P. Kinnunen, A. Korhonen, N. Myller, J. Sorva, Simon, R. Bednarik, J. Helminen, P. Kinnunen, A. Korhonen, N. Myller, J. Sorva, and A. Taherkhani, "Theoretical Underpinnings of Computing Education Research: What is the Evidence?," in Proceedings of the Tenth Annual Conference on International Computing Education Research, 2014, pp. 27–34.
[10] M. Knobelsdorf, E. Isohanni, and J. Tenenberg, "The Reasons Might Be Different: Why Students and Teachers Do Not Use Visualization Tools," in Proceedings of the 12th Koli Calling International Conference on Computing Education Research, 2012, pp. 1–10.
[11] S. Fincher and M. Petre, Computer science education research. Taylor & Francis, 2004.
[12] C. A. Shaffer, M. L. Cooper, A. J. D. Alon, M. Akbar, M. Stewart, S. Ponce, and S. H. Edwards, "Algorithm Visualization: The State of the Field," ACM Trans. Comput. Educ., vol. 10, no. 3, pp. 1–22, Aug. 2010.
[13] J. E. Hannay, D. I. K. Sjoberg, and T. Dyba, "A Systematic Review of Theory Use in Software Engineering Experiments," IEEE Trans. Softw. Eng., vol. 33, no. 2, pp. 87–107, Feb. 2007.
[14] A. Bikner-Ahsbahs and S. Prediger, "Networking of Theories—An Approach for Exploiting the Diversity of Theoretical Approaches," in Theories of Mathematics Education, Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 483–506.
[15] A. Lishinski, J. Good, P. Sands, and A. Yadav, "Methodological Rigor and Theoretical Foundations of CS Education Research," in Proceedings of the 2016 ACM Conference on International Computing Education Research - ICER '16, 2016, pp. 161–169.
[16] B. Hanks, S. Fitzgerald, R. McCauley, L. Murphy, and C. Zander, "Pair programming in education: a literature review," Comput. Sci. Educ., vol. 21, no. 2, pp. 135–173, 2011.
[17] L. Malmi, J. Sheard, Simon, R. Bednarik, J. Helminen, A. Korhonen, N. Myller, J. Sorva, and A. Taherkhani, "Characterizing Research in Computing Education: A Preliminary Analysis of the Literature," in Proceedings of the Sixth International Workshop on Computing Education Research, 2010, pp. 3–12.
[18] L. Malmi, "Theory --- what is it for?," ACM Inroads, vol. 5, no. 4, pp. 34–35, Dec. 2014.
[19] S. Charters and B. Kitchenham, "Guidelines for performing systematic literature reviews in software engineering," 2007.
[20] S. Gregor, "The Nature of Theory in Information Systems," MIS Q., vol. 30, no. 3, pp. 611–642, 2006.