Mob Programming Challenges, Success Factors, and Practices: A Multivocal Literature Review Protocol with Preliminary Results

Rafi Ullah*, Muhammad Ilyas†

Abstract

Mob programming (MP) is an emerging and relatively new and unexplored programming technique while getting more and more popularity and attention in the software industry. Due to the infancy of this area, there is a significant requirement to identify critical success factors (CSFs), that may be beneficial to be adopted by developers and critical barriers (CBs) that may hinder in the way of MP in software industry. Similarly, there is also a strong need to identify practices that may be beneficial for adopting the CSFs and for addressing the CBs. – The objective of this paper is to design a Multivocal literature review (MLR) protocol to systematically review the literature to identify CSFs, CBs, and their associated practices. Primary objective of the author is to develop MP maturity model (MPMM). Crruntly protocol of this study is in the implementation stage of the MPMM. The anticipated outcome of this protocol is to identify CSFs, CBs, and associated practices in MP with the aim of enhancing, productivity, and quality of software development in the software industry. The identified CSFs lead to develop a MPMM. The proposed model is expected to assist software vendors in addressing the challenges faced by them during software development.

Keywords: Mob Programming; Mobbing; Collaboration; Vendors; Multivocal Literature Review.

Introduction

The software engineering (SE) process consists of a set of activities/phases in which the developer follows a set of guidelines and techniques to develop high-quality software within the stipulated schedule and budget (Hneif et al., 2009). Agile software development is one such approach of them. It is an iterative and flexible approach in which the developer develops software in collaboration. The goal of the agile software development methodology is to produce high-quality software (Balijepally et al., 2017; Ilyas et al., 2016; Rao et al., 2011). Recently a new development paradigm of agile methodologies known as Mob programming (MP) has been introduced in the software industry. MP is a collaborative software development method. MP as described by Woody Zuill (2106), is "the whole team working on the same problem, at the same time, in the same space, and using the same workstation". It extends the idea of pair programming, where two people collaborate to carry out a

*Corresponding Author: Department of CS & IT, University of Malakand, Chakdara 18800, Pakistan, rafiullah.cs@uom.edu.pk

†Department of CS & IT, University of Malakand, Chakdara 18800, Pakistan, milyasmkd@uom.edu.pk

task, while in MP, collaboration is extended to the entire team (Zuill et al., 2016). MP as a software development technique is essential for the software industry as it is a collaborative method and provides help in creating quality software. It provides better results because team members work in collaboration and extend their cooperation (Balijepally et al.,2017; Lilienthal 2017; Zuill et al., 2016).

MP, due to its novel and creative working method, is being adopted more frequently in the software industry. In the software industry, this approach has been adopted in the last decade due to its prevailing benefits, i.e. working in collaboration quality of coding and code design methods have improved, efficient use of tools, increase in team's technical level, and comprehensive knowledge of system. MP also increases satisfaction, confidence, and coordination among team members. Furthermore, Mobbing provides maximum feedback, and communication and increases the learning of developers (Björklund et al., 2020; Hohman & Slocum, 2001; Kattan, 2019).

Due to the infancy of this area, there is a strong need to identify the CBs that may hinder in the way of MP in software industry. The objective of this study is to identify those barriers and their associated practices or solutions for addressing them. Similarly, there is also a need to identify the CSFs and their practices that may be beneficial for adopting MP. The following sections present background and motivation, methodology, preliminary results, and conclusion of the study.

Software development teams are required to follow certain methodologies or approaches to achieve their maximum productivity level. In SE, currently, various software development methodologies, tools, techniques, and best practices are adopted by software developers for application development (Ilyas,et al., 2023; Lucian, 2021). Traditional methods are used in the early age of computer technology to develop software. In these methods, a module or a part of a module is developed by a single programmer (Erickson et al., 2005). The use of agile software development methods has rapidly increased specifically, during the last two decades (dragos & lundberg, 2021). Pair and MP methodologies are examples of agile methodologies. Pair programming is a collaborative method, where two programmers develop software together using a single workstation computer (Kattan et al., 2017).

MP or "Mobbing" term is first used at the beginning of the past two decades (Balijepally et al., 2017). It is a collaborative working method and is first introduced in the literature of software development i.e. in the Extreme Programming (XP) community by Moses Hohman and Slocum in 2001 (Hohman & Slocum, 2001). In the first talk of Woody Zuill at the Open Jam at Agile 2012 held in Dallas, TX, used "Whole Team

Programming" for the MP. They call it Software Teaming (Zuill & Meadows, 2016). Later on, in 2014, the MP term is elaborated on, and popularized in the literature by Woody Zuill and Kevin Meadows, being the pioneers of MP (Woody & Meadows, 2016). The research community is paying more and more attention to MP, it is still an emerging and relatively new and unexplored programming technique in the software developing world (Stahl & Torvald, 2021). The rest of the section gives a brief overview of some recent studies on MP.

A case study is carried out on MP development practice by Kattan et al. (2017), in which various characteristics and practices that are involved in MP are evaluated. It is concluded that productivity increases with workers' interest, satisfaction, enjoyment, coordination among team members, and learning in collaboration. Gareth (2018) conducted a control experiment to investigate the impact on team communication and learning using the MP software development approach. In this study, it is evaluated that, using the MP as a software development method, it facilitates better team communication and learning. It is also concluded according to the data gathered; individuals engaged in MP felt more motivated, secure, and satisfied as well as communicated and learn more efficiently than non-MP.

Systematic Literature Review (SLR) is conducted by Shiraishi et al. (2019) to clearly state the definition of MP and other aspects such as advantages and disadvantages and tools required for MP. The authors reported that MP strengthens the participants' learning style regardless of their background, i.e., coding expertise and experience, which promotes the development of high-quality software. The resultant Code had fewer bugs and is more efficient and maintainable. Similarly, Balijepally et al. (2017) have also performed a review study. They provide an overview of the motivation for adopting MP as a programming method. Furthermore, they explore its benefits and text analysis of MP articles. The authors reported that MP increases learning, and increase quality of code, and developers are highly satisfied with MP Sessions. For the effective use of the MP and to be used by more developers as a programming technique, an empirical validation.

Various studies have already been conducted in the area of MP using different research methodologies, i.e., SLR, thematic analysis, experiments, and empirical studies, to investigate the multiple prospectives (Björklund et al., 2020; Stahl & Torvald, 2021; Shiraishi et al., 2019). These studies only include the state of the art in published literature, while the practitioner's experiences are not included. The aim of this research study is to include the grey literature (GL), which incorporates both formal published literature and GL, to find the success

factors, barriers, and associated practices in MP that have not been investigated by the researchers in this area of interest.

Despite the growing interest in this field, a MLR is lacking to identify the success factors, barriers, and practices for effective usage of MP methodology in software industry. Our research is based on a MLR, followed by an Empirical study carried out in the software industry via a questionnaire survey.

Research Methodology

Our research objective is to develop a MP Maturity Model (MPMM). This study is conducted in four phases. MLR is the first phase of our research work. This research work aims to design MLR protocol for discovering CSFs, CBs, and their Practices in MP methodology from both white literarure and GL. During the development of this MLR protocol, various already published MLR protocols and studies are studied for guidance (Akbar et al., 2021; Matouq et al., 2020; Antil, (2020); Calderón et al., 2018; Garousi, et al., 2019; Lohrasbinasab et al., 2020; Ilyas et al., 2020; Myrbakken & Palacios, 2017; Nylund, 2020; Vale et al., 2021; Scheuner & Leitner, 2020).

Multivocal Literature Review

In the area of SE, Systematic Mapping (SM) and SLR have gained widespread popularity. These types of studies are very useful as they provide an overview of the existing evidence and also identify gaps in the primary studies of a particular field of study. However, these studies only focus on formally published research and lack attention to the considerable amount of GL generated by software developing practitioners (Matouq et al., 2020; Ilyas & Khan, 2016).

A MLR is a type of review that, includes both formal white (academic) literature e.g., journal and conference papers as well as GL which includes unpublished work such as blog postings, audio-video contents, data sets, lectures, and non-peer-reviewed work such as white papers. The MLR is conducted to summarize both the most recent research work and practice literature related to a particular field, it is helpful for both researchers and practitioners (Garousi et al., 2019).

Conducting a review study according to guidelines provided by Kitchenham et al. (2007). This methodology involves three primary stages: planning, conducting, and reporting. Planning is the first phase of MLR that includes protocol development, i.e., research questions (RQs), search string designing, data sources, IC and EC of sources, and quality assessment criteria for the included sources. Conducting phase is the second phase, also known as the conduction phase. It focuses on the

execution of the protocol for a study. It includes primary study selection, data extraction from primary selection, and data synthesis. The last phase is the reporting and its aim is to analyze and present the results. Various phases and their respective steps in conducting this MLR study are explained and represented in Figure 1.

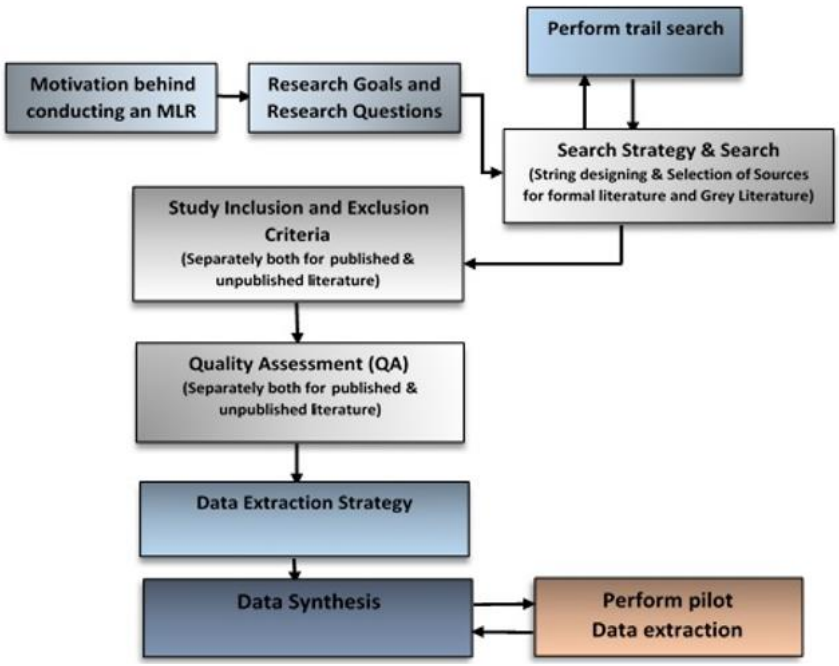

Figure 1: MLR conducting procedure.

Motivation for conducting an MLR

Before conducting a SLR or MLR, researchers need to determine which one's systematic review is necessary. This involves identifying and reviewing the existing literature on the phenomenon of interest (Kitchenham & Charters, 2007). Once the need for a review is established, the next step is to decide whether to conduct SLR or MLR. Various MLR guidelines from different disciplines have tackled the dilemma of whether to incorporate the GL and proceed with an MLR rather than an SLR. MLR can conduct a more inclusive, adaptive, and relevant reviewing of literature in SE as compared to SLR. Due to the addition of GL and multiple sources, MLRs provide a better understanding of current practice and aid in the development of more effective solutions for the same (Garousi et al., 2019).

Nowadays, MLR in the contemporary landscape of computer science, especially in SE, is a useful tool for investigations because new

developments in SE take place. When there is limited academic literature in a particular research area, MLR provides substantial benefits (Garousi, et al., 2019).

GL should be incorporated in a review study particularly when relevant knowledge is inadequately presented in academic publications to validate outcomes via practical experience. It also helps challenge assumptions and allows for the use of academic research (Adams et al., 2017). Additionally, researchers also reported that MLRs help to bridge the gap between academic research and professionals by including GL (Ogawa & Malen, 1991). Garousi et al. (2019), suggested that due to the prevalence of GL as a primary means of sharing knowledge, advice, and experiences among SE practitioners regarding new techniques, approaches, and technology-driven developments, an MLR should be carried out as a preliminary study in SE. If GL is not included, researchers may miss out on crucial updated knowledge and information regarding continuously evolving interesting real-world occurrences. Additionally, the inclusion of GL is beneficial for both researchers and practitioners (Garousi et al., 2019; Mahood et al., 2014).

MLR Protocol Development

MLR in SE is quite young (Garousi et al., 2019). The MLR process is used during this study based on recommendations from Garousi et al. (2019) and Petersen et al. (2008). They provided particular steps that are used to address the GL selection and filtering process. These guidelines are based on Kitchenham's recommendations, which are created for the conduct of SLR studies (Kitchenham & Charters, 2007).

According to Kitchenham and Charters (2007), a protocol must be developed for a systematic study. The protocol describes the procedure adopted for conducting MLR in SE. The protocol is defined and the researchers determine how to work and interact to conduct the review study. A pre-defined protocol is necessary to reduce the potential for researcher bias. The key processes in developing MLR protocol for this study are described in the sections below.

Goals of the MLR

The main aim of this MLR is to uncover the following valuable information about agile MP methodology. The main goals of this MLR are:

- To determine the CSFs, from the literature in MP methodology in SE
- To identify the CBs, from the literature in MP methodology in SE

- To identify practices/solutions, from the literature for the implementation of the CSFs and for addressing the CBs identified in MP methodology in SE

Consequently, this knowledge gained enables us for a focused study to be conducted on a specific aspect of MP that has received relatively limited focus to date. The duration of MLR in this study is limited to publications published and unpublished from 2001 to 2023. The reason for searching literature from 2001 is that, MP as a programming approach is introduced in 2001 by Mosses Hohman and Slocum (Hohman & Slocum, 2001).

Research Questions

Our study is founded on a set of RQs that motivated this research work. The answers to these RQs provide recent information and analyze the current state of knowledge on the topic that is under consideration. In order to achieve the aforementioned goals, three RQs have been set out.

RQ1: What are the CSFs, as identified from the literature, which may be beneficial to be adopted by developers in MP methodology?

RQ2: What are the CBs, as identified from the literature, which may need to be avoided by developers in MP methodology?

RQ3: What are the practices, identified from the literature, for adopting the CSFs identified in RQ1 and for avoiding/addressing the CBs identified in RQ2?

Search Strategy

A robust search strategy is very essential for conducting a successful review study. The first stage in conducting an MLR is to retrieve and select the pertinent sources for a review study. The primary goal of this step is to establish the search and evaluation strategies for categorizing the primary studies in an MLR. The search and assessment techniques are crucial for conducting an exhaustive search for GL that can address the suggested RQs. The search strategy consists of several steps and is used to define how relevant sources are selected. The key steps in the search strategy consist of the following steps.

Construction and Identification of the Search Terms

At the outset, RQs are analyzed from four different perspectives i.e. population, intervention, outcome of relevance, and experimental design in order to derive the search terms by evaluating the keywords in the RQs. Based on keywords in RQs first, search terms are derived and formulated search strings using a framework called PICO recommended by Kitchenham and Charters (2007), and other steps proposed by Brereton

et al. (2007). The mechanism adopted for the construction and identification of the search terms is as follows:

a. First, by determining the population, intervention, and result that are relevant, different search terms are extracted from the study questions.
Population: SE & MP
Intervention: Success factor or CSFs, Barriers or CBs, and Practices in MP
Comparison: For the current study, no comparisons take place
Outcome of Relevance: Our objective is to investigate the success factors and barriers and to improve the practices in MP to increase the efficiency, productivity, and quality of software development in the software industry
Experimental Design: In our study experimental design contains MLR, empirical studies or evaluative studies, case studies, white papers, expert opinions, blog posts, and observations

RQ1:
[What are the CSFs] INTERVENTION
as identified from the literature, which may be
[beneficial to be adopted by developers in] OUTCOMES OF RELEVANCE
[MP Development] POPULATION
RQ2:
[What are the CBs] INTERVENTION
as identified from the literature,
[which may need to be avoided by developers in] OUTCOMES OF RELEVANCE
[MP Development] POPULATION
RQ3:
[What are the Practices/Solutions] INTERVENTION
identified from the literature,
[for adopting CSFs identified in RQ1 and for avoiding CBs identified in RQ2] OUTCOMES OF RELEVANCE
[MP Development] POPULATION

Furthermore, the following search strategy is also incorporated while creating the search keyword.

b. Find alternative words or synonyms for major or key terms
c. The keyword/major terms are verified in some pertinent sources/papers
d. The Boolean "OR" operator is used to concatenate the alternate spelling or synonyms and the Boolean "AND" operator to concatenate the Major/Keyword
e. Incorporate the search string into a concise form if required

The aforementioned plan is adopted to identify the major terms from RQs, and also used for searching purposes in this research study.

Results for (a):

- RQ1. Success factors, MP, Developer
- RQ2. Barriers, MP, Developer
- RQ3. Practices, MP, Developer

Results for (b): Keywords Synonyms used in RQs.

Mob Programming: Mobbing

Note: In some research articles on MP the synonym “mobbing” is also used. However, due to the retrieval of irrelevant results by the search string, I have excluded this synonym from our search string.

Factor: Factor OR "success factors" OR "key factor" OR "critical success factor" OR CSFs OR Benefit.

Barriers: Challenge OR problem OR risk OR issue OR barrier OR trouble OR obstacle OR Failure OR CBs.

Practices: Practice OR exercise OR solution OR advice OR "implementation initiative" OR Standard.

Vendor: Supplier OR Developer OR Seller OR "service-provider".

Note: Due to the potential loss of relevant research articles, the word “vendor” and its synonyms is removed from the search string because numerous papers did not explicitly refer to the term "vendor" or its equivalent synonym. Another reason is that programming is always done on the vendor’s side.

Results for (c): MP, Success Factors, Barriers, Practices, Developer.

Results for (d): Combining various synonyms using Boolean AND and OR operators

Constructing the Search Strings

In our MLR, various synonyms of keywords in RQs are combined by using Boolean operators AND and OR. Finally, searching for various arrangements of these keywords, in this study the subsequent search string is finalized for each question separately.

RQ1: ("Mob Programming") AND (Factor OR "success factors" OR "key factor" OR"critical success factor"OR CSFs OR Benefit).

RQ2: ("Mob Programming") AND (Challenge OR problem OR risk OR issue OR barrier OR trouble OR obstacle OR Failure OR CBs).

RQ3: ("Mob Programming") AND (Practice OR exercise OR solution OR advice OR"implementation initiative" OR Standard).

Search String Breakup

It is important to note that certain databases, such as "Google Scholar" and "CiteSeer", have restrictions on the length of search strings. Consequently, if needed, the string created in the generated in the preceding steps are split into three or more smaller strings. The search string in this MLR study consists of three sub-strings.

Substring 1: ("Mob Programming") AND (Factor OR "success factors" OR "key factor" OR "critical success factor" OR CSFs OR Benefit).

Substring 2: ("Mob Programming") AND (Challenge OR problem OR risk OR issue OR barrier OR trouble OR obstacle OR Failure OR CBs).

Substring 3: ("Mob Programming") AND (Practice OR exercise OR solution OR advice OR "implementation initiative" OR Standard).

Final Search String

The above sub-search strings are combined in one final search string by using the Boolean AND operator.

Final Search String = ("Mob Programming") AND (Factor OR "success factors" OR "key factor" OR "critical success factor" OR CSFs OR Benefit OR Challenge OR risk OR problem OR issue OR barrier OR trouble OR Failure OR CBs OR Practice OR exercise OR solution OR advice OR "implementation initiative" OR Standard).

Trail Search

A preliminary search is conducted using the search string on Google Scholar and other databases. In our trial search, there are some known sets of papers in our study that are extracted from the above-mentioned databases. Then search string is tried in these digital libraries. Trial search is successful as it retrieved already-known papers. The search string is formulated as follow.

("Mob Programming") AND (Factor OR "success factors" OR "key factor" OR "critical success factor" OR CSFs OR Benefit OR Challenge OR risk OR problem OR issue OR barrier OR trouble OR Failure OR CBs OR Practice OR exercise OR solution OR advice OR "implementation initiative" OR Standard).

Sources to be Searched for Literature

Resources that are searched for to locate the related documents in MLR are categorized into White or Published Literature Resources and GL or Unpublished Literature Resources.

Published Literature Resources for White Literature

As per Kitchenham and Charters (2007), to retrieve relevant white literature (WL) in MLR, two search strategies are followed.

Automatic Search: Automatic search is used to locate the maximum number of formally published documents in various digital electronic sources. In this study, the following relevant online digital venues are used to find the relevant literature.

- Google Scholar https://scholar.google.com
- ACM Digital Library https://dl.acm.org
- Springer Link https://www.springerlink.com/
- Science Direct https://www.sciencedirect.com/
- IEEE Xplore https://ieeexplore.ieee.org
- Wiley Online Library https://onlinelibrary.wiley.com/
- CiteSeer Digital Library https://citeseerx.ist.psu.edu
- IET digital library https://digitallibrary.theiet.org

These are among the most well-known and prominent academic electronic databases in SE. The main reason to include the above digital databases in the literature search is the possibility of accessing their contents. Furthermore, these databases offer extensive coverage of the SE field in general and provide access to high-impact full-text journals and conference proceedings (Kitchenham & Charters, 2007).

Snowballing technique: The snowballing technique is used to find literature that may have been missed by automatic search. It finds relevant studies by analyzing the references in the selected papers and authors (Wohlin, 2014). The snowballing (forward & backward) technique is used on all selected articles in accordance with (Kitchenham & Charters, 2007) SLR principles in order to incorporate as many pertinent or relevant sources as possible. Finding articles that reference the pertinent articles that are part of the research is called backward snowballing, while forward snowballing involves finding more relevant articles to include in the study by looking through the reference lists of pertinent publications (Afzal et al., 2009; Wohlin, 2014).

Unpublished Literature Resources for Grey Literature

In accordance with Garousi et al. (2019) MLR criteria, aim of the study is to locate GL sources relevant to research objectives of the study. The following two strategies is used to incorporate the GL in this study.

Automatic Search (Search Engines)

To identify GL, as recommended by Garousi et al. (2019), it is make sure the use of the search terms listed in the section of final search

string above. The most popular search engine is Google and it is used to select the most appropriate GL materials for review studies.

- Google search engine: (www.google.com)

Google search engine is used as recommended by Garousi et al.(2019), to conduct a sophisticated pre-search, for locating and retrieval of all available sources that are relevant to our search keywords to obtain the practitioner's website or sources that publish documents related to MP.

Stopping Criteria

Google search engine provide a lot of links to different sources, but it is necessary to limit the sources to a specific size that can be easily managed. The search query produced a few sites that are relevant to our research study; the remaining pages, however, primarily contain unrelated content. The page rank algorithm is used to assess the relevancy of the pages that are retrieved (Langville & Meyer, 2006). Searching is stopped until a page contains no relevant information about MP.

Digital Databases

Following the guidelines suggested by Garousi et al. (2019), an advanced search is conduct across various online databases.

- ProQuest Dissertations and Thesis Global: A database, where relevant Ph.D./Master thesis may be retrieved
- Opengrey https://www.opengrey.eu/
- Arvix https://www.arvix.org/
- Agilealliance https://www.agilealliance.org/
- Istbq https://www.istbq.org/
- Idc https://www.idc.com/

Furthermore, according to Garousi et al. (2019), "Stack Overflow" is the premier and most reliable online community for developers to enhance their programming skills, exchange knowledge, and advance their careers, which can be explored for GL in SE.

Manual Search (Methods' creator websites)

A review must include practitioners' websites that publish GL, as recommended by Garousi et al. (2019). A substantial category of publications is known as "GL" which comprises various kinds of documents, i.e. web pages, reports and blog posts, technical reports, research projects, and organization annual reports related to the research area and also incorporates government documents such as white papers, evaluation papers, videos, and blog posting. Typically, these materials are not controlled by or produced by commercial publishers. Other relevant

resources for GL are Social websites, online forums, blogs, institutional repositories, government websites, etc.

Managing Search Results

Search results are managed using a special format for primary sources. A separate abbreviations is used for each digital library i.e. SL-1, ACM-1and SD-1 for Springer Link, ACM, and Science Direct respectively, and also a source is assigned a paper ID, which is the number of articles collected from a specific digital library by using a search string and paper title, while the paper title is the title of the primary source article. In case of duplicate articles, it is removed if the same article is collected from other primary sources. The results are stored in a table form as shown in Table 1 below.

Table 1: Managing Search Results

S.No.	Primary Source Name	Paper ID	Title of Paper
1	SPRINGER LINK	SL-1	
2	ACM DIGITAL LIBRARY	ACM-1	
3	IEEE Xplore	IEEE-1	
4	…….	……	

Study Inclusion and Exclusion Criteria

It's really difficult to choose papers for review research from a lot of sources. For this purpose, the developed inclusion criteria (henceforth referred to as IC) and exclusion criteria (henceforth referred to as EC) to select the most relevant studies and exclude pertinent studies. A research's inclusion and EC specify whether a source found through a search must meet to be included in the study or excluded. As recommended by Kitchenham and Charters (2007), the IC and EC are determined by our study questions (Kitchenham & Charters, 2007). Primary sources selection is divided into published and unpublished and also defined separate IC and EC for each source.

Published Sources Inclusion Criteria

Using the search string across several digital libraries, it retrieved many studies. The extracted data from various sources are checked with the IC and EC. Only those sources are included that are related to our RQs. All those sources that discuss, success factors, barriers, and practices in MP are considered for inclusion. The criteria for including sources are outlined below.

1. IC1: Studies in the area of MP in SE are included
2. IC2: Studies, relevant to RQs as defined in section 3.2.1(b)

3. IC3: Studies that are written in the English language
4. IC4: Studies published from the years 2001 to 2023, as MP practice is introduced in the literature of software development in 2001
5. IC5: Peer-reviewed papers i.e. journals and peer-reviewed experienced reports i.e. conference papers, and Ph.D./Master thesis

Published Sources Exclusion Criteria

Sources that fail to meet the IC or are not relevant to the RQs are excluded. The EC, which are explained used below to decide to exclude sources.

1. EC1: Studies that are not written in English
2. EC2: Studies that fall outside the domain of SE
3. EC3: Studies that don't pertain to our RQs
4. EC4: Studies that do not focus on collaborative programming methods
5. EC5: Studies whose Full text is not accessed
6. EC6: Duplicated studies if retrieved several times from other sources are excluded from the final set

Unpublished Sources Inclusion Criteria (IC)

The IC for GL are listed below.

1. IC1: The website provides Ph.D./Master thesis, case studies, and blog posts, etc.
2. IC2: Sources that are relevant to and focus on MP
3. IC3: The source/material pertains to the search terms
4. IC4: The language used in writing the study is English

Unpublished Sources Exclusion Criteria

For exclusion, the following criteria are used for GL.

1. EC1: The study fails to meet the research objective
2. EC2: Studies that do not focus on MP
3. EC3: A study that doesn't clearly state its research approach
4. EC4: Study that is not written in English
5. EC5: Sources that are not fully accessible
6. EC6: Textbooks and duplicate studies

Formal or GL that did not meet the aforementioned IC or met any of the EC is excluded from our analysis.

Conducting a Review Study

Conducting a review study is the execution of the study protocol. Various steps in the conduction phase are explained below (Garousi et al., 2019).

Formal Primary Sources Selection

The first step of primary source selection involves applying the developed search string to all selected digital databases. Once the studies are retrieved from automatic search, the next task is to select primary publication sources. For that purpose, in this study utilizing the "tollgate approach" suggested by Afzal et al. (2009). Phases of this method are as follows.

Pahse-1. Initial Search: Search terms are used to find relevant articles in digital resources.

Phase-2. Remove Duplicate: The title and abstract of the study is comprehensively examined based on study IC or EC criteria to discard the irrelevant studies and focus on identifying only those sources that are most pertinent to our RQs.

Phase-3. First Selection Process: The inclusion and exclusion of relevant primary sources are determined based on the introduction and conclusion.

Phase-4. Second Selection Process: The full text of all selected studies is reviewed and evaluated against the aforementioned IC and EC, resulting in the refinement of source selection to discard irrelevant studies. In case of uncertainty about whether a source, is to be included or not, that is discussed with the secondary reviewer for further evaluation against IC and EC in order to make optimal decisions. A detailed record is kept for each source selection to justify why a primary source is included or excluded from the final evaluation.

After conducting both automated and manual searches, the snowballing technique is performed to identify additional relevant studies that meet our IC and EC. Finally, all the identified studies are consolidated, and any duplicates or twice-collected studies are eliminated.

Unpublished Sources Selection

Unpublished sources i.e. GL are also located by using automatic search and applied to a widely used search engine "Google" for the selection of appropriate sources. GL sources are included or excluded according to the criteria given in the above sections.

Quality Assessment (QA)

The primary sources included in this study are assessed for the purpose of making sure they are valid and unbiased. Quality assessment criteria is a very important activity as this strengthens the study selection. According to Calderón and Education (2015), by evaluating the relevance and informational value of the assessment results, the number of primary studies included in the review can be reduced. When the number of primary studies retrieved is high, this step can be useful. Primary sources

in this study is assessed in accordance with the provided guidelines of Garousi et al.(2019), furthermore, a separately publication quality assessment criteria is set both for published or formal literature and GL which include non-peer-reviewed sources as mentioned in the following section.

Quality Assessment Criteria for Academic/formal Literature

The included formal or WL sources are assessed to ensure their quality and also to verify their validity and impartiality. Quality assessment is a crucial activity that is also carried out to increase the strength of a study. To enhance the quality of our work, a short questionnaire is developed to evaluate the quality of candidate sources. This questionnaire is distributed among senior researchers of the SE group at, the the University of Malakand (SERG_UOM). Furthermore, to assess the quality of the chosen primary formal literature, i.e. WL, the guidelines proposed by Garousi et al.(2019) and the mechanism of Khan et al.(2021) and Alqmase et al.(2022) is adopted to assess the quality of formal literature. The scoring plane of the quality assessment is presented in Table 2, and it consists of columns a, b, c, and d and score. Ultimately, a checklists is introduced to additional enhance the formal literature selection process and evaluate source quality. The quality assessment plan for our research work is presented below.

Table 2: A classification scheme for academic literature

				Quality Assessment				
P-Id	Publication venue	Publication Year	Research Method	A	B	C	D	Score
1	Journal	20__	SLR	1	1	2	1	5
2	Conference	20__	Review	1	1	2	1	5
3	Workshop	20__	Experiment	1	1	1.5	1	4.5
4	----	----	----	--	--	--	--	---
5	----	----	----	--	--	--	--	---

a) Is the study published from a recognized source of publications? a score of 1 for “Yes” and 0 for “No” is assigned

The research is accepted for publication in a prestigious conference or journal, with its ranking determined based on the Computer Science Conference ranking in Computing Research Education CORE (2023), for conference and the Journal Citation Report (JCR) by Clarivate (2023) list of computer science for journal papers. Futhermore, Scimago

Journal & Country Rank (SJR) is also used for ranking journals if a journal is not identified or not ranked by JCR (SCImago, 2023).

b) As per CORE (2023), conferences, workshops, seminars, and symposiums are categorized based on rankings into the following categories.
   - A*-Flagship conference: These conferences are considered premier events in a particular field, a acategory of “Very relevant”, and a score of 3 is assigned
   - A- Excellent conference: Conferences that are excellent and highly respected in a specific area, a category of “Relevant”, and a score of 2 is assigned
   - B-Good conference: Conferences that are good and well regarded in a specific area, a category assigned to it as “Somewhat Relevant” and a score of 1.5
   - C-conferences that come across minimum standards in a specific area and assign it as “Not relevant” and a score of 1
   - Conferences that are not presented in the CORE ranking are assigned a score of 0

Similarly, the quality assessment of journal papers is carried out according to the JCR ranking published by Thomas Reuter Clarivate (2023) JCR provides rankings of journals based on impact Factor (IF). A quartile (Q) is a journal's specific ranking in a database, determined by the journal's impact factor (IF), citations, and indexing. It is divided into four quadrants i.e. quartiles.

   - Q1-denoting the top 25% IF distribution and is assigned the category of “Very relevant” and a score of 4
   - Q2-represents the intermediate and high position between the top 50% to top 25% IF distribution, categorized as "Relevant" with a score of 3
   - Q3-indicates the middle-low position, specifically the top 75% to top 50%, and is categorized as "Somewhat Relevant”
   - Q4- represents the lowest position, specifically the bottom 25% of the IF distribution, categorized as "Not relevant" with a score of 1.5
   - A journal not included in the JCR list is assigned a score of 1

c) Investigating factors that influence the MP approach is the primary goal of this study. It is investigated whether a source contributes any or not and assign a score of 1 for “Yes” and 0 for “No” respectively

d) The source discusses any practices for the success of MP methodology. It is given a score of 1 for “Yes” and 0 for “No” and 0.5 for “partially”

Grey Literature Quality Assessment Criteria

Using Garousi et al.(2019), recommended guidelines, to evaluate the quality of GL sources in the area of SE. A checklist of criteria is used as presented in Table 3 for the selection of GL. The checklist comprises criteria such as Authority of the producer, Methodology, Objectivity, Date, Novelty, Impact, and outlet controls for the selection of GL. The quality assessment criteria suggested by Abrar et al. (2023), is adopted in this study. This method has also been utilized by other researchers to evaluate the quality of GL (Akbar et al., 2021; Antil (2020), Lohrasbinasab et al., 2020; Nylund, 2020; Pereira et al., 2021). There are 21 questions on the checklist, and a score of 1 is given to each. If the total score exceeds 50%, the score quality of the source is considered acceptable; otherwise, it is considered a failure.

Table 3: Criteria for grey literature quality assessment

Criteria	Questions
Significance/ Novelty	C1: Does the source contribute to the current research? Or increase the current stock of knowledge in the area? C2: Is without this research the area be lesser?
Authority of the Producer	C3: Is the reputation of the author and publishing organization reliable e.g. SE Institute? Microsoft Corporation and IBM etc. C4: Is the organization or author frequently cited by others? C5: Is the author knowledgeable and produced any other work in the field?
Courage	C6: Do limitations in the study are clearly stated?
Methodology	C7: Does the source find/discuss any barriers, Success Factors, and Practices related to MP? C8: Did a certain question get addressed in the source? C9: Is an empirical evaluation conducted on the study? C10: Does the findings and results of the study clearly stated? C11: Do the results justify the conclusions drawn?
Objectivity	C12: Does the work is presented in a balanced manner? C13: Identifying bias is crucial, especially if it is not acknowledged or stated C14: Is the information or statements provided in the sources a subjective view or is it as objective as possible? C15: It is still an opinion even if the source expresses the opinion of an expert; is the author's viewpoint expressed enough?
Accuracy	C16: Are the goals and objectives of the source described clearly? C17: Is the methodology of the source stated clearly? C18: Does the source have mention references to support its arguments?
Date of Publication	C19: Does the source clearly mention the publication date? C20: Has the bibliography included key contemporary material?
Publication Type and their Outlet	C21: In SE Adams et al.,(2017), sources of GL are classified into different types i.e., • 1st tier GL (measure = 1): High outlet control/ High credibility: It includes books, magazines, theses, government reports, and white papers • 2nd tier GL (measure = 0.5): Moderate outlet control/ Moderate credibility: it includes annual reports, news articles, presentations, videos, Q/A sites (such as Stack Overflow), and Wiki articles • 3rd tier GL (measure = 0): Low outlet control/ Low credibility: it includes blogs, emails, and tweets

Data Extraction Strategy

During this phase of MLR, selected sources is studied to extract the data. Data is gathered from two primary sources: data gathered during the extraction process and data gathered during the search process. To differentiate publications obtained from the scientific and gray selection processes, generic information is required during the search process.

This information serves in decision-making when selecting the primary studies for review and documenting the selection process. By contrast, specific data is collected during the extraction process to fulfill the goals of the study and provide answers to the questions that are being investigated. Cruzes and Dybå (2011), recommendations are used to carry out the data extraction method. The data extraction process consists of the following steps.

Primary Study Data Extraction

The first set of papers from SLR's initial search, the second set of papers from snowballing, and the final set of URLs from GL are used for the data extraction. The primary studies are a combination of academic or WL and GL. In order to answer the RQs stated in the review, the purpose of this phase is to gather information from pertinent publications. Each publication's full text are studied and data related to the proposed idea are extracted from the primary studies. The extraction process consist of data extraction form which include fileds S.No, title of paper, authors name, reference no, methodology (case study, survey, report, interview etc.), database or Search engines for GL, population, publications quality description or overview, venue of paper, publication year, type of organization, medium type, company level, success factors, challenges/barriers, and practices in the area of MP.

Data Extraction Process

In the data extraction process, two key roles i.e. primary and secondary reviewers are involved. The primary reviewer is to extract the data and approaches the secondary reviewer in case of any issues or guidance needed in the extraction process. An inter-rater reliability test is conducted once the primary reviewer has finished the data extraction process to ensure the reliability of the collected data. The secondary reviewer randomly selects a subset of publications from the sources selected by the primary reviewer, which is then be compared to the primary reviewer's extracted data. The inter-rater reliability test is considered positive if the results are similar on the other hand, the extracted data need to be reviewed by the primary reviewer once more.

Collected Data Storage

A standard data collection form in Microsoft Excel are designed to save the extracted data i.e. success factors, barriers, and practices while a statistical tool SPSS is used for further analysis.

Data Synthesis

Before reporting the outcomes of the study, the last step is the data synthesis and analysis. In data synthesis, data gathered during the extraction phase are examined to compare data, provide answers to the RQs the study addresses, and produce results. The following way is adopted to synthesize the data in our review study. As there are three RQs in this study, the synthesis process is categorized into three parts and are presented in a table form. For Research question (RQ).1 one summary table is created that consists of columns having (S. No, Success Factors, Frequency/Occurrences, and Percentage) and the list show success factors in MP. A separate table documents the comprehensive details of every success factor noted in the summary table. The separate table contains the following columns i.e.CSFs, Reference S. No, Paper reference/Paper title.

The process employed for the Research Question (RQ1), as mentioned earlier, repeates for the next Research Question (RQ2) and Research Question (RQ3).

Preliminary Results

Implementing the MLR protocol and the outcomes listed according to the aforementioned sections of the study protocol. Following the implementation of the search strategy as mentioned above on specified digital libraries, obtained 523 results for formal literature in six digital libraries and about 8360 results for GL. Table 4 provides details about each digital library's primary and final selection. Finally, after applying the IC and EC, 60 and 80 sources are selected for formal and GL respectively.

Table 4: MLR Sources Search Results

S.No	Source Name	Total search results found	Primary Selection	Final Selection
1	Science Direct	93	10	08
2	IEEE Xplore	04	03	03
3	SPRINGER LINK	78	50	07
4	ACM Digital library	22	15	07
5	Google Scholar	318	250	28
6	IET library	08	07	07
7	Google Search Engine	8360	200	80
	Total Results	8883	870	140

Validation of the protocol

The last step in conducting a review is the Validation of the protocol. Initially, the developed protocol is submitted to the secondary reviewer for review and finally, the protocol is presented to the SE Research Group at the University of Malakand (SERG_UOM). The suggestions/feedback are taken into account, leading to the appropriate revisions of the protocol.

Divergences

A new appendix is provide to this document with any deviations from the protocol that occur throughout the research.

Conclusion

Despite the growing interest in the field of MP, a MLR is lacking to identify the success factors, barriers, and practices for effective use of MP methodology. In this paper to begin with, it outlined our plan in the form of an MLR protocol and currently the study in the implementation process of this protocol. In addition, preliminary results are obtained in this study. The anticipated outcome of this protocol is to identify CSFs, CBs, and associated practices in MP with the aim of enhancing, productivity, and quality of software development in the software industry. Furthermore, the identified Success factors lead us to develop a MPMM. The proposed model is expected to assist vendors in addressing the challenges faced by them during software development. Our research is based on a MLR. Lastly, empirical study in the form of a questionnaire survey in the software industry is also conducted.

Acknowledgments

We are thankful to SERG_UOM, for their feedback and comments on this MLR protocol.

Author contribution

Rafi Ullah is a PhD research scholar in the Dept. of CS & IT at the University of Malakand under the supervision of Dr. Muhammad Ilyas assistant professor Dept. of CS & IT at the University of Malakand. The protocol is designed by the primary author and collects all the related data. The secondary author reviews and validates the protocol and gives valuable suggestions. The final results of the paper are reviewed by the authors and approved the manuscript.

References

Abrar, M. F., Khan, M. S., Khan, I., Ali, G., & Shah, S. (2023). Digital Information Credibility: Towards a Set of Guidelines for Quality

Assessment of Grey Literature in Multivocal Literature Review. Applied Sciences, 13(7), 4483.

Adams, R. J., Smart, P., & Huff, A. S. (2017). Shades of grey: guidelines for working with the grey literature in systematic reviews for management and organizational studies. International Journal of Management Reviews, 19(4), 432-454.

Afzal, W., Torkar, R., & Feldt, R. (2009). A systematic review of search-based testing for non-functional system properties. Information and Software Technology, 51(6), 957-976.

Akbar, M. A., Mahmood, S., Meshram, C., Alsanad, A., Gumaei, A., AlQahtani, S. A. J. M. T., & Applications. (2021). Barriers of managing cloud outsource software development projects: a multivocal study. Multimedia Tools and Applications, 1-24.

Al-Matouq, H., Mahmood, S., Alshayeb, M., & Niazi, M. (2020). A Maturity Model for Secure Software Design: A Multivocal Study. IEEE Access, 8, 215758-215776.

Alqmase, M., Alshayeb, M., & Ghouti, L. (2022). Quality assessment framework to rank software projects. Journal of Automated Software Engineering, 29(2), 41.

Antil, P. (2020). Requirements engineering in global scaled agile software development environment: a multi-vocal literature review protocol (SERL_TR_2018_01).

Balijepally, V., Chaudhry, S., & Nerur, S. P. (2017). Mob programming–A promising innovation in the agile toolkit. Association for Information Systems AIS Electronic Library(AMCIS), 1-9.

Barbara Kitchenham, S. M. C. (2007). Guidelines for performing systematic literature reviews in software engineering. Retrieved from University of Durham Durham, UK.

Björklund, P. F., Jacob, & Dalipi, F. (2020). Exploring the Impact of Mob Programming on the Well-Being of the Developers: Insights from a Software Company. Paper presented at the NOKOBIT-Norsk konferanse for organisasjoners bruk av informasjonsteknologi, November 24-25, 2020.

Brereton, P., Kitchenham, B. A., Budgen, D., Turner, M., & Khalil, M. (2007). Lessons from applying the systematic literature review process within the software engineering domain. Journal of systems software 80(4), 571-583.

Calderón, A., Ruiz, M., O'Connor, R. V., & Interfaces. (2018). A multivocal literature review on serious games for software process standards education. Computer Standards & Interface, 57, 36-48.

Calderón, A., Ruiz, M. J. C., & Education. (2015). A systematic literature review on serious games evaluation: An application to software project management. Computers & Education, 87, 396-422.

Clarivate. (2023). Web of Science data. Available from ISI Retrieved 18.11.2023, from ISI https://mjl.clarivate.com/home?PC=MASTER&Error=1

CORE. (2023). ICORE Conference Portal. Retrieved 18.11. 2023, from ACM https://portal.core.edu.au/conf-ranks/.

Cruzes, D. S., & Dyba, T. (2011). Recommended steps for thematic synthesis in software engineering. Paper presented at the 2011 international symposium on empirical software engineering and measurement.

Daniel, S., & Torvald, M. (2021). Mob programming: From avant-garde experimentation to established practice. Journal of Systems and Software, 180(1), 1-12.

Dingsøyr, T., Nerur, S., Balijepally, V., & Moe, N. B. (2012). A decade of agile methodologies: Towards explaining agile software development. The Journal of Systems and Software, 85(6), 1213-1221.

Garousi, V., Felderer, M., Mäntylä, M. V. J. I., & technology, s. (2019). Guidelines for including grey literature and conducting multivocal literature reviews in software engineering, Information and Software Technology, 106, 101-121.

Hneif, M., & Ow, S. H. (2009). Review of agile methodologies in software development. International Journal of Research and Reviews in Applied Sciences,, 1(1), 1-8.

Hohman, M. M., & Slocum, A. C. (2001). Mob Programming and the Transition to XP. Paper presented at the Proceedings First XP Universe Conference.

Ilyas, M., & Khan, S. U. (2016). Practices for software integration success factors in GSD environment. Paper presented at the 2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS).

Ilyas, M., Khan, S. U., Khan, H. U., & Rashid, N. (2023). Software integration model: An assessment tool for global software development vendors. Journal of Software: Evolution Process, e2540.

John Erickson, K. L., Keng Siau. (2005). Agile modeling, agile software development, and extreme programming: the state of research. Journal of Database Management (JDM), 16(4), 88-100. Source: DBLP

Kattan, H. M. (2019). Software development practices patterns: from pair to mob programming. Anais da III Escola Regional de Engenharia de Software, 147-156.

Kattan, H. M., Oliveira, F., Goldman, A., & Yoder, J. W. (2017). Mob programming: the state of the art and three case studies of open source software. Paper presented at the Brazilian Workshop on Agile Methods.

Khan, R. A., Khan, S. U., Khan, H. U., & Ilyas, M. (2021). Systematic mapping study on security approaches in secure software engineering. IEEE Access, 9, 19139-19160.

Langville, A. N., & Meyer, C. D. (2006). Google's PageRank and beyond: The science of search engine rankings. USA Princeton University press.

Lilienthal, C. (2017). From pair programming to mob programming to mob architecting. Paper presented at the Software Quality. Complexity and Challenges of Software Engineering in Emerging Technologies.

Lohrasbinasab, I., Acharya, P. B., & Colomo-Palacios, R. (2020). BizDevOps: a multivocal literature review. Paper presented at the Computational

Science and Its Applications–ICCSA 2020: 20th International Conference, Cagliari, Italy, July 1–4, 2020, Proceedings, Part VI 20.

Lucian, D. (2021). Mob vs Pair : Comparing the two programming practices - a case study. (Independent thesis Basic level (degree of Bachelor) Student thesis), Linköping University, Digitala Vetenskapliga Arkivet. DiVA database.

Mahood, Q., Van Eerd, D., & Irvin, E. (2014). Searching for grey literature for systematic reviews: challenges and benefits. Research synthesis methods, 5(3), 221-234.

Muhammad Ilyas, S. U. K. (2016). An Exploratory Study of Success Factors in Software Integration for Global Software Development Vendors. Proceedings of the Pakistan Academy of Sciences, 53(3), 239-253.

Muhammad Ilyas, S. U. K., Nasir Rashid (2020). Empirical Validation of Software Integration Practices in Global Software Development. SN Computer Science, 1(3), 157-179.

Myrbakken, H., & Colomo-Palacios, R. (2017, 2017). DevSecOps: A Multivocal Literature Review. Paper presented at the Software Process Improvement and Capability Determination, Cham.

Nylund, A. (2020). A multivocal literature review on developer experience. (Master), Aalto University, School of Science Helsinki, Finland. Retrieved from https://aaltodoc.aalto.fi/handle/123456789/43567

Ogawa, R. T., & Malen, B. (1991). Towards rigor in reviews of multivocal literatures: Applying the exploratory case study method. Review of educational research, 61(3), 265-286.

Pereira-Vale, A., Fernandez, E. B., Monge, R., Astudillo, H., Márquez, G., & Security. (2021). Security in microservice-based systems: A multivocal literature review. journal of Computer and Security, 103, 102200.

Petersen, K., Feldt, R., Mujtaba, S., & Mattsson, M. (2008). Systematic mapping studies in software engineering. Paper presented at the 12th International Conference on Evaluation and Assessment in Software Engineering (EASE) 12.

Rao, K. N., Naidu, G. K., & Chakka, P. (2011). A study of the Agile software development methods, applicability and implications in industry. International Journal of Software Engineering its applications, 5(2), 35-45.

Scheuner, J., & Leitner, P. (2020). Function-as-a-Service performance evaluation: A multivocal literature review. Journal of Systems and Software, 170, 110708.

SCImago. (2023). SJR —SCimago journal & Country Rank. Retrieved 18.11.2023,https://www.scimagojr.com/journalrank.php?category=1710

Shiraishi, M., Ishizaki, H., Fukazawa, Y., & Yoder, J. (2019). Mob programming: a systematic literature review. Paper presented at the 2019 IEEE 43rd Annual Computer Software and Applications Conference (COMPSAC).

Stephenson, G. A. (2018). Mob programming: The impact on team communication and learning. Retrieved from Johannesburg, South Africa.

Williams, L., & Kessler, R. (2003). Pair programming illuminated. United States: Addison-Wesley Professional.

Wohlin, C. (2014). Guidelines for snowballing in systematic literature studies and a replication in software engineering. Paper presented at the Proceedings of the 18th international conference on evaluation and assessment in software engineering.

Woody, Z., & Meadows, K. (2016). Mob programming: A whole team approach. Paper presented at the Agile 2014 Conference, Orlando, Florida.

Woody Zuill, K. M. (2016). Mob Programming A Whole Team Approach: Leanpub.