Pontifícia Universidade Católica
do Rio de Janeiro

Danilo Almeida Felipe

On Psychometric Instruments in Software Engineering Research Regarding Personality

Dissertação de Mestrado

Dissertation presented to the Programa de Pós–graduação em Informática of PUC-Rio in partial fulfillment of the requirements for the degree of Mestre em Informática.

Advisor: Prof. Marcos Kalinowski

Rio de Janeiro
June 2022

Pontifícia Universidade Católica do Rio de Janeiro

Danilo Almeida Felipe

On Psychometric Instruments in Software Engineering Research Regarding Personality

Dissertation presented to the Programa de Pós–graduação em Informática of PUC-Rio in partial fulfillment of the requirements for the degree of Mestre em Informática. Approved by the Examination Committee:

Prof. Marcos Kalinowski
Advisor
Departamento de Informática – PUC-Rio

Profª. Simone Diniz Junqueira Barbosa
Departamento de Informática – PUC-Rio

Prof. Jean Carlos Natividade
Departamento de Psicologia – PUC-Rio

Rio de Janeiro, June 1st, 2022

Danilo Almeida Felipe

Danilo graduated in Information Systems from the Federal University of Ceará (2018). His main research interests are: Experimental Software Engineering (focusing in human factors) and Software Quality.

Bibliographic data

Felipe, Danilo Almeida

On Psychometric Instruments in Software Engineering Research Regarding Personality / Danilo Almeida Felipe; advisor: Marcos Kalinowski. – Rio de Janeiro: PUC-Rio, Departamento de Informática, 2022.

66 f: il. color. ; 30 cm

Dissertação (mestrado) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Informática, 2022.

Inclui bibliografia

1. Informática – Teses. 2. Engenharia de Software Comportamental. 3. Personalidade. 4. Mapemento de literatura. I. Kalinowski, Marcos. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Informática. III. Título.

CDD: 004

To my parents, friends, and colleagues
for their support and encouragement.

Acknowledgments

I want to thank the Department of Informatics at PUC-Rio for the opportunity to conduct high-level research and CAPES for its sponsorship. Without them, it would be much more difficult.

I would like my gratitude to my advisor Prof. Dr. Marcos Kalinowski for the shared knowledge, the guidance during these years as his master's student, and not least for always being kind and patient.

Besides my advisor, I would like the rest of my master thesis committee: Prof. Dr. Simone Diniz Junqueira Barbosa and Prof. Dr. Jean Carlos Natividade, for agreeing to participate and contribute to improve this study since its initial design.

Continuing, I am immensely grateful for all the people present in my life these past few years. There are so many people and events... Undoubtedly, facing the adult life together with a worldwide pandemic was scary. Having these people with me made things easier.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

Abstract

Felipe, Danilo Almeida; Kalinowski, Marcos (Advisor). **On Psychometric Instruments in Software Engineering Research Regarding Personality**. Rio de Janeiro, 2022. 66p. Dissertação de Mestrado – Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

Context: Although software development is an inherently human activity, research in Software Engineering (SE) has focused mostly on processes and tools, thus failing to recall the human factors behind it. Even when explored, researchers typically do not properly use the psychological background to understand better human factors in SE, such as the psychometric instruments, which aim to measure human factors.
Objective: Our goal is to present an overview and reflections on psychometric instruments in SE research regarding personality.
Method: We conducted a systematic mapping of the literature to generate a catalog of the psychometric instruments used.
Results: This dissertation contributes with an update of an existing secondary study to cover fifty years of SE research (1970 to 2020). We observed remaining discrepancies between one of the most popular adoption instruments (MBTI) and existing recommendations in the literature on the use of this instrument.
Conclusion: The findings lead us to conclude that the adoption of psychometric instruments regarding personality in SE needs to be improved. Future work directs us to analyze the mapped literature under the lens of social science specialists and researchers.

Keywords

Behavioral Software Engineering; Personality; Mapping study.

Resumo

Felipe, Danilo Almeida; Kalinowski, Marcos. **Instrumentos Psicométricos na Pesquisa em Engenharia de Software Sobre Personalidade**. Rio de Janeiro, 2022. 66p. Dissertação de Mestrado – Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

Contexto: Embora o desenvolvimento de software seja uma atividade humana, a pesquisa em Engenharia de Software (ES) concentra-se principalmente em processos e ferramentas, esquecendo-se dos fatores humanos por trás. Ainda quando explorados, os pesquisadores não tem adotado adequadamente referencial da psicologia para entender melhor os fatores humanos em ES, bem como dos instrumentos psicométricos, que visam medir algum tipo de fator humano.
Objetivo: Nosso objetivo é apresentar uma visão geral e reflexões sobre o uso dos instrumentos psicométricos na pesquisa da ES em relação a personalidade.
Método: Foi realizado um mapeamento sistemático da literatura para gerar um catálogo dos instrumentos psicométricos utilizados.
Resultados: Esta dissertação contribui com a atualização de um estudo secundário existente para cobrir cinquenta anos de pesquisa em ES (de 1970 a 2020). Observamos discrepâncias remanescentes entre um dos instrumentos mais popularmente adotado (MBTI) e as recomendações existentes na literatura sobre o seu uso.
Conclusão: A adoção de instrumentos psicométricos relativos a personalidade em ES precisa ser aprimorada. Trabalhos futuros nos direcionam a analisar a literatura mapeada sob a ótica de especialistas e pesquisadores das ciências sociais.

Palavras-chave

Engenharia de Software Comportamental; Personalidade; Mapemento de literatura.

Table of contents

List of figures

List of tables

I will be more than my survival,
Own these scars on my heart

Evanescence, *Part of Me.*

1
Introduction

In this chapter, we provide the context of the work presented in this dissertation, along with our motivation to tackle the problem herein described. Next, we present the research goal and methodology followed in order to answer our research question, followed by the contributions of the work.

1.1
Context and Motivation

Software Engineering (SE) activities are primarily performed by humans. However, many empirical studies have only focused on proposing new methods and technologies to support SE activities leaving human and social factors behind them underexplored (FELDT et al., 2008), impeding a more holistic view of the area.

Behavioral Software Engineering (BSE), proposed by Lenberg et al. (2015), is the body of knowledge of SE research that attempts to understand human aspects related to the activities of software engineers, software developers, and other stakeholders. The topic has been the subject of recent research in the SE domain. Nevertheless, because it is relatively immature, some approaches adopted misled researchers mainly by not properly combining SE research with social sciences backgrounds, such as psychology, to address human factors (GRAZIOTIN et al., 2015b). Graziotin et al. (2018) enlighten a set of existing research on relating developer happiness to productivity, software quality, and social interactions.

Furthermore, in SE research, it is inherent to note that measurement activities are an essential part. In empirical studies the researcher(s) must be sure when adopting ways of measuring the study's resources in question. These resources can be personnel, hardware, or software for an activity or process (WOHLIN et al., 2012)

BSE research has encouraged the use of psychometric instruments as support to the understanding of human factors in a more systematic way (FELDT et al., 2008; LENBERG et al., 2015). In its turn, the Psychoempirical Software Engineering proposed in Graziotin et al. (2015b) deals with "denoting research in Software Engineering with proper theory and measurements from psychology". A problem addressed by the authors is the misuse of theoretical backgrounds of psychology, such as assuming a certain theory as the only truth in the research foundation; and also the improper use of psychometric instruments, some not validated from psychology or used to evaluate wrong human factors.

However, information about how psychometric instruments are adopted in SE research remains vague and dispersed in many studies. As far as we know, one study has partially synthesized this knowledge concerning a specific construct, personality[1]: a period of forty years (1970 to 2010) about personality in SE research is mapped by Cruz et al. (2015). Still, there is only a brief discussion and characterization of the instruments and their use in the software engineering context (*e.g.*, education and pair programming), missing a critical assessment. This status quo remains unchanged more than ten years later and deserves to be challenged.

There is also a need to find out whether SE research over the years has been adopting these instruments coherently with well-known recommendations. Although some authors have outlined this on a smaller scale (MCDONALD; EDWARDS, 2007; USMAN; MINHAS, 2019), a large-scale study has not been done to get a big picture.

1.2 Research Goal

In order to synthesize this knowledge in a structured manner, **our objective is to present an overview and reflections on the use of psychometric instruments in SE research on personality**[2].

We intend to consolidate findings on the use of psychometric instruments in SE research in a catalog. Therefore, we updated a systematic mapping study (secondary study) on personality-related SE research. It is noteworthy that our study focuses on the use of psychometric instruments. More specifically, we classify and discuss the objective of the studies, reported limitations on their use, SE constructs related to the psychometric instruments, the type of research, and empirical evaluations. Additionally, we discuss aspects of the use of the most used instrument under the lens of literature guidelines.

1.3 Contributions

The main contribution of this dissertation concerns updating a systematic mapping study on personality-related SE research to cover fifty years of research, with a particular focus on psychometric instruments. We carefully

[1]Personality is one of the most studied concepts in BSE research as pointed by Lenberg et al. (2015)

[2]Hereafter we refer to psychometric instruments related to personality simply as "psychometric instrument(s)".

assessed the need for an update (MENDES et al., 2020) and followed guidelines on the search strategy to update systematic literature studies (WOHLIN et al., 2020). The previous mapping study covered forty years of research and identified 90 research papers (CRUZ et al., 2015). Applying the search strategy led us to identify 106 additional papers published within the ten subsequent years (2011 to 2020). More specific contributions related to the psychometric instruments include:

– Observing remaining discrepancies between the application of the psychometric instruments within recent SE research and existing recommendations in the literature. We also identified the most common objectives of studies employing these instruments and reported limitations.

– Relating the use of psychometric instruments within recent SE research to theoretical SE constructs, aiming at providing a better understanding on how such instruments are used within the context of *actors* applying *technologies/interventions* to perform *activities* on *software systems.*

– Summarizing the type of research and the empirical evaluations in SE research employing psychometric instruments.

1.4 Organization

The remainder of this dissertation is organized as follows.

Chapter 2 provides a basic background on Behavioral Software Engineering, common personality models, and secondary studies which treat personality in SE, focusing on psychometric instruments.

Chapter 3 presents in detail the systematic mapping protocol, derived research questions, update strategy for an existing mapping study, and documented execution of the protocol.

Chapter 4 presents the results from executed protocol organized by defined research questions and added by a review based on guidelines existing in the literature and identified threats to validity.

Finally, Chapter 5 presents the concluding remarks, limitations, and future work of this dissertation.

2
Background and Related Work

This chapter introduces the theoretical foundation and related work to this dissertation. To approach the theoretical foundation, we describe the broader context of this dissertation (Behavioral Software Engineering), provide a definition of personality, and describe frequently used personality models in SE. Regarding related work, we focus on personality studies in SE, emphasizing existing secondary studies.

2.1
Behavioral Software Engineering (BSE)

BSE is defined as "the study of cognitive, behavioral and social aspects of software engineering performed by individuals, groups or organizations" (LENBERG et al., 2015). It involves dealing with existing relationships between SE and disciplines from social sciences, such as *work and organizational psychology*, the *psychology of programming*, and *behavioral economics* to get a broader understanding of SE practices.

Although software is developed by humans, for a long-time, SE research has focused intensively on the technical aspects (such as processes and tools) and less on human and social aspects (FELDT et al., 2008). A summarization conducted in Graziotin et al. (2017) points out existing research, in the scope of BSE, on relating developer happiness to productivity, software quality, and also social interactions.

BSE defines several constructs called BSE concepts. When operationalized in empirical research, they could provide insights to researchers and practitioners. Also, with adequate knowledge adopted from other disciplines, such as social sciences, it is possible to better understand the software engineer's practices, as a human, in the execution of their activities. It is worth mentioning that BSE is restricted to software engineers and aggregates, and not human aspects related to the use of the software (LENBERG et al., 2015).

Still in the context of BSE, the authors present a definition for the body of knowledge research described earlier and also conducted a systematic literature review based on the definition. The findings report lack of research in some SE knowledge areas (*e.g.*, requirements, design, and maintenance) and rare collaboration between SE and social science researchers. A list of 55 BSE concepts and respective units of analysis is raised and detailed. Concepts such as *cognitive style*, *job satisfaction*, *communication*, and ***personality*** are seen as more frequently studied, while concepts such as *intentions to leave* are underexplored.

2.2 Personality: Adopted Definition and Common Models

In the present study, we focused on mapping the literature on BSE with a restricted scope in personality, observed as one of the most studied concepts and the one with the most significant relationship with other concepts (LENBERG et al., 2015). Despite being a human factor with different definitions, we need to adopt a consistent one with the consolidated literature in SE. A deeper review and discussion are not our scope and goal. However, we need to adopt definitions to support our decisions. For personality, we rely on the following definition used in Cruz et al. (2015):

> Personality is generally viewed as a dynamic organization, inside the person, of psychophysical systems that create the person's characteristic patterns of behavior, thoughts, and feelings. Ryckman (2012) defined personality as "the dynamic and organized set of characteristics possessed by a person that uniquely influences his or her cognitions, motivations, and behaviors in various situations". We use these definitions because they are general enough to allow the inclusion of studies covering a wide range of personality theories and research methods. [...] The dispositional perspective encompasses the traits and types theory, which is one of the most used theories in organizational psychology (ANDERSON et al., 2001) and in studies on personality in software engineering.

Complementarily, Barroso et al. (2017) summarize personality models commonly used to identify personality traits in SE in dispositional perspective. Those models are the **Myers-Briggs Type Indicator**, the **Big Five Model**, and the **Five Factor Model**.

The Myers-Briggs Type Indicator (MBTI) is a model based on Jung's theory of personality types adapted by Isabel Myers and Katharine Briggs in a personality inventory, with the purpose of identifying dominant individual preferences over four dichotomous dimensions (MYERS, 1998):

- *I-E dimension*: refers to the way that an individual directs their energy towards the world. Introversion (I) directs to the inner world of experiences, ideas, and internal experiences (imaginative world). Extraversion (E) to the outer world of people and objects (the real world).

- *S-N dimension*: refers to functions or the way of an individual's perception work. Sensing (S) people tends to rely on what can be perceived by the five senses; on the other side, iNtuitive (N) people rely on patterns and relationships.
- *T-F dimension*: refers to the processes of judging and make conclusions. Thinking (T) people tend to perform logical and objective impartial analysis. Feeling (F) people highlight personal or social values, in a harmonic way.
- *J-P dimension*: is an extended dimension based on Jung's theory, this refers to the way that an individual prefers to deal with the outside world. Judging (J) people prefer to be decisive using the judging processes (T-F dimension). Perceiving (P) people prefer to be more spontaneous using perception processes (S-N dimension).

Given the dimensions described earlier, an individual can be placed in one of 16 possible combinations of personality (INTP, ESFJ, ISFJ, and so on).

Following, we have the Big Five Model and the Five Factor Model. Often the mentioned models are treated in the literature as being the same. However, Barroso et al. (2017) points out a distinction in theoretical basis, causality, and measurement between the two models, distinguishing the Five Factor Model as a derivation of the Big Five Model, in which the latter assumes that personality traits are important for social interaction. At the same time, the former is a model that provides causes and contexts. In Table 2.1, we can compare the descriptions of the models according to Barroso et al. (2017) synthesis.

Table 2.1: Description of the Big Five and Five Factor Models

Big Five Model		Five Factor Model	
Insurgency	Refers to the orientation of an individual in relation to others. Individuals with insurgency traits tend to be talkative, bold, assertive and sociable.	**Extraversion**	Refers to a person involved with the outside world. Extroverts feel comfortable in social relations, are enthusiastic, friendly and active.
Agreeableness	Refers to the sympathy and social interaction of a person. They are nice guys, get along well with others, they are reliable and useful.	**Agreeableness**	Refers to cooperation ability of an individual.
Conscientiousness	Refers to the organization. Conscientious individuals are suitable for hard working; they are organized and able to complete tasks in the proposed time.	**Conscientiousness**	Refers to how individuals manage, regulate and direct their impulses.
Neuroticism	Refers to stress, anxiety, fear, and the volatility of a person. Individuals with this trait tend to not let emotion interfere with their work.	**Neuroticism**	Refers to how an individual experiences negative feelings. Those who have low neuroticism are emotionally stable, calm, confident and secure.
Openness to experience	Refers to imagination, curiosity and wit of an individual. Individuals with this trait tend to be curious, open-minded and arts connoisseur	**Openness to experience**	Refers to an individual's imaginative and creative traits.

To achieve our goal, we treat these two last models together in our analysis of the psychometric instruments, given their similarity. We emphasize that a deeper review of personality models is beyond our scope.

Given the definition of personality and an overview of common personality models, in the next section we highlight secondary studies on the dispositional perspective of personality. Some of them were captured by our secondary study protocol described in Chapter 3.

2.3
Personality and Psychometrics in Software Engineering

According to Michell (1999), "psychometrics is concerned with theory and techniques for quantitative measurement in psychology and social sciences" (Michell, 1999 *apud* Feldt et al., 2008). In addition, Feldt et al. (2008) state that "[...] in practice, this often means the measurement of knowledge, abilities, attitudes, emotions, **personality**, and motivation". The use of psychometric instruments in SE is encouraged, especially in empirical research, as a way to emphasize hitherto unexplored human factors and to help understand how they affect the research landscape (FELDT et al., 2008). This is our focus in this dissertation, given the importance of measurement activities in empirical research.

In Barroso et al. (2017), personality models utilized in 21 papers are mapped onto three main ones: Myers-Briggs Type Indicator (MBTI), Big Five Personality Dimensions, and Five-Factor Model (a variation of the Big-Five Model). The study covers the period of 2003 to 2016 and includes peer-reviewed publications in the IEEE, ACM, and Elsevier digital libraries. In addition to the personality models used, inconclusive findings are also identified on the influence of software engineers' personalities on professional activities. However, there is no mapped information about the psychometric instruments that operationalize these models.

As far as we know, McDonald & Edwards (2007) is the first study that brings to attention on the use of psychometric instruments in SE research, beyond providing guidelines for the use of two of them (MBTI and 16PF). In addition, is highlighted that one of the authors is from social sciences and a certified professional regarding these instruments.

Cruz et al. (2015) performed a systematic mapping on personality in SE research using the dispositional perspective. In addition to reporting on the most common SE topics addressed, such as *education* and *extreme programming*, they reported which personality tests (a.k.a psychometric instruments) were most commonly used, which resembles this study. However, the authors only reported brief information on personality-related psychometric instruments without deep discussion about them regarding their use in SE. They provide a valuable list of instruments and relate them to some SE topics, but lack a critical review involving the use of these instruments.

In its turn, Usman & Minhas (2019) investigate ethical topics raised by McDonald & Edwards (2007) on the adoption of MBTI-based tests in a sample of 8 studies obtained in the final set compiled by Cruz et al. (2015) published after 2007, and complemented with 7 studies returned in string-based search on Scopus[1] in the years of 2016 and 2017, totaling a sample of 15 studies. Their results indicate that the use of psychometric instruments in SE is inadequate. The authors found problems in all of the analyzed studies, including the reliability and validity of MBTI (there are different versions of this instrument). The authors also highlight possible causes, such as not exploring literature guidelines and lack of collaboration with social science researchers. However, the study reported is initial and limited to analyzing only the use of MBTI in a small sample of studies.

Still, Graziotin et al. (2015a) claims that the use of psychometric instruments should be cautious, in addition to the proper theoretical background used. The authors then propose the *Psychoempirical Software Engineering* that aims "to denote research in SE with proper theory and measurement from psychology". In the same study, the authors provide broader steps when adopting psychometric instruments in SE research and exemplify scenarios using the *affect* construct. Nonetheless, the steps initially do not cover personality.

2.4 Concluding Remarks

In this chapter we presented BSE and related work on personality in SE. Although we are researchers without formal qualifications in social sciences, we rely on consolidated methodological tools of SE to present an overview on the use of psychometric instruments in SE research. Our reflections are narrowed to SE literature guidelines and limitations reported within the analyzed studies. The next chapter presents our systematic mapping study protocol.

[1]https://www.scopus.com/

3
Systematic Mapping Protocol

Systematic mapping is a method to build a classification scheme of an area providing a visual summary of the state of research in a structured way (PETERSEN et al., 2008). It aims at providing an auditable and replicable process with minimal bias.

This chapter describes each step of our research method based on guidelines in the literature. Section 3.1 introduces the mapping goal and research questions. Section 3.2 describes the search strategy for collecting new evidence. Section 3.3 presents the study selection criteria and discusses quality assessment, and Section 3.4 presents the Data Extraction Form and our classification scheme. Concluding, Section 3.5 documents how the mapping protocol was started.

3.1
Mapping Goal and Research Questions

Our systematic mapping aims at providing an overview of the use of psychometric instruments in SE research. To guide our investigation, and to obtain an overview of the state-of-the-art, trends and gaps, we describe the main Research Question (RQ) as follows:

RQ1: Which psychometric instruments have been applied in SE research regarding personality?

Based on the main RQ, we derived five secondary RQs in order to further characterize the field as follows.

- **RQ1a:** What are the objectives of the studies?
- **RQ1b:** What are the limitations faced by the use of psychometric instruments reported in the studies?
- **RQ1c:** To which SE constructs are those psychometric instruments related?
- **RQ1d:** Which types of research do the studies reference?
- **RQ1e:** Which types of empirical studies have been conducted?

In the following section, the search strategy is presented. It was developed by this dissertation author and reviewed by the advisor.

3.2 Search Strategy

3.2.1 Existing Mapping Study and the Need to Update

This mapping started in a traditional way of conducting secondary studies (string-based search in digital libraries with snowballing steps), according to consolidated literature (PETERSEN et al., 2008; PETERSEN et al., 2015; KITCHENHAM, 2007; MOURAO et al., 2017). Later, new guidelines emerged, and we noticed that they could help conduct our study (MENDES et al., 2020; WOHLIN et al., 2020), given the awareness we had about comprehensive secondary studies on human factors in SE that could be updated (CRUZ et al., 2015; LENBERG et al., 2015).

We defined Cruz et al. (2015) as a candidate for the update as we wanted to start our immersion into BSE using a narrower scope, focused on personality, to allow a comprehensive overview and a focused critical assessment. Cruz et al. (2015) identified 90 studies published within a time range of forty years, whereas Lenberg et al. (2015) defined 55 BSE concepts (*e.g.*: *personality*, *job satisfaction*, *communication*, etc.) in a large scale study that considered 250 papers. The narrower focus by Cruz et al. (2015) would also allow us to apply our update search strategy (discussed in Section 3.2.2) with reasonable efforts.

We also argue that personality is a BSE concept presented in Lenberg et al. (2015) as one of the most studied together with others (such as *group composition*, *communication*, and *organizational culture*). We believe that updating Cruz et al. (2015) yields significant results regarding our objective described in Chapter 1, which is also within BSE's scope. Thus, we decided to update the mapping study by Cruz et al. (2015).

To evaluate the need for updating the mapping study of Cruz et al. (2015), we used the 3PDF framework recommended by guidelines to evaluate the possibility of updating secondary studies in SE (MENDES et al., 2020)[1]. We conducted the evaluation process by answering the same seven RQs used in Mendes et al. (2020), listed and answered hereafter (Steps 1.a. to 3.b.) and ilustrated in Figure 3.1.

Step 1.a.: *Does the published SLR still address a current question?* Yes. Human factors, such as personality, have been the subject of SE research despite being historically poorly explored (GRAZIOTIN et al., 2018). A list of arguments follows: recent efforts to consolidate a body of knowledge in SE as

[1]This study proposes guidelines for updating Systematic Literature Reviews (SLR). Despite this, we believe that our mapping goal is comprehensive enough to apply it in the context of systematic mappings.

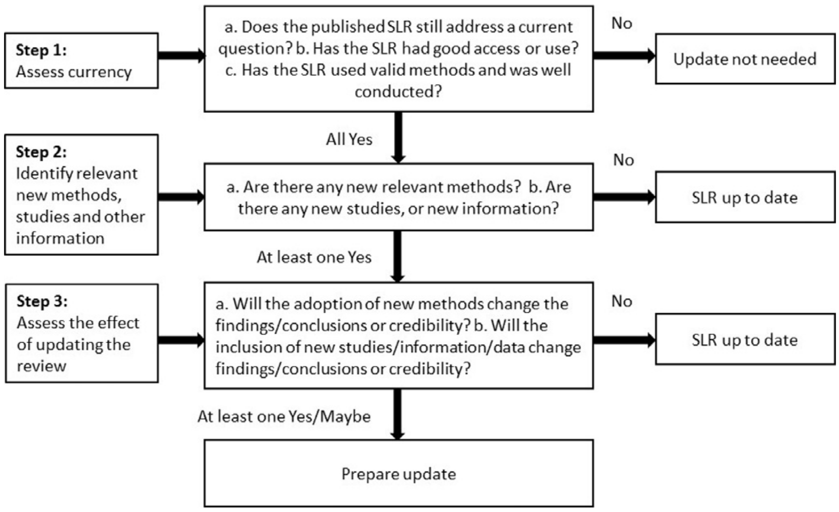

Figure 3.1: Framework recommended by Mendes et al. (2020), adapted from Garner et al. (2016)

proposed by Lenberg et al. (2015) defining BSE; existing conferences such as CHASE (Workshop on Cooperative and Human Aspects of Software Engineering), subsidized by ICSE (International Conference on Software Engineering)[2], the largest conference in SE; award-winning papers on the topic in main software engineering related venues, such as (GRAZIOTIN et al., 2018), awarded in the Journal of Systems and Software[3]. All these arguments show the importance of human factors as a relevant research topic in SE.

Step 1.b.: *Has the SLR had good access or use?*

Yes. Like Mendes et al. (2020), we used in the cut-off point the same yearly average citation value of 6.82 documented by Garousi & Fernandes (2016) to consider a paper for good access or use. In August of 2020, Cruz et al. (2015) had a yearly average citation value of 30.8 in Google Scholar.

Step 1.c.: *Has the SLR used valid methods and was it well-conducted?*

Yes. Regarding the methods, Cruz et al. (2015) present an extension of preliminary results published previously in Cruz et al. (2011) with improvements, such as a refined search string increasing the sensitivity and coverage; adding backward snowballing steps; review of RQs and extended presentation of results. Finally, the authors present clear steps in their mapping protocol and are based on well-recognized guidelines for conduct secondary studies in SE (KITCHENHAM, 2007).

Step 2.a.: *Are there any new relevant methods?*

Yes. Concerning new methods about our mapping protocol, we adopted guidelines presented as the best way to search for evidence to update secondary

[2]http://www.icse-conferences.org/
[3]https://www.journals.elsevier.com/journal-of-systems-and-software

studies in SE (MENDES et al., 2020). We believe that we used good literature references to help answer our RQs and, consequently, reflect on the results' presentation. However, different from Cruz et al. (2015) mapping study, our focus is totally on the psychometric instruments within SE research, not a characterization of an aspect (personality) in general.

Step 2.b.: *Are there any new studies or new information?*

Yes. The papers included in the original study had each a considerable number of citations in a preliminary verification in Google Scholar. In addition to having a five-year time interval since the publication of the mapping and the started conduction of the update in this present study (2015 to 2020), beyond the period of ten years (2011 to 2020), not incorporated by Cruz et al. (2015).

Step 3.a.: *Will the adoption of new methods change the findings, conclusions or credibility?*

Yes, potentially. We adopted a new method concerning the mapping protocol and addressed different RQs to get a big picture of psychometric instruments in SE research, which we believe generates new and important findings. Cruz et al. (2015) has a relevant RQ on the psychometric instruments/personality tests used, but there is little discussion of its results beyond listing and counting frequencies.

Step 3.b.: *Will the inclusion of new studies/information/data change findings, conclusions or credibility?*

Yes. Regarding new potential findings, we had prior knowledge of a series of studies (GRAZIOTIN et al., 2015b; GRAZIOTIN et al., 2017; GRAZIOTIN et al., 2018) used as control papers as a strategy to ensure good literature coverage in our protocol. These studies are not covered by Cruz et al. (2015) because they were published later. They discuss the use of psychometric instruments in SE, and on theoretical basis of other areas (such as social sciences and psychology), which can support SE research in general.

Next, we describe the search strategy to collect new evidence from a secondary study update in SE.

3.2.2 Strategy to Collect New Evidence

We adopted the guidelines proposed in Wohlin et al. (2020) as a strategy to search for new evidence from a secondary study update. They are the following:

- **Use a seed set containing the original secondary study and its included primary studies:** Cruz et al. (2015) included 90 papers in their final set. However, one of them was excluded (S86) because it is a book chapter, and we did not find evidence of publication in a scientific journal or conference to be approved in IC1 (see Table 3.1). As suggested, the secondary study itself was included, obtaining a seed set of 90 studies.

– **Use Google Scholar to search for papers and apply Forward Snowballing (FS), without iteration:** We used the Publish or Perish 7 tool (HARZING, 2020) to assist this step. The tool has features related to bibliometric analysis, in which one is to retrieve citations from publications using Google Scholar (a FS feature). Thus, we conducted the FS in the seed set using the tool in August 2020 and exported the results for treatment in JabRef[4], a bibliographic reference manager. Also, a new FS step was conducted in January 2021 to ensure full 2020 indexation. All screening steps were conducted using JabRef.

– **Include more than one researcher in the initial screening to minimize the risk of removing studies that should be included (false negatives):** One researcher was included to assist in the initial screening of studies and discussions were held with a third researcher.

3.3 Study Selection

Petersen et al. (2015) argue that only studies that are relevant to answer the RQs must be considered. Our inclusion criteria consists of **primary studies published in journals, conferences, and workshops reporting SE research using psychometric instruments regarding personality that were published after 2010**[5]. The exclusion criteria applied to filter the raw set of studies from FS are presented in Table 3.1.

This mapping study aims to provide an overview of the use of psychometric instruments in SE research published in peer-reviewed venues. Therefore, we focus on classifying the type of contribution by discovering objectives, the use of psychometric instruments, and the type of research to understand the overall publication landscape without applying a formal quality assessment. The procedure involved reading titles and abstracts and looking for evidence of psychometric instruments. If it was not enough for clarification, the paper's introduction and the conclusion were read. Still, if not sufficient, the full text of the study was read.

As shown in Table 3.1, we only included papers written in English (EC1), peer-reviewed (EC2), and complete (EC3). Due to the volume of papers, it was necessary to adopt an objective and impartial filter strategy based on

[4]https://www.jabref.org/
[5]Cruz et al. (2015) already had mapped 1970 to 2010.

Table 3.1: Exclusion criteria

Criteria	Description
EC1	Papers that are not written in English.
EC2	Grey literature. Such as books, theses (bachelor's degree, MSc or PhD), technical reports, occasional papers, and manuscripts without peer-review evidence.
EC3	Papers that are only available in the form of abstracts, posters, short versions, and presentations. First, we check whether the paper's information is a short version according to the venue. If not available, we excluded papers with less than six pages.
EC4	Papers that did not include in their title or abstract terms defined by Cruz et al. (2015) as regarding personality. There are: "personality", "psychological typology", "psychological types", "temperament type", and "traits".
EC5	Papers addressing uniquely other psychometric constructs (*e.g.*, behavior, cognition, abilities, roles, etc.)
EC6	Papers that do not meet the inclusion criteria, *i.e.*, papers that do not contribute to SE.
EC7	Papers that use secondary personality data (available datasets, reused data collected in previous studies, etc).

keywords in the title and abstract (EC4). Focusing on the scope of the study itself, papers that did not deal with personality or related concepts (EC5) in SE studies (EC6) were also eliminated. Papers that used simulated or secondary personality data without applying a psychometric instrument (EC7) were also not considered.

3.4 Data Extraction and Classification Scheme

The data extracted from each paper of the final set is shown in Table 3.2.

3.5 Applying the Systematic Mapping Protocol

The first step to execute the mapping protocol was to conduct FS in the seed set as described in the Section 3.2.2, which generated an entire of 6702 entries (step 1 of Figure 3.2). Between September and October of 2020 we conducted an initial screening of duplicates and of studies with year less than or equal to 2010, given that Cruz et al. (2015) cover a range from 1970 to 2010.

Many entries were provided by Google Scholar/Publish or Perish 7 export feature with incomplete or incorrect data (e.g., journal studies categorized in the entry as books or miscellaneous, and truncated title or abstracts). After removing duplicate entries and when possible, the data for each publicated study were complemented and registered in JabRef to perform a more reliable

Table 3.2: Data Extraction Form

Information	Description
Study Metadata	Paper title, author's information, venue, psychometric instrument (name, version, and application process), and year of publication.
Objective (RQ1a)	Study objective: we employed open coding (STOL et al., 2016) to extract information.
Limitations (RQ1b)	Limitations on the use of psychometric instruments (if exists), such as what were the difficulties of adoption/application and data interpretation. We employed open coding (STOL et al., 2016) to extract data.
Purpose of the psychometric instrument in the study (RQ1c)	What constructs represent the purpose of the psychometric instrument in the study. In SE, constructs are derived from one of the classes: *people*, *organizations*, *technologies*, *activities*, or *software systems* (SJøBERG et al., 2008). We employed open coding (STOL et al., 2016) to extract data.
Research Type (RQ1d)	For research type facets we used the taxonomy proposed by Wieringa et al. (2005), containing the following categories: *evaluation research*, *solution proposal*, *validation research*, *philosophical paper*, *opinion paper*, or *experience paper*. Petersen et al. (2015) recommendations were followed in this categorization.
Empirical Evaluation (RQ1e)	Classification of the empirical study in the following categories of Wohlin et al. (2012): *experiment/quasi-experiment*, *case study*, or *survey*.

exclusion per year. The result of this initial screening resulted in 2974 entries (step 2 of Figure 3.2).

Thereafter, another screening was conducted regarding the exclusion criteria EC1 and EC2. The removal was performed based on the metadata provided in the title, abstract, and journal/booktitle field entries. When it was not possible to easily identify, a verification was made through the URL of the entry or by searching the source on the internet. This exclusion was conducted between October and November of 2020 (step 3 of Figure 3.2). Each entry was analyzed individually; these exclusion steps reduced the set to 1718 entries.

Thereafter we removed entries from 2020 and applied EC4, which resulted in 369 entries of 2011 to 2019, step 4 of Figure 3.2). The removal was conducted in order to assure a full ten-year index coverage by replacing them with a new FS conducted in January 2021 to cover the entire year of 2020. In this new FS we considered only papers from 2020 and applied the same previous ECs, which resulted in a candidate set of 403 entries (step 5 of Figure 3.2).

In 2021, we applied the other exclusion criteria (EC3, EC5, EC6, and EC7) by reading the remaining studies in the candidate set and extracting data from the selected ones (step 6 of Figure 3.2). Two additional researchers assisted in this step covering two years each (2017 to 2018 and 2019 to 2020), using a prepared web form with detailed advice for data extraction. All exclusions/inclusions and extracted data were carefully reviewed. As a result, 106 studies were included and had their data extracted. The list of papers and the extracted data can be found packaged online in Felipe & Kalinowski (2022).

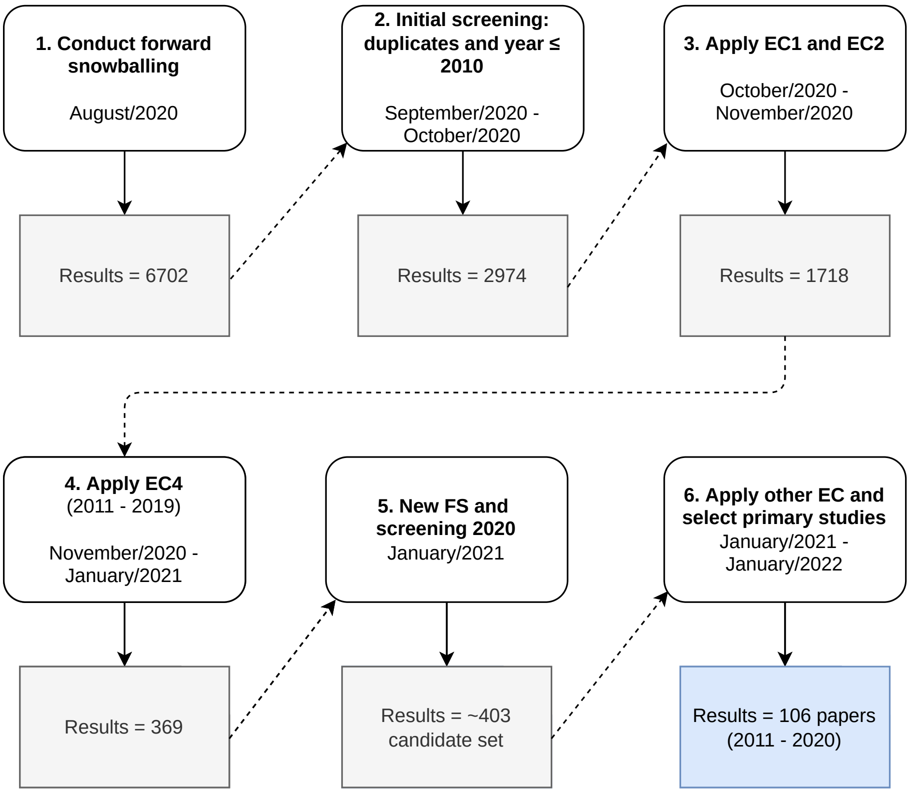

Figure 3.2: Steps of mapping execution

In the next chapter we present the results extracted and organized by our RQs. We mapped the years from 2011 to 2020. All the steps mentioned above involved the guidance and agreement of the advisor. Bibliographic references of the selected studies are presented in Appendix A.

3.6 Concluding Remarks

In this chapter we detailed the systematic mapping protocol and its application. The next chapter presents the systematic mapping study results.

4
Systematic Mapping Study Results

This chapter presents the results of the mapping study. First, we provide an overview of the included studies, followed by sections answering the defined RQs based on the information extracted from the included studies (Sections 4.1 to 4.7). We complement this chapter by discussing aspects of the employment of the most used psychometric instrument in SE research based on existing guidelines (Section 4.8), followed by threats to the validity of the mapping (Section 4.9).

4.1
Overview

We identified 106 additional primary studies that employ psychometric instruments in SE research, ranging from 2011 to 2020[1]. The temporal distribution of the studies is depicted in Figure 4.1. We can observe that the time range of 2014 to 2016 holds the highest frequency of studies. Indeed, when screening and extracting data process, this period required more effort due to also having a larger volume of papers to be analyzed. The main related work in this dissertation has also been published in the same period (CRUZ et al., 2015; LENBERG et al., 2015).

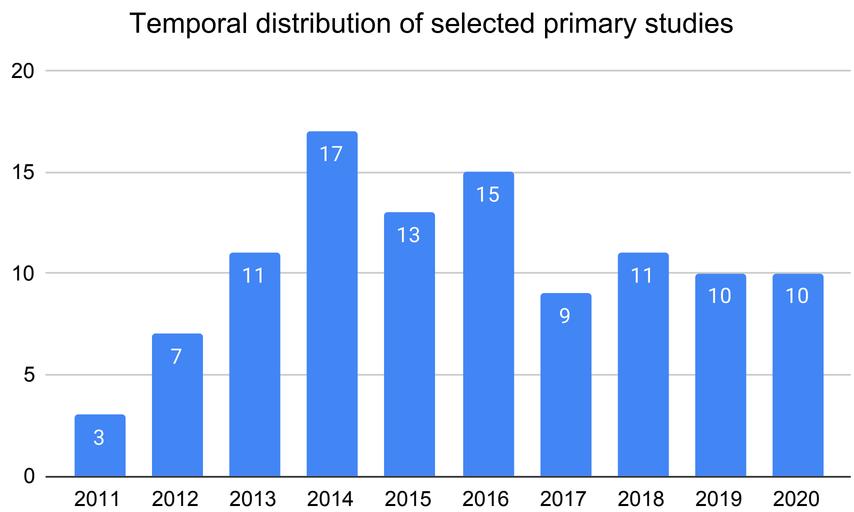

Figure 4.1: Temporal distribution of selected primary studies

[1]We refer to the identification of studies from our protocol as S(NUMBER), where (NUMBER) begins to account from 91, given that Cruz et al. (2015) has 90 mapped studies.

Nevertheless, it is possible to observe that the topic is still consistently being researched. In fact, while we found 106 studies in our investigated ten-year range (2011 to 2020), Cruz et al. (2015) found 90 studies in the previous 40-year range (1970 to 2010).

Figure 4.2 shows that most of the studies (62 out of 106) were published in journals, which reinforces our effort in data extraction due to fewer restrictions on document size generally required. Followed by conference (39 out of 106) and workshop (5 out of 106) publications. The latter could have a low number due to more restrictions in document size, hence not approved by our ECs, as we were looking for complete research papers.

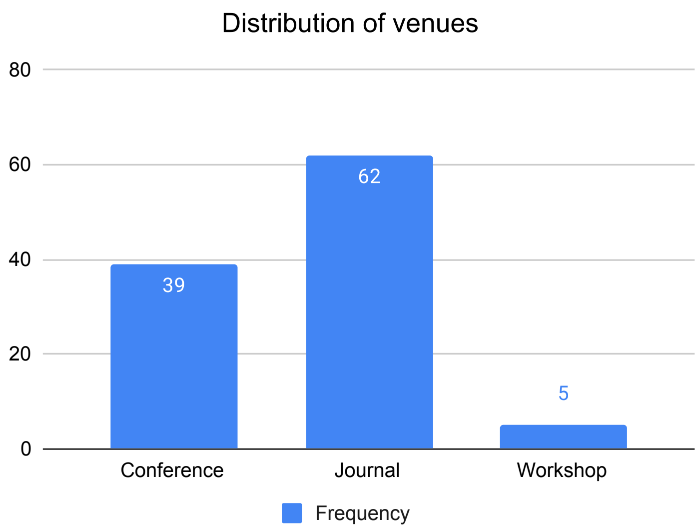

Figure 4.2: Distribution of venues

4.2 RQ1. Which psychometric instruments have been applied in SE research regarding personality?

The frequency of instruments is illustrated in Figure 4.3. The most used ones were versions of the Myers-Briggs Type Indicator (MBTI), a finding also reported in Cruz et al. (2015). For illustration purposes, instruments used in only one study are not shown in the chart. An overview of the complete list can be found in Table 4.1.

To facilitate the instruments' categorization, we compared details of the versions reported through bibliographic references and specific information available to understand if they were referring to the same instrument. In these cases, we kept the details on the version in our repository but consolidated them for analysis purposes (*e.g.*, MBTI, mapped across multiple versions).

It is also noteworthy that a variety of instruments (*e.g.*, IPIP, BFI, and NEO-FFI) that operationalize the Big Five/Five-Factor Models of personality are used. Together with the MBTI, these models stand out as the main theoretical backgrounds for instruments applied in SE research, as pointed out by Barroso et al. (2017). Although the instruments based on the Big Five/Five Factor Model are the majority in sum, we consider each one of the instruments shown in Table 4.1 as an individual instrument in our analysis.

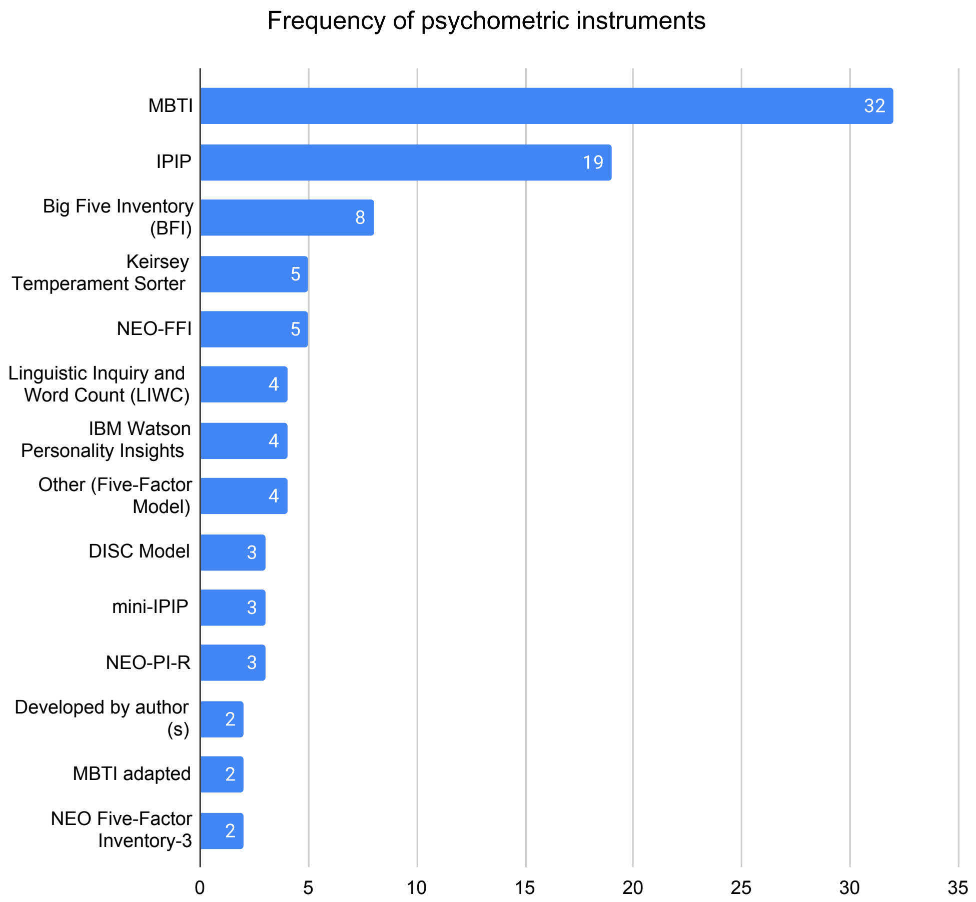

Figure 4.3: Frequency of used psychometric instruments

4.3 RQ1a. What are the objectives of the studies?

During the data extraction, it was possible to observe the following major objectives with open and axial coding procedures (STOL et al., 2016).

Investigate the effect of personality: data on personality is used as an intervention to investigate phenomena. We identified this objective with several specificities illustrated, as shown in Figure 4.4. Studies such as [S92, S96, S103, S106, S113, S115, S193, S196] investigate the effect of personality in pair pro-

Table 4.1: Psychometric instruments mapped organized by studies

Pyschometric instrument	Studies	Count
MBTI	S91, S93, S99, S101, S110, S111, S112, S119, S120, S122, S123, S126, S127, S129, S130, S137, S138, S151, S155, S158, S162, S163, S165, S167, S168, S169, S172, S173, S177, S179, S181, S196	32
IPIP	S92, S106, S113, S124, S128, S132, S135, S136, S139, S146, S150, S161, S164, S171, S174, S178, S182, S186, S195	19
Big Five Inventory (BFI)	S100, S104, S114, S125, S141, S175, S183, S189	8
Keirsey Temperament Sorter (KTS)	S96, S103, S108, S131, S147	5
NEO-FFI	S117, S133, S140, S153, S174	5
Linguistic Inquiry and Word Count (LIWC)	S116, S134, S149, S194	4
IBM Watson Personality Insights	S152, S176, S180, S188	4
Other (Five-Factor Model)	S107, S143, S145, S192	4
DISC Model	S109, S148, S157	3
mini-IPIP	S118, S142, S156	3
NEO-PI-R	S95, S102, S110	3
Developed by author(s)	S121, S184	2
MBTI adapted	S94, S154	2
NEO Five-Factor Inventory-3	S97, S98	2
NEO-PI3	S105	1
IPIP based	S115	1
Five-Factor Stress	S144	1
Other (Unspecified)	S159	1
Student Styles Questionnaire	S166	1
EPQ-R	S170	1
IPI	S185	1
16 Personality Factors (16PF)	S187	1
HEXACO Model	S191	1
Quick Big Five (QBF)	S193	1

gramming teams compositions; in the quality of software developed [S95, S123, S133, S140, S164, S173, S177, S182] and perceived satisfaction [S95, S133]; regarding influence on project management activities [S100, S130], including communication [S134] and collaboration [S177]; in academic contexts, such as analyzing achievements [S102, S120, S122, S125, S169], resilience [S141] or learning outcomes [S144, S175] of students; in development preferences [S104, S110, S114]; on the influence on project success [S108, S117, S157]; in performance of software teams [S136, S153, S160, S165, S169, S172, S187]; finally, in activities on distributed software development [S149, S152, S180].

Furthermore, the effect of personality has also been investigated in other topics, such as the use of CASE [S101, S170] and static analysis tools [S143], implementing a new technology [S109], performance of software engineer [S127, S146], programming styles [S135, S150], software testing performance [S105], team climate and productivity [S139, S195], requirements engineering activities [S142], knowledge management activities [S159, S162], software engineer burnout tendencies [S179], collaboration [S176], creativity [S189], and on the use of software repositories [S194].

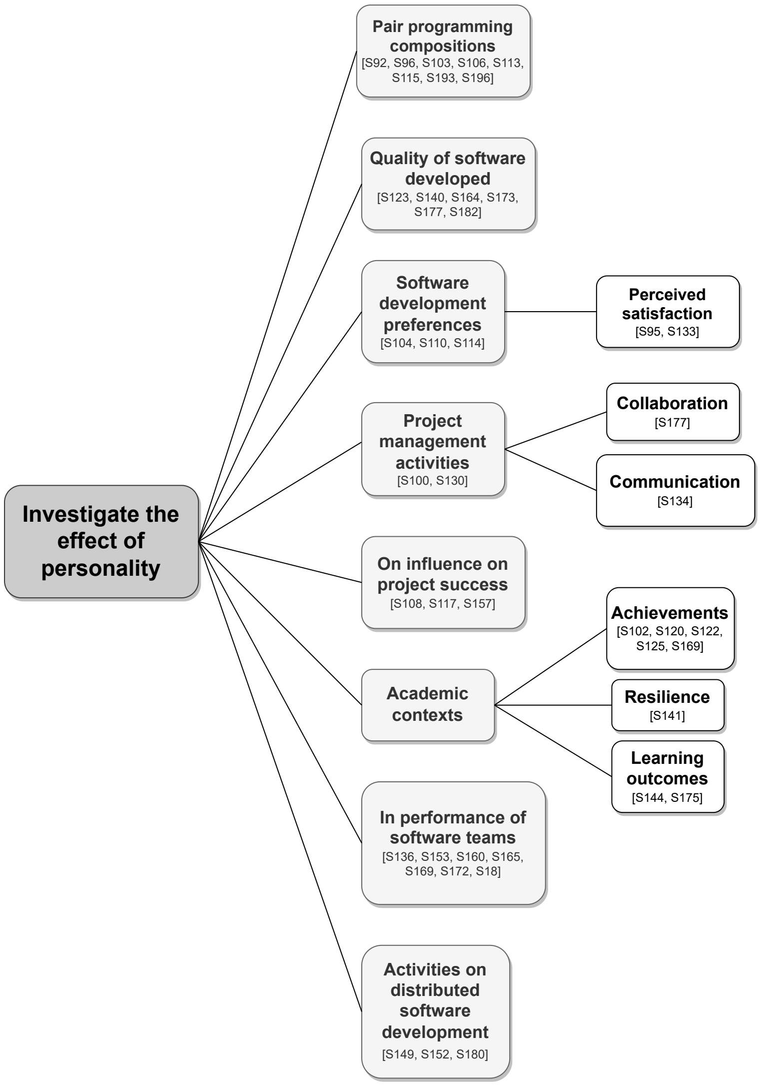

Figure 4.4: Tree structure of *Investigate the effect of personality* major code

Characterize software engineer personality: as shown in Figure 4.5, some studies aimed to discover patterns in personality of software engineering professionals [S93, S119] and more specifically to mapping these patterns to roles and required skills [S99, S107, S111, S129, S132, S138, S158, S161, S179, S184, S185, S188, S190, S191, S192] or preferences in software aspects [S112, S126, S166]. Characteristics of software engineering personality are also explored in distributed software development [S116]; considering common profiles of personality by region or organization-context [S131, S137]; to relate them with other psychometric constructs over time [S118]; and with respect to distinguishing profiles within different computer major courses [S147].

Predict performance or preferences: four studies used information on personality to perform predictions (see Figure 4.6). This objective involves development team's performance algorithms and tools in order to optimize resources [S91, S97, S98] and predict preferred roles of software engineering

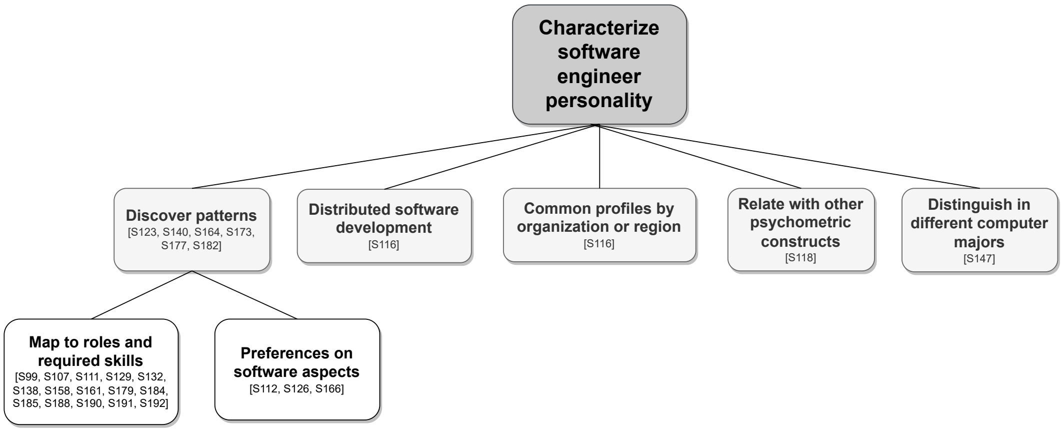

Figure 4.5: Tree structure of *characterize software engineer personality* major code

professionals [S181].

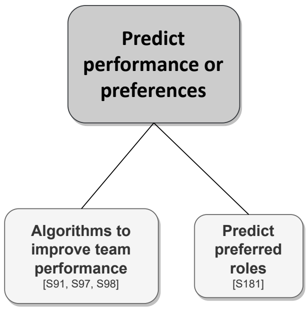

Figure 4.6: Tree structure of *predict performance or preferences* major code

Propose an approach: in this objective, some proposal is given to support research or practical activities in SE (Figure 4.7). For instance, proposing approaches to measure personality [S121, S148, S154, S175, S186] and to interpret statistically data in domain of SE [S156]; to personalize SE activities [S124]; to create links between human factors (like personality) with software diversity [S128]; to investigate turnover intentions [S145]; to relate individual characteristics to development performance [S183]; to perform team compositions [S151, S173]; and to conduct task assignments [S167, S168].

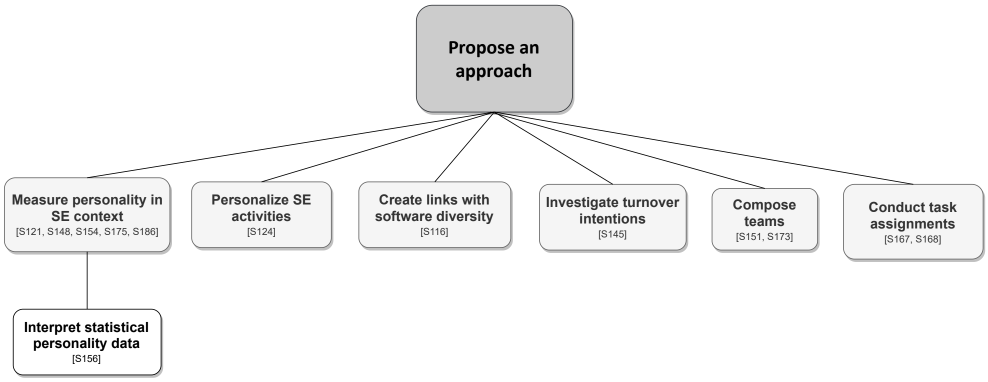

Figure 4.7: Tree structure of *propose an approach* major code

Investigate the relation of personality: this objective is related to check whether personality has any relationship with a research object (see the structure in Figure 4.8). Personality relationships are investigated with learning outcomes [S144], burnout tendency [S178], choices in major degree in computing [S155], class grades considering gender [S163], learning effectiveness [S174], project success [S157], and task selections [S162].

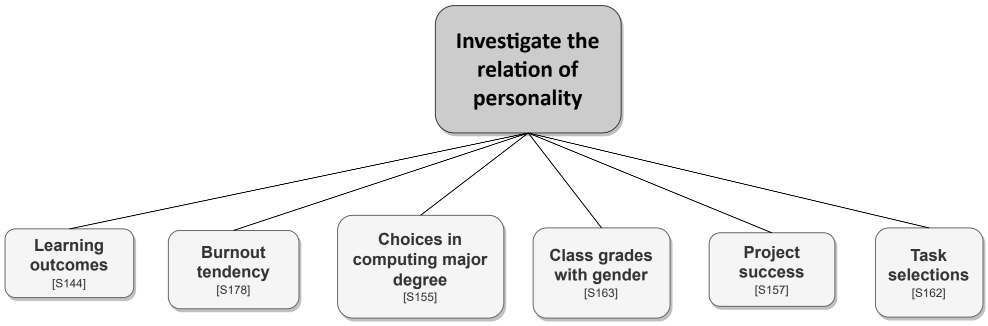

Figure 4.8: Tree structure of *investigate the relationship of personality* major code

We can compare our mapped objectives with the most frequent research topics mapped by Cruz et al. (2015). Listing them (highlighted): ***pair programming***, in *investigate the effect of personality*; ***education***, during the data extraction, we identified studies that contributed to educational purposes or used an academic scenario to reach the research goal in all major codes described earlier; ***team effectiveness***, in *predicting team performance*; ***software process allocation***, ***software engineer personality characteristics***, ***individual performance***, and ***team process*** in all of our major codes. It can show us that these topics are still being researched.

4.4
RQ1b. What are the limitations faced by the use of psychometric instruments reported in the studies?

Less than half of the primary studies (43 out of 106) reported limitations related to the adoption of psychometric instruments in SE research. An overview of these coded limitations (STOL et al., 2016) is described in the following.

Possible misuse of psychometric instrument: the authors of [S95, S132, S152, S157, S187] declare that adopting the psychometric instrument to the objectives of research could affect the validity of the study, but this limitation is mitigated by relying on the literature. Other threats concern not involving psychologists in the research design [S130, S148, S187], use of non-intuitive platforms, and poor instructions on instruments' application [S178]. The lack of a data set for performing benchmarks is also reported [S180].

Choice of short version of psychometric instrument: the statistical power of personality data in the studies may be compromised by the adoption of shorter versions of the instruments, so they could result in less accurate results [S118, S156, S191], hence compromising the research goal.

Personality may not be a representative construct: the use of personality as investigated human factor may not be a good choice for the research design [S128] and encountered correlations may not assure causality [S135, S150].

Bias on subjects responses: the authors indicate that subjects' administration of psychometric instruments can become a threat in cases of factors like lack of honesty or loss of concentration of the subjects [S96, S97, S110, S113, S143, S146, S149, S164, S172, S177, S178, S184, S190, S195] sometimes caused by the absence of control of researcher employing certain empirical approaches (like surveys). This limitation is generally mitigated by ensuring they made aware that the response data obtained is anonymized and used only for research purposes.

Statistical power of psychometric instrument: the dichotomous approach of MBTI based instruments could affect results in scenarios when a person is in a center of a scale due to instruments' statistical structure [S96, S103, S182]. Others report that personality traits data should be measured by adopting other psychometric instruments in order to achieve better results [S100, S104, S111, S158, S163, S173] but without suggesting any other instrument; in case of identification of personality traits from textual analysis, the amount of chunk text may not be sufficient [S176]. Moreover, the one scale/trait of the psychometric instrument showed low internal validity, being excluded from the study [S170].

Paid subjects: participants were paid to participate in the study, which may have influenced them somehow [S105].

Issues with dictionary to measure personality from text: adequacy of language dictionaries to measure personality from textual analysis may be a threat. Also, the data extracted to measure personality may not be representative of a person [S116, S134, S194].

Construction issues on proposed psychometric instrument: the adequacy of psychometric instrument to SE context may deal with validity issues. These issues were mitigate using expert judgments and a series of incremental refinements [S160].

4.5 RQ1c. To which SE constructs are those psychometric instruments related?

Regarding SE constructs, we used the framework to describe theories (SJøBERG et al., 2008) and its archetypal classes as support for open coding (STOL et al., 2016) constructs. This framework is largely used in SE research to present theories. In it, the archetypal classes interact together in the following way: an *actor* applies a *technology* (we believe that *intervention* is more appropriate in the context of psychometric instruments) to perform certain *activities* on a *software system*. In Tables B.1 to B.4 (Appendix B) we list and describe the coded SE constructs to which the psychometric instruments are related within the identified studies, organized by archetypal class.

Figure 4.9 depicts the relationship network between the constructs described in the aforementioned tables. We can observe that constructs of class **Actor** *researcher* and *academic setting* are the majority, indicating the application of studies in an academic setting due to scope limitations of the research design or for purely investigative purposes by researchers. In the **Intervention** class, *personality traits data* represents the most common coded construct and is typically measured by some psychometric instrument and used to investigate its influence on SE activities.

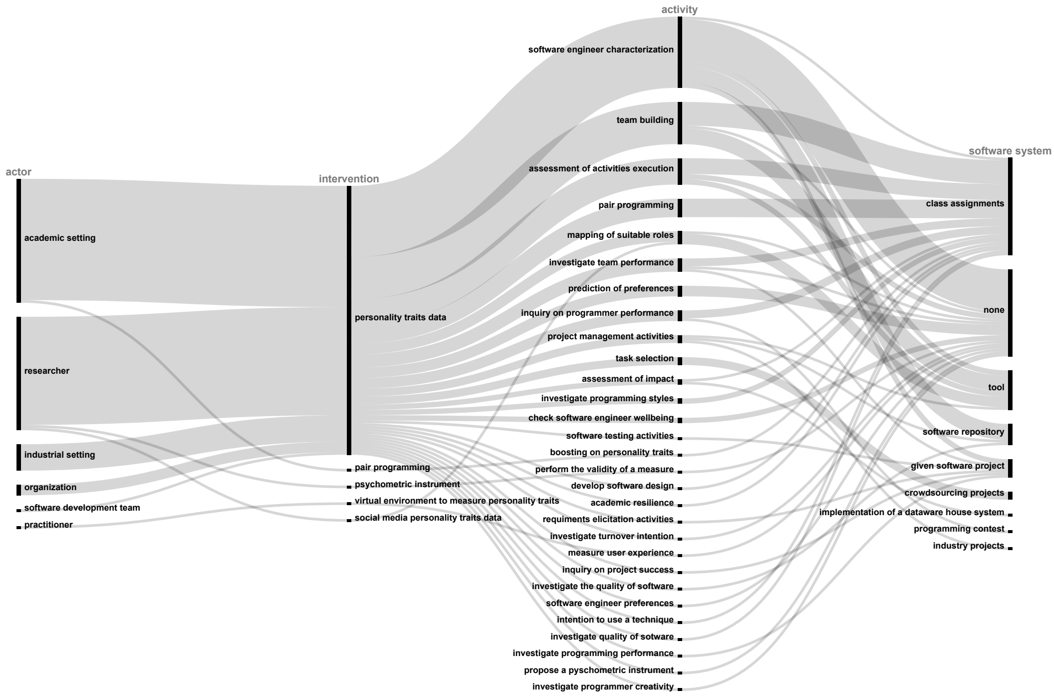

Figure 4.9: SE constructs related to the psychometric instruments

Regarding the **Activity** class, the constructs *software engineer characterization*, *team building*, *assessment of activities execution*, *pair programming*, and *mapping of suitable roles* have a higher frequency. All of them are related to the previously mentioned constructs in classes **Actor** (*academic setting* and *researcher*) and **Intervention** (*personality traits data*). These connections between **Actor** and **Activity** are shown more clearly in the matrix representation in Figure C.2 available in the Appendix C.

Still, the activities described earlier are also strongly related to the constructs of the **Software System** class. It is possible to observe that *software engineer characterization* is typically not related to any specific software system (code *none*), indicating no direct reference or use to software systems in these studies. The code *tool* indicates the use of some technique using software/algorithms/games/logic rules to support the studies. Moreover, *class assignments* were specially related to *team building* and *pair programming*, where SE teams were built based in academic contexts based on personality data. It is noteworthy to mention the use of a *software repository*, in which software artifacts of different kinds are stored.

Please note that RQ1c aims at answering what parts of SE theory the psychometric instruments are related to. There may be similarities with the overall objectives of the mapped studies (RQ1a), but RQ1c is specifically focused on the context of used instruments in the primary studies and their relations to SE theory elements.

For a better understanding of the relationships between the coded constructs, we invite the reader to check the matrices relating the frequencies of the coded constructs two by two, provided in Figures C.1 to C.6 (Appendix C). These figures allow observing, for instance, the relationship between the constructs of the classes **Actor** and **Software System**: *academic setting* is largely related to *class assignments* (what is an expected connection) and *researcher* with *none* software, which denotes more investigations in SE research without using software artifacts.

4.6
RQ1d. Which types of research do the studies reference?

Figure 4.10 depicts the distribution of research types facets by year. It is possible to note that validation research over-represented the set of mapped studies either in distribution per year and in total (85 out of 106). This facet includes empirical studies, as well as the less frequent evaluation research (11 out of 106). The difference between them indicates that most empirical research has been conducted in academic scenarios for initial validation purposes and does not propose and measure new proposals in industrial scenarios. This is somehow expected due to possible difficulties in using industrial practitioner subjects as part of research designs.

Few solution proposals have also been mapped (10 out of 106), which typically represent new research proposals with some limited evaluation required for publication or no presentation of empirical evaluation, therefore they were not classified as either evaluation or validation research.

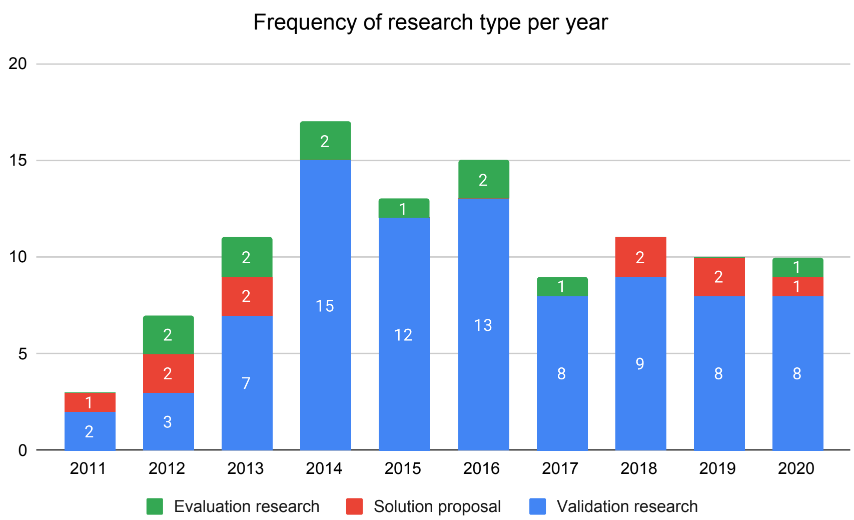

Figure 4.10: Frequency of research type per year

4.7
RQ1e. Which types of empirical studies have been conducted?

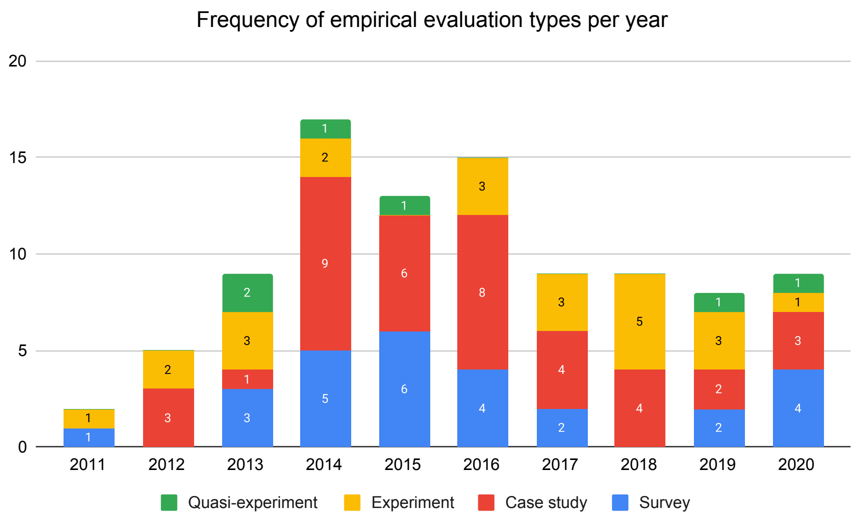

Figure 4.11: Frequency of empirical evaluation types per year

Figure 4.11 depicts the frequency of empirical evaluations adopted by year, with 96 out of 106 studies. When analyzing the figure, it is possible to observe that survey and case study strategies have been frequently adopted through the years. In the case of surveys, studies generally use this empirical strategy to apply psychometric instruments.

It is noteworthy that case studies represent a high frequency of empirical approaches. Many studies document the research design as experiments (which we see in figure less frequent), but in an inconsistent way with the definition of experimental/quasi-experimental design (WOHLIN et al., 2012). In these cases, we classified the empirical evaluation type as case studies. Experiments and quasi-experiments are shown less frequently. This may be explained due to the complexity of handling personality as a variable in controlled experiments.

4.8
Discussion and Review of the Results

The findings resulting from two large systematic literature studies (the one herein reported and the one by Cruz et al. (2015)) show that MBTI is still the most used psychometric instrument to measure personality in SE research. Nonetheless, there is no consensus in literature regarding the validity of this instrument (WYRICH et al., 2019; CALEFATO et al., 2019).

However, the focus in this section is not to provide a deep and critical psychometric assessment of the mapped instruments, but to discuss an overview of how the use of MBTI based instruments in SE research relates to the guidelines reported by McDonald & Edwards (2007). Therefore, we extracted additional details of instruments (actual name, version/bibliographic references) and how they were employed (covering from administration to data interpretation), whenever possible. Not following recommendations of these guidelines, most studies did not report any bibliographic references and explicit versions of the instrument.

Furthermore, most studies also did not report anything different from "we use x to measure personality" regarding instruments application or "x is widely used in SE research to measure personality" to support the choice of the instrument.

Following the guidelines, we extracted data from a reader's perspective looking for "explicit details of types of test used, administration process, the qualifications of the testers" (MCDONALD; EDWARDS, 2007) by answering the following derived questions:

Has the study documented participation of a qualified tester? Only one study claims the participation of an MBTI certified practitioner to process data [S138]; however by using survey as empirical approach more refinements in data interpretation were limited.

Are there details to justify the choice of MBTI? No, the choice is primarily justified by the widely use and acceptance of this instrument by previous SE studies [S151, S163, S167, S168, S169, S172, S173] or by brief claims about the validity of the instrument and professional widespread use [S155].

Are there details about versions of MBTI? One study claims the use of *Form M* or *Form G*, but no references to specific versions were documented [S112, S138, S181]. Others document the use of some free test [S129] in an unspecified version [S126]. Also, some applied the psychometric instrument through a website [S169, S172, S177] or in a printed version [S155]. Most studies document Myers (1998) as bibliographic reference when they refer to MBTI, but this reference is about the model and its theoretical foundation, not the specific psychometric instrument.

How was MBTI administered? One study mentions having provided a participation consent form [S165]. Other studies document that instruments were self-administered by subjects through a survey empirical strategy [S130, S137, S138, S181] or in a range of time without further details (longitudinal study) [S122]. One study did not measure personality by means of MBTI, but proposes a solution mapping its dimensions against adequate soft skills in requirements elicitation techniques [S179].

How were the results of MBTI interpreted? In general, the identified studies did not report on the result interpretation. In one study, an interview followed the administration of the psychometric instrument to obtain refinements of the resulting personality traits [S112]. Additionally, one study reported the algorithm to interpret the results [S137].

In sum, we can see that despite existing literature, there is a lack of concern about how to handle personality and psychometric instruments in SE, especially the MBTI. McDonald & Edwards' guidelines date from 2007. More recently, we had a critical review by Usman & Minhas (2019) including a sample of Cruz et al. (2015) studies (see our background and related works in Section 2.3). The results of their review are consistent with our observations, leading us to the same conclusion that there was no progress in improving the adoption of the MBTI in SE research over the years.

We understand that researchers are usually restricted by document size in research papers. However, given the majority of studies mapped in journals, which are generally more extensive and detailed in terms of text, we believe that more details on how a relevant human factor as personality is handled could be provided (or at least be made available in open science repositories).

4.9
Threats to Validity

In this section, we discuss the findings of the updated mapping study regarding its threats to validity. We list the possible threats and procedures we took to mitigate those issues hereafter according to Petersen et al. (2015).

Theoretical validity: with respect to our search strategy, we relied on empirically assessed guidelines to search for new evidence to update secondary studies in SE. Based on the set of 90 studies covering forty years of personality research in SE, we identified 6702 new studies to be analyzed in a single forward snowballing iteration using Google Scholar and included 106 additional studies of them. We believe that as a result, we had a good coverage of the literature within the last fifty years.

Regarding study selection, the exclusion of short papers and grey literature can threaten the representativeness of the sample of selected studies. However, we adopted this strategy to prioritize complete and peer-reviewed studies. We noticed that short papers frequently did not provide the necessary information to answer our research questions during initial data extraction efforts.

Furthermore, given the huge quantity of papers to be analyzed (6702) we filtered out papers that did not include terms related to personality in the title and abstract. This decision was taken to make the study selection effort viable. We included synonyms and believe that this did not lead to relevant studies being excluded.

Finally, concerning the data extraction process, a threat can be the main control of one researcher in the mapping study execution, which can bring some bias to results in different ways. This threat was mitigated by exhausting reviewing extracted data and consensus meetings with the dissertation advisor. We had support from two additional researchers in the extraction process to cover years 2017 to 2020, but the researcher that extracted data for the remaining years (2011 to 2016) reviewed their extraction. Still, the data extraction process is error-prone. To improve the reliability in this process, all extracted data is auditable and openly available to the community.

Descriptive validity: our protocol is based on solid guidelines and an update of a comprehensive mapping study. The open coding method for answering our research questions may not help to provide an easily understandable overview. We incorporated some axial coding procedures when answering research questions that primarily use open coding accordingly to Table 3.2 (RQ1a and RQ1b).

Regarding transparency, we documented the entire process and packaged all generated artifacts organized by the followed steps (see Figure 3.2) in order to turn it available to the community. They allow further analyses and replication of our protocol. Studies that were not included are flagged with their respective exclusion criteria.

Generalizability: The present mapping study is restricted to the dispositional personality perspective in SE research. More perspectives that could interest some target audiences may have been adopted in the SE literature. However, they were not captured by this protocol, and it is not our focus.

Further, we do not have any formal psychology or social sciences qualifications. We relied on literature guidelines to conduct this study and help consolidate a body of knowledge. Consequently, our view of personality and psychometric instruments may not be comprehensive enough.

5
Conclusions

This dissertation aimed to provide a comprehensive literature mapping concerning psychometric instruments used in SE research regarding personality, updating a broad existing secondary study (CRUZ et al., 2015). We provided a detailed protocol based on specific guidelines that met the need for an update and a search strategy for new evidence. While the updated secondary study included 90 studies covering 1970 to 2010, we identified 106 studies covering 2011 to 2020. In the following, we discuss the contributions, limitations, and future work.

5.1
Contributions

By answering our protocol's established research questions (Chapter 3), we contribute to the Behavioral Software Engineering (LENBERG et al., 2015) body of knowledge on the following topics with respect to personality:

- *Common objectives*: we observed common objectives of mapped studies that use personality, employing coding procedures to provide an overview of the most studied topics. *Investigate the effect of personality* in some SE activity contexts had the highest frequency, followed by *characterize software engineer personality*, which aims to discover and distinguish the personality of software engineering professionals and systematizing it, mostly for mapping roles and skills.
- *Limitations*: the limitations regarding the adoption of psychometric instruments are poorly reported in the mapped research. In fact, less than half of the primary studies in our set reported some limitations on the adoption of psychometric instruments.
- *Theoretical constructs*: we mapped the use of psychometric instruments within recent SE research related to archetypal classes of constructs. We observed that the instruments are mainly used within the context of **actors** *academic setting* and *researcher* who applied the **intervention** *personality traits data* to perform **activities** such as *software engineer characterization*, *team building*, *assessment of activities execution*, *pair programming*, and *mapping of suitable roles*, mostly without considering a **software system** (*none*) and sometimes related to *class assignments* and varied *tools*.

– A summary of the type of research and the empirical evaluations employing psychometric instruments. Validation research is the most common type of the mapped studies. It depicts research conducted in academic scenarios or not proposing something new and measuring in practice. With respect to the empirical evaluations, surveys and case studies are generally adopted.

– Finally, we observed remaining discrepancies between the application process of MBTI based instruments within recent SE research and existing recommendations in the literature. Even with long time existing guidelines on this instrument, we can not see any improvements so far.

5.2 Limitations

Although we performed an update of the secondary study by Cruz et al. (2015), we did not triangulate our data with their study since our goal and research questions differ. The authors of the updated study focused on characterizing the SE research on personality more broadly. In our turn, we focused on psychometric instruments. Although their study also mapped the instruments by answering the research question “What personality tests are administered in the studies, and to what type of participants (professionals or students)?”, they did not openly provide the detailed spreadsheet of extracted data. This limited us to work only with the information reported in their study paper.

We did not have psychology background, limiting our overall understanding. Even with the variety of instruments mapped among the personality models presented (especially the ones from the Big-Five/Five-Factor Model), we only conducted a partial review involving the use of the MBTI based on the guidelines of the SE literature that fit a personality perspective (dispositional). Although we defined this focus, we believe that the review is still valuable due to the findings presented on the MBTI, which is the most used psychometric instrument.

5.3 Future Work

Future work includes extending the synthesis presented by involving experts, collecting the point of view of social science researchers on the use of psychometric instruments in SE regarding personality, also covering other psychometric instruments. In particular, we propose the extended work to investigate the following additional research question:

> **RQ2:** How do social sciences researchers perceive the adoption of psychometric instruments regarding personality in SE research?

This could, for instance, involve applying a survey-based approach or conducting focus group sessions to gather further insights, producing guidelines that could be steering future SE research regarding personality in more promising directions.

6
Bibliography

ANDERSON, N. et al. **Handbook of Industrial, Work & Organizational Psychology: Volume 1: Personnel Psychology**. [S.l.]: SAGE, 2001. Google-Books-ID: Q8V4UVbQ_MYC. ISBN 978-1-84787-950-9.

BARROSO, A. S. et al. Influence of Human Personality in Software Engineering - A Systematic Literature Review. In: . [s.n.], 2017. p. 53–62. ISBN 978-989-758-249-3. Disponível em: <https://www.scitepress.org/Link.aspx?doi=10.5220/0006292000530062>.

CALEFATO, F.; LANUBILE, F.; VASILESCU, B. A large-scale, in-depth analysis of developers' personalities in the Apache ecosystem. **Information and Software Technology**, v. 114, p. 1–20, out. 2019. ISSN 0950-5849. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0950584918301216>.

CRUZ, S.; SILVA, F. Q. B. da; CAPRETZ, L. F. Forty years of research on personality in software engineering: A mapping study. **Computers in Human Behavior**, v. 46, p. 94–113, maio 2015. ISSN 0747-5632. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0747563214007237>.

CRUZ, S. S. J. O. et al. Personality in software engineering: Preliminary findings from a systematic literature review. In: **15th Annual Conference on Evaluation Assessment in Software Engineering (EASE 2011)**. [S.l.: s.n.], 2011. p. 1–10.

FELDT, R. et al. Towards Individualized Software Engineering: Empirical Studies Should Collect Psychometrics. In: **Proceedings of the 2008 International Workshop on Cooperative and Human Aspects of Software Engineering**. New York, NY, USA: ACM, 2008. (CHASE '08), p. 49–52. ISBN 978-1-60558-039-5. Event-place: Leipzig, Germany. Disponível em: <http://doi.acm.org/10.1145/1370114.1370127>.

FELIPE, D. A.; KALINOWSKI, M. **On Psychometric Instruments in Software Engineering Research Regarding Personality**. Zenodo, 2022. Type: dataset. Disponível em: <https://zenodo.org/record/6567567>.

GARNER, P. et al. When and how to update systematic reviews: consensus and checklist. **BMJ**, v. 354, jul. 2016. ISSN 1756-1833. Publisher: British Medical Journal Publishing Group Section: Research Methods & Reporting. Disponível em: <https://www.bmj.com/content/354/bmj.i3507>.

GAROUSI, V.; FERNANDES, J. M. Highly-cited papers in software engineering: The top-100. **Information and Software Technology**, v. 71, p. 108–128, mar. 2016. ISSN 0950-5849. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0950584915001871>.

GRAZIOTIN, D. et al. Consequences of Unhappiness While Developing Software. In: **Proceedings of the 2Nd International Workshop on Emotion Awareness in Software Engineering**. Piscataway, NJ, USA: IEEE Press, 2017. (SEmotion

'17), p. 42–47. ISBN 978-1-5386-2793-8. Event-place: Buenos Aires, Argentina. Disponível em: <https://doi.org/10.1109/SEmotion.2017...5>.

GRAZIOTIN, D. et al. What happens when software developers are (un)happy. **Journal of Systems and Software**, v. 140, p. 32–47, jun. 2018. ISSN 0164-1212. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0164121218300323>.

GRAZIOTIN, D.; WANG, X.; ABRAHAMSSON, P. Understanding the Affect of Developers: Theoretical Background and Guidelines for Psychoempirical Software Engineering. In: **Proceedings of the 7th International Workshop on Social Software Engineering**. New York, NY, USA: ACM, 2015. (SSE 2015), p. 25–32. ISBN 978-1-4503-3818-9. Event-place: Bergamo, Italy. Disponível em: <http://doi.acm.org/10.1145/2804381.2804386>.

GRAZIOTIN, D.; WANG, X.; Pekka Abrahamsson. The Affect of Software Developers: Common Misconceptions and Measurements. In: . [S.l.: s.n.], 2015.

HARZING, A.-W. **Publish or Perish**. 2020. Disponível em: <https://harzing.com/resources/publish-or-perish>.

KITCHENHAM, B. **Guidelines for performing Systematic Literature Reviews in Software Engineering, version 2.3**. [S.l.], 2007.

LENBERG, P.; FELDT, R.; WALLGREN, L. G. Behavioral software engineering: A definition and systematic literature review. **Journal of Systems and Software**, v. 107, p. 15–37, set. 2015. ISSN 0164-1212. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0164121215000989>.

MCDONALD, S.; EDWARDS, H. M. Who should test whom? **Communications of the ACM**, v. 50, n. 1, p. 66–71, jan. 2007. ISSN 0001-0782. Disponível em: <https://doi.org/10.1145/1188913.1188919>.

MENDES, E. et al. When to update systematic literature reviews in software engineering. **Journal of Systems and Software**, v. 167, p. 110607, set. 2020. ISSN 0164-1212. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0164121220300856>.

MICHELL, J. **Measurement in Psychology: A Critical History of a Methodological Concept**. [S.l.]: Cambridge University Press, 1999. Google-Books-ID: oNIcvjpDQeQC. ISBN 978-1-139-42560-5.

MOURAO, E. et al. Investigating the Use of a Hybrid Search Strategy for Systematic Reviews. In: **2017 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM)**. [S.l.: s.n.], 2017. p. 193–198.

MYERS, I. B. **Mbti Manual: A Guide to the Development and Use of the Myers - Briggs Type Indicator**. Palo Alto, Calif: Consulting Psychologists Press, 1998. ISBN 978-0-89106-130-4.

PETERSEN, K. et al. Systematic mapping studies in software engineering. In: **Proceedings of the 12th international conference on Evaluation and Assessment in Software Engineering**. Swindon, GBR: BCS Learning & Development Ltd., 2008. (EASE'08), p. 68–77.

PETERSEN, K.; VAKKALANKA, S.; KUZNIARZ, L. Guidelines for conducting systematic mapping studies in software engineering: An update. **Information and Software Technology**, v. 64, p. 1–18, ago. 2015. ISSN 0950-5849. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0950584915000646>.

RYCKMAN, R. M. **Theories of Personality**. [S.l.]: Cengage Learning, 2012. Google-Books-ID: uckKAAAAQBAJ. ISBN 978-1-285-40181-2.

SJøBERG, D. I. K. et al. Building Theories in Software Engineering. In: SHULL, F.; SINGER, J.; SJøBERG, D. I. K. (Ed.). **Guide to Advanced Empirical Software Engineering**. London: Springer, 2008. p. 312–336. ISBN 978-1-84800-044-5. Disponível em: <https://doi.org/10.1007/978-1-84800-044-5_12>.

STOL, K.; RALPH, P.; FITZGERALD, B. Grounded Theory in Software Engineering Research: A Critical Review and Guidelines. In: **2016 IEEE/ACM 38th International Conference on Software Engineering (ICSE)**. [S.l.: s.n.], 2016. p. 120–131.

USMAN, M.; MINHAS, N. M. Use of Personality Tests in Empirical Software Engineering Studies A Review of Ethical Issues. In: **Proceedings of the Evaluation and Assessment on Software Engineering**. New York, NY, USA: Association for Computing Machinery, 2019. (EASE '19), p. 237–242. ISBN 978-1-4503-7145-2. Disponível em: <https://doi.org/10.1145/3319008.3319032>.

WIERINGA, R. et al. Requirements Engineering Paper Classification and Evaluation Criteria: A Proposal and a Discussion. **Requir. Eng.**, v. 11, n. 1, p. 102–107, dez. 2005. ISSN 0947-3602. Disponível em: <http://dx.doi.org/10.1007/s00766-005-0021-6>.

WOHLIN, C. et al. Guidelines for the search strategy to update systematic literature reviews in software engineering. **Information and Software Technology**, v. 127, p. 106366, nov. 2020. ISSN 0950-5849. Disponível em: <http://www.sciencedirect.com/science/article/pii/S095058491930223X>.

WOHLIN, C. et al. **Experimentation in Software Engineering**. Berlin Heidelberg: Springer-Verlag, 2012. ISBN 978-3-642-29043-5. Disponível em: <https://www.springer.com/gp/book/9783642290435>.

WYRICH, M.; GRAZIOTIN, D.; WAGNER, S. A theory on individual characteristics of successful coding challenge solvers. **PeerJ Computer Science**, v. 5, p. e173, fev. 2019. ISSN 2376-5992. Publisher: PeerJ Inc. Disponível em: <https://peerj.com/articles/cs-173>.

A
Final list of selected primary studies

[S91] Omar, M., Syed-Abdullah, S.-L., et al. (2011). *"Developing a Team Performance Prediction Model: A Rough Sets Approach"*. Informatics Engineering and Information Science (pp. 691–705). Springer.

[S92] Salleh, N., Mendes, E., et al. (2011). *"The effects of openness to experience on pair programming in a higher education context"*. 2011 24th IEEE-CS Conference on Software Engineering Education and Training (CSEET), 149–158.

[S93] Raza, A., UlMustafa, Z., et al. (2011). *"Personality Dimensions and Temperaments of Engineering Professors and Students – A Survey"*. Journal of Computing, 3(12), 13–20.

[S94] Yilmaz, M.,& OConnor, R. V. (2012). *"Towards the Understanding and Classification of the Personality Traits of Software Development Practitioners: Situational Context Cards Approach"*. 2012 38th Euromicro Conference on Software Engineering and Advanced Applications, 400–405.

[S95] Gómez, M. N., Acuña, S. T., et al. (2012). *"How Does the Extraversion of Software Development Teams Influence Team Satisfaction and Software Quality?: A Controlled Experiment"*. International Journal of Human Capital and Information Technology Professionals (IJHCITP), 3(4), 11–24.

[S96] Sfetsos, P., Adamidis, P., et al. (2012). *"Investigating the Impact of Personality and Temperament Traits on Pair Programming: A Controlled Experiment Replication"*. 2012 Eighth International Conference on the Quality of Information and Communications Technology, 57–65.

[S97] Stylianou, C.,& Andreou, A. S. (2012). *"A Multi-objective Genetic Algorithm for Software Development Team Staffing Based on Personality Types"*. Artificial Intelligence Applications and Innovations (pp. 37–47). Springer.

[S98] Stylianou, C., Gerasimou, S., et al. (2012). *"A Novel Prototype Tool for Intelligent Software Project Scheduling and Staffing Enhanced with Personality Factors"*. 2012 IEEE 24th International Conference on Tools with Artificial Intelligence, 1, 277–284. https://doi.org/10.1109/ICTAI.2012.45

[S99] Montequín, V. R., Balsera, J. V., et al. (2012). *"Using Myers-Briggs Type Indicator (MBTI) as a Tool for Setting up Student Teams for Information*

Technology Projects". Journal of Information Technology and Application in Education, 1, 7.

[S100] Mair, C., Martincova, M., et al. (2012). *"An Empirical Study of Software Project Managers Using a Case-Based Reasoner"*. 2012 45th Hawaii International Conference on System Sciences, 1030–1039.

[S101] Gorla, N., Chiravuri, A., et al. (2013). *"Effect of personality type on structured tool comprehension performance"*. Requirements Engineering, 18(3), 281–292.

[S102] Koroutchev, K., Acuña, S. T., et al. (2013). *"The Social Environment as a Determinant for the Impact of the Big Five Personality Factors and the Group's Performance"*. International Journal of Human Capital and Information Technology Professionals (IJHCITP), 4(1), 1–8.

[S103] Sfetsos, P, Adamidis, P., et al. (2013). *"Heterogeneous Personalities Perform Better in Pair Programming: The Results of a Replication Study"*. Software Quality Professional Magazine, 15(4)

[S104] Bishop, D. (2013). *"Personality Theory as a Predictor for Agile Preference"*. MWAIS 2013 Proceedings.

[S105] Kanij, T., Merkel, R., et al. (2013). *"An empirical study of the effects of personality on software testing"*. 2013 26th International Conference on Software Engineering Education and Training (CSEE T), 239–248.

[S106] Radhakrishnan, P.,& Kanmani, S. (2013). *"Improvement of programming skills using pair programming by boosting extraversion and openness to experience"*. International Journal of Teaching and Case Studies, 4(1), 13–35.

[S107] Varathan, K. D.,& Thiam, L. T. (2013) *"Mining Facebook in Identifying Software Engineering Students' Personality and Job Matching"*. ISSN:2188-272X – The Asian Conference on Society, Education and Technology 2013 – Official Conference Proceedings.

[S108] Karapıçak, Ç. M.,& Demirörs, O. (2013). *"A Case Study on the Need to Consider Personality Types for Software Team Formation"*. Software Process Improvement and Capability Determination (pp. 120–129). Springer.

[S109] Thiti Phiriyayotha, Siriluck Rotchanakitumnuai. (2013). *"Data Warehouse Implementation Success Factors and the Impact of Leadership and Personality on* *the Relationship between Success Factors"*. Journal of Business and Economics, 4(10), 948–956.

[S110] Luse, A., McElroy, J. C., et al. (2013). *"Personality and cognitive style as predictors of preference for working in virtual teams"*. Computers in Human Behavior, 29(4), 1825–1832.

[S111] Martínez, L. G., Licea, G., et al. (2013). *"Using MatLab's fuzzy logic toolbox to create an application for RAMSET in software engineering courses"*. Computer Applications in Engineering Education, 21(4), 596–605.

[S112] Raza, A., Capretz, L.F., Ul-Mustafa, Z., 2014. *"Personality Profiles of Software Engineers and Their Software Quality Preferences"*. International Journal of Information Systems and Social Change (IJISSC) 5, 77–86.

[S113] Salleh, N., Mendes, E., Grundy, J., 2014. *"Investigating the effects of personality traits on pair programming in a higher education setting through a family of experiments"*. Empirical Software Engineering 19, 714–752.

[S114] Bishop, D., Deokar, A., 2014. *"Toward an Understanding of Preference for Agile Software Development Methods from a Personality Theory Perspective"*. 47th Hawaii International Conference on System Sciences, pp. 4749–4758.

[S115] Venkatesan, V., Sankar, A., 2014. *"INVESTIGATION OF STUDENTS PERSONALITY ON PAIR PROGRAMMING TO ENHANCE THE LEARNING ACTIVITY IN THE ACADEMIA"*. Journal of Computer Science 10, 2020–2028.

[S116] Licorish, S.A., MacDonell, S.G., 2014. *"Personality profiles of global software developers"*, in: Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering, EASE '14. Association for Computing Machinery, New York, NY, USA, pp. 1–10.

[S117] Gómez, M.N., Acuña, S.T., 2014. *"A replicated quasi-experimental study on the influence of personality and team climate in software development"*. Empir Software Eng 19, 343–377.

[S118] Kosti, M.V., Feldt, R., Angelis, L., 2014. *"Personality, emotional intelligence and work preferences in software engineering: An empirical study"*. Information and Software Technology 56, 973–990.

[S119] Okike, E.U., 2014. *"Bipolar Factor and Systems Analysis skills of Student Computing Professionals at University of Botswana, Gaborone"*. International Journal of Advanced Computer Science and Applications (IJACSA) 5.

[S120] Okike, E.U., 2014. *"Investigating Students' Achievements in Computing Science Using Human Metric."* International Journal of Advanced Computer Science and Applications (IJACSA) 5.

[S121] Yilmaz, M., O'Connor, R.V., Clarke, P., 2014. *"An Exploration of Individual Personality Types in Software Development"*, in: Barafort, B., O'Connor, R.V., Poth, A., Messnarz, R. (Eds.), Systems, Software and Services Process Improvement, Communications in Computer and Information Science. Springer, Berlin, Heidelberg, pp. 111–122.

[S122] Kruck, S.E., Sendall, P., Ceccucci, W., Peslak, A., Hunsinger, S., 2014. *"Does Personality Play a Role in Computer Information Systems Performance?"*. Issues in Information Systems 15, 383–392.

[S123] Ezekiel U. Okike and Olanrewaju A. Amoo, 2014. *"Problem Solving and Decision Making: Consideration of Individual Differences in Computer programming Skills Using Myers Briggs Type Indicator (MBTI) and Chidamber and Kemerer Java Metrics (CKJM)"*. 7, 8.

[S124] Papatheocharous, E., Belk, M., Nyfjord, J., Germanakos, P., Samaras, G., 2014. *"Personalised continuous software engineering"*, in: Proceedings of the 1st International Workshop on Rapid Continuous Software Engineering, RCoSE 2014. Association for Computing Machinery, New York, NY, USA, pp. 57–62.

[S125] Gramß, D., Frank, T., Rehberger, S., Vogel-Heuser, B., 2014. *"Female characteristics and requirements in software engineering in mechanical engineering"*, in: 2014 International Conference on Interactive Collaborative Learning (ICL). Presented at the 2014 International Conference on Interactive Collaborative Learning (ICL), pp. 272–279.

[S126] Okike, E.U., 2014. *"A Code Level Based Programmer Assessment and Selection Criterion Using Metric Tools"*. International Journal of Advanced Computer Science and Applications (IJACSA) 5.

[S127] Vitó Ferreira, N.N., Langerman, J.J., 2014. *"The correlation between personality type and individual performance on an ICT Project"*, in: 2014 9th International Conference on Computer Science Education. Presented at the 2014 9th International Conference on Computer Science Education, pp. 425–430.

[S128] Huang, F., Liu, B., Song, Y., Keyal, S., 2014. *"The links between human error diversity and software diversity: Implications for fault diversity seeking"*. Science of Computer Programming 89, 350–373. https://doi.org/10.1016/j.scico.2014.03.004

[S129] Ribaud, V., & Saliou, P. (2015). *"Relating ICT Competencies with Personality Types"*. Systems, Software and Services Process Improvement (pp. 295–302). Springer International Publishing.

[S130] Branco, D. T. M. C., Oliveira, E. C. C. de, Galvão, L., Prikladnicki, R., & Conte, T. (2015). *"An Empirical Study about the Influence of Project Manager Personality in Software Project Effort"*. 102–113. https://www.scitepress.org/Link.aspx?doi=10.5220/0005373001020113

[S131] Yilmaz, M., & O'Connor, R. V. (2015). *"Understanding personality differences in software organisations using Keirsey temperament sorter"*. IET Software, 9(5), 129–134.

[S132] Kanij, T., Merkel, R., & Grundy, J. (2015). *"An Empirical Investigation of Personality Traits of Software Testers"*. 2015 IEEE/ACM 8th International Workshop on Cooperative and Human Aspects of Software Engineering, 1–7.

[S133] Acuña, S. T., Gómez, M. N., Hannay, J. E., Juristo, N., & Pfahl, D. (2015). *"Are team personality and climate related to satisfaction and software quality? Aggregating results from a twice replicated experiment"*. Information and Software Technology, 57, 141–156.

[S134] Licorish, S. A., & MacDonell, S. G. (2015). *"Communication and personality profiles of global software developers"*. Information and Software Technology, 64, 113–131.

[S135] Karimi, Z., Baraani-Dastjerdi, A., Ghasem-Aghaee, N., & Wagner, S. (2015). *"Influence of Personality on Programming Styles an Empirical Study"*. Journal of Information Technology Research (JITR), 8(4), 38–56.

[S136] Monaghan, C., Bizumic, B., Reynolds, K., Smithson, M., Johns-Boast, L., & van Rooy, D. (2015). *"Performance of student software development teams: The influence of personality and identifying as team members"*. European Journal of Engineering Education, 40(1), 52–67.

[S137] Dargis, M., Finke, A., & Penicina, L. (2015). *"Relationship between Personality Types Conceptualized by C. G. Jung and Latvian IT Specialist Preferences"*. Complex Systems Informatics and Modeling Quarterly, 4, 1–11.

[S138] Capretz, L. F., Varona, D., & Raza, A. (2015). *"Influence of personality types in software tasks choices"*. Computers in Human Behavior, 52, 373–378.

[S139] Soomro, A. B., Salleh, N., & Nordin, A. (2015). *"How personality traits are interrelated with team climate and team performance in software engineering? A* *preliminary study"*. 2015 9th Malaysian Software Engineering Conference (MySEC), 259–265.

[S140] Suslow, W., Kowalczyk, J., Statkiewicz, M., Boinska, M., & Nowak, J. (2015). *"Psychosocial Correlates of Software Designers' Professional Aptitude"*. International Journal of Advanced Computer Science and Applications (IJACSA), 6(8), Article 8.

[S141] McDermott, R., Daniels, M., & Cajander, Å. (2015). *"Perseverance Measures and Attainment in First Year Computing Science Students"*. Proceedings of the 2015 ACM Conference on Innovation and Technology in Computer Science Education, 302–307.

[S142] Murukannaiah, P. K., Ajmeri, N., et al. (2016). *"Acquiring Creative Requirements from the Crowd: Understanding the Influences of Personality and Creative Potential in Crowd RE"*. 2016 IEEE 24th International Requirements Engineering Conference (RE), 176–185.

[S143] Ostberg, J.-P., Wagner, S., et al. (2016). *"Does Personality Influence the Usage of Static Analysis Tools? An Explorative Experiment"*. 2016 IEEE/ACM Cooperative and Human Aspects of Software Engineering (CHASE), 75–81.

[S144] Mujkanovic, A., & Bollin, A. (2016). *"Improving Learning Outcomes through Systematic Group Reformation—The Role of Skills and Personality in Software Engineering Education"*. 2016 IEEE/ACM Cooperative and Human Aspects of Software Engineering (CHASE), 97–103.

[S145] Eckhardt, A., Laumer, S., et al. (2016). *"The Effect of Personality on IT Personnel's Job-Related Attitudes: Establishing a Dispositional Model of Turnover Intention across IT Job Types"*. Journal of Information Technology, 31(1), 48–66.

[S146] Karimi, Z., Baraani-Dastjerdi, A., et al. (2016). *"Using Personality Traits to Understand the Influence of Personality on Computer Programming: An Empirical Study"*. Journal of Cases on Information Technology (JCIT), 18(1), 28–48.

[S147] Hsu, W.-C., & Barnes, G. M. (2016). *"Are computer information technology majors different than computer science majors in personality, learning style, or academic performance?"*. Journal of Computing Sciences in Colleges, 31(4), 100–107.

[S148] Agung, A. A. G., & Yuniar, I. (2016). *"Personality assessment website using DISC: A case study in information technology school"*. 2016 International Conference on Information Management and Technology (ICIMTech), 72–77.

[S149] Rastogi, A., & Nagappan, N. (2016). *"On the Personality Traits of GitHub Contributors"*. 2016 IEEE 27th International Symposium on Software Reliability Engineering (ISSRE), 77–86.

[S150] Karimi, Z., Baraani-Dastjerdi, A., et al. (2016). *"Links between the personalities, styles and performance in computer programming"*. Journal of Systems and Software, 111, 228–241.

[S151] Gilal, A. R., Jaafar, J., et al. (2016). *"Balancing the Personality of Programmer: Software Development Team Composition"*. Malaysian Journal of Computer Science, 29(2), 145–155.

[S152] Paruma Pabón, O. H., González, F. A., et al. (2016). *"Finding Relationships between Socio-Technical Aspects and Personality Traits by Mining Developer E-mails"*. 2016 IEEE/ACM Cooperative and Human Aspects of Software Engineering (CHASE), 8–14.

[S153] Aeron, S., & Pathak, S. (2016). *"Personality composition in Indian software teams and its relationship to social cohesion and task cohesion"*. International Journal of Indian Culture and Business Management, 13(3), 267–287.

[S154] Yilmaz, M., Yilmaz, M., et al. (2016). *"A Gamification Approach to Improve the Software Development Process by Exploring the Personality of Software Practitioners"*. Software Process Improvement and Capability Determination (pp. 71–83). Springer International Publishing.

[S155] John H. Reynolds, D. Robert Adams, et al. (2016). *"The Personality of a Computing Major: It Makes a Difference"*. EDSIGCON Proceeding 2016, 2.

[S156] Kosti, M. V., Feldt, R., et al. (2016). *"Archetypal personalities of software engineers and their work preferences: A new perspective for empirical studies"*. Empirical Software Engineering, 21(4), 1509–1532.

[S157] Xia, X., Lo, D., et al. (2017). *"Personality and Project Success: Insights from a Large-Scale Study with Professionals"*. 2017 IEEE International Conference on Software Maintenance and Evolution (ICSME), 318–328.

[S158] Gilal, A. R., Jaafar, J., et al. (2017). *"Making Programmer Effective for Software Development Teams: An Extended Study"*. Journal of Information Science and Engineering, 33(6), 17.

[S159] Rehman, M., Safdar, S., et al. (2017). *"Personality traits and knowledge sharing behavior of software engineers"*. 2017 8th International Conference on Information Technology (ICIT), 6–11.

[S160] Yilmaz, M., O'Connor, R. V., et al. (2017). *"An examination of personality traits and how they impact on software development teams"*. Information and Software Technology, 86, 101–122.

[S161] Çelikten, A.,& Çetin, A. (2017). *"Assigning Product Development Roles to Software Engineers Based on Personality Types and Skills"*. Academic Journal of Science, 7(3), 475–486.

[S162] Tunio, M. Z., Luo, H., et al. (2017). *"Impact of Personality on Task Selection in Crowdsourcing Software Development: A Sorting Approach"*. IEEE Access, 5, 18287–18294.

[S163] Gilal, A. R., Jaafar, J., et al. (2017). *"Suitable Personality Traits for Learning Programming Subjects: A Rough-Fuzzy Model"*. International Journal of Advanced Computer Science and Applications (IJACSA), 8(8), Article 8.

[S164] Barroso, A. S., Madureira, J. S., et al. (2017). *"Relationship between Personality Traits and Software Quality—Big Five Model vs. Object-oriented Software Metrics"*. 63–74.

[S165] Gilal, A. R., Omar, M., et al. (2017). *"Software development team composition: Personality types of programmer and complex network"*. 153–159.

[S166] Jevsikova, T., Dagienė, V., et al. (2018). *"On Preferences of Novice Software Engineering Students: Temperament Style and Attitudes Towards Programming Activities"*. Fundamentals of Computer Science and Software Engineering (pp. 101–113). Springer International Publishing.

[S167] Tunio, M. Z., Luo, H., Wang, C.,& Zhao, F. (2018). *"Crowdsourcing Software Development: Task Assignment Using PDDL Artificial Intelligence Planning"*. Journal of Information Processing Systems, 14(1), 129–139.

[S168] Tunio, M. Z., Luo, H., Wang, C., Zhao, F., et al. (2018). *"Task Assignment Model for Crowdsourcing Software Development: TAM"*. Journal of Information Processing Systems, 14(3), 621–630.

[S169] Perez-Gonzalez, H. G., Nunez-Varela, A., et al. (2018). *"Investigating the Effects of Personality on Software Design in a Higher Education Setting Through an Experiment"*. 2018 6th International Conference in Software Engineering Research and Innovation (CONISOFT), 72–78.

[S170] Sánchez, M. G. T., Cachero, C., et al. (2018). *"Evaluating the Effect of Developers' Personality and Productivity on their Intention to Use Model-Driven*

Web Engineering Techniques: An Exploratory Observational Study". Journal of Web Engineering, 483–526.

[S171] Anderson, G., Keith, M. J., et al. (2018). *"The Effect of Software Team Personality Composition on Learning and Performance: Making the "Dream" Team"*. https://doi.org/10.24251/HICSS.2018.059

[S172] Garousi, V.,& Tarhan, A. (2018). *"Investigating the Impact of Team Formation by Introversion/Extraversion in Software Projects"*. Balkan Journal of Electrical and Computer Engineering, 6(2), 132–140.

[S173] Gilal, A. R., Jaafar, J., et al. (2018). *"Finding an effective classification technique to develop a software team composition model"*. Journal of Software: Evolution and Process, 30(1), e1920.

[S174] Shuto, M., Washizaki, H., et al. (2018). *"Personality and Learning Effectiveness of Teams in Information Systems Education Courses"*. EAI Endorsed Transactions on E-Learning, 5(17).

[S175] Ting, T. L.,& Varathan, K. D. (2018).*"JOB RECOMMENDATION USING FACEBOOK PERSONALITY SCORES"*. Malaysian Journal of Computer Science, 31(4), 311–331.

[S176] Chowdhury, S., Walter, C., et al. (2018). *"Toward Increasing Collaboration Awareness in Software Engineering Teams"*. 2018 IEEE Frontiers in Education Conference (FIE), 1–9.

[S177] Barroso, A. S., de J. Prado, K. H., et al. (2019). *"How personality traits influences quality of software developed by students"*. Proceedings of the XV Brazilian Symposium on Information Systems, 1–8.

[S178] Mellblom, E., Arason, I., et al. (2019). *"The Connection Between Burnout and Personality Types in Software Developers"*. IEEE Software, 36(5), 57–64.

[S179] Khan, A., Gilal, A. R., et al. (2019). *"Does Software Requirement Elicitation and Personality make any relation?"*. Journal of Advanced Research in Dynamic and Control Systems, Volume 11(08-Special Issue), 1162–1168.

[S180] Calefato, F., Lanubile, F., et al. (2019). *"A large-scale, in-depth analysis of developers' personalities in the Apache ecosystem"*. Information and Software Technology, 114, 1–20.

[S181] Hasan, A., Moin, S., et al. (2019). *"Prediction of Personality Profiles in the Pakistan Software Industry–A Study"*. Psych, 1(1), 320–330.

[S182] Qamar, N.,& Malik, A. A. (2019). *"Birds of a Feather Gel Together: Impact of Team Homogeneity on Software Quality and Team Productivity"*. IEEE Access, 7, 96827–96840.

[S183] Wyrich, M., Graziotin, D., et al. (2019). *"A theory on individual characteristics of successful coding challenge solvers"*. PeerJ Computer Science, 5, e173.

[S184] Akarsu, Z., Orgun, P., et al. (2019). *"Assessing Personality Traits in a Large Scale Software Development Company: Exploratory Industrial Case Study"*. Systems, Software and Services Process Improvement, 192–206.

[S185] Celikten, A., Kurt, E., et al. (2019). *"A Decision Support System for Role Assignment in Software Project Management with Evaluation of Personality Types"*. Artificial Intelligence and Applied Mathematics in Engineering Problems (pp. 200–210). Springer International Publishing.

[S186] Khan, I. A., Khan, A., et al. (2019). *"Urdu Translation: The Validation and Reliability of the 120-Item Big Five IPIP Personality Scale"*. Current Psychology, 38(6), 1530–1541.

[S187] Gomes, A., Silva, M., et al. (2020). *"Evaluating the Relationship of Personality and Teamwork Quality in the Context of Agile Software Development"*. SEEK 2020 Proceedings, 6. https://ksiresearch.org/seke/seke20paper/paper158.pdf

[S188] Chiu, Y.-T., Day, M.-Y., et al. (2020). *"Exploring impacts of project leaders' written expressions in virtual and fluid projects: The role of personality and emotion"*. International Research Workshop on IT Project Management 2020. https://aisel.aisnet.org/irwitpm2020/5

[S189] Amin, A., Basri, S., et al. (2020). *"The impact of personality traits and knowledge collection behavior on programmer creativity"*. Information and Software Technology, 128, 106405.

[S190] Akarsu, Z.,& Yilmaz, M. (2020). *"Managing the social aspects of software development ecosystems: An industrial case study on personality"*. Journal of Software: Evolution and Process, 32(11), e2277.

[S191] Russo, D.,& Stol, K.-J. (2020). *"Gender Differences in Personality Traits of Software Engineers"*. IEEE Transactions on Software Engineering, 48(3), 819–834.

[S192] Doungsa-ard, C.,& Chaiwon, V. (2020). *"The Software Engineering Position Mapping From Personality Traits"*. 2020 Joint International Conference on Digital

Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI DAMT NCON), 194–199.

[S193] Demir, Ö.,& Seferoglu, S. S. (2020). *"The Effect of Determining Pair Programming Groups According to Various Individual Difference Variables on Group Compatibility, Flow, and Coding Performance"*. Journal of Educational Computing Research, 59(1), 41–70.

[S194] Papoutsoglou, M., Kapitsaki, G. M., et al. (2020). *"Modeling the effect of the badges gamification mechanism on personality traits of Stack Overflow users"*. Simulation Modelling Practice and Theory, 105, 102157.

[S195] Vishnubhotla, S. D., Mendes, E., et al. (2020). *"Investigating the relationship between personalities and agile team climate of software professionals in a telecom company"*. Information and Software Technology, 126, 106335.

[S196] Tehlan, K., Chakraverty, S., et al. (2020). *"A genetic algorithm-based approach for making pairs and assigning exercises in a programming course"*. Computer Applications in Engineering Education, 28(6), 1708–1721.

B
Dictionary of coded constructs related to SE

Table B.1: Coded constructs in class Actor

Actor coded construct	Description	Studies	Count
academic setting	The study has an educational purpose or is applied in an academic setting due to scope limitations	S91, S92, S95, S96, S99, S101, S102, S103, S105, S106, S107, S110, S111, S113, S115, S120, S122, S125, S126, S133, S135, S136, S140, S141, S143, S144, S146, S147, S148, S150, S151, S155, S158, S162, S163, S165, S166, S169, S171, S172, S173, S174, S176, S177, S182, S193, S196	47
researcher	The study author(s) act primarily for investigative purposes	S93, S104, S108, S112, S114, S116, S117, S118, S119, S121, S123, S124, S128, S129, S131, S132, S134, S137, S138, S139, S145, S149, S152, S156, S159, S161, S167, S168, S170, S175, S178, S179, S180, S181, S183, S185, S186, S187, S188, S189, S191, S192, S194	43
industrial setting	The study is applied in a industrial setting	S109, S130, S142, S153, S157, S160, S164, S184, S190, S195	10
organization	The scope of the study is not clear (where does the study data come from?). The code was adopted to be comprehensive.	S97, S98, S100, S127	4
practitioner	The study is clearly applied and focused on practioners	S154	1
software development team	A software developent team was the interventor in the study	S94	1

Table B.2: Coded constructs in class Intervention

Intervention coded construct	Description	Studies	Count
personality traits data	The study has data on personality traits measured by some psychometric instrument	S91, S92, S93, S94, S95, S96, S97, S98, S99, S100, S101, S102, S103, S104, S105, S107, S108, S109, S110, S111, S112, S113, S114, S115, S116, S117, S118, S119, S120, S122, S123, S124, S125, S126, S127, S128, S129, S130, S131, S132, S133, S134, S135, S136, S137, S138, S139, S140, S141, S142, S143, S144, S145, S146, S147, S148, S149, S150, S151, S152, S153, S155, S156, S157, S158, S159, S160, S161, S162, S163, S164, S165, S166, S167, S168, S169, S170, S171, S172, S173, S174, S176, S177, S178, S179, S180, S181, S182, S183, S184, S185, S186, S187, S188, S189, S190, S191, S192, S193, S194, S195, S196	102
pair programming	The study adopted pair programming to assess some impact	S106	1
psychometric instrument	The study adopted a psychometric instrument to be adapted/validated	S121	1
social media personality traits data	The study has data on personality measure using social media data as source	S175	1
virtual environment to measure personality	The study used an virtual enviroment to measure personality as intervention	S154	1
software development team	A software developent team was the interventor in the study	S94	1

Table B.3: Coded constructs in class Activity

Activity coded construct	Description	Studies	Count
characterization of software engineer	The study characterizes, in some extent, the software engineering professional. It means: comparison of SE and other professionals, discover (and/or compare) personalities for some purpose.	S93, S108, S112, S116, S118, S119, S127, S129, S131, S132, S134, S137, S138, S147, S148, S149, S152, S155, S156, S160, S165, S179, S180, S184, S191, S194, S195	27
team building	The main activity of the study is to build software development teams with more than two members	S91, S94, S95, S97, S98, S99, S102, S111, S124, S133, S144, S153, S158, S171, S172, S174	16
assessment of activities execution	The main activity of the study is to assess software artifacts and processes in a software activity	S101, S117, S120, S122, S125, S126, S143, S159, S169, S182	10
pair programming	The main activity of the study is to build pair programming teams	S92, S96, S103, S113, S115, S193, S196	7
mapping of suitable roles	The main activity of the study is to mapping software engineer roles (developer, tester, project manager, etc.)	S107, S161, S175, S185, S192	5
investigate team performance	The main activity of study is investigate performance of a software team	S136, S139, S173, S176, S190	5
inquiry on programmer performance	The main activity is assess programmer performance in coding activities	S123, S128, S146, S163	4
project management activities	The main activity is related to software project management activities	S100, S130, S188	3
task selection	The main activity is related to selection of tasks to development	S162, S167, S168	3

Table B.4: Coded constructs in class Software System

Software system coded construct	Description	Studies	Count
class assignments	The study used a software system for academic settings, previously developed for some specific purpose (be tested, refactored, ...) or developed during the conduction of the study by students	S92, S95, S96, S99, S102, S103, S106, S113, S115, S117, S120, S123, S125, S126, S133, S135, S136, S140, S141, S144, S146, S147, S150, S151, S158, S163, S169, S170, S171, S172, S173, S174, S176, S177, S182, S193, S196	37
none	No software system was used in the study	S93, S104, S110, S112, S114, S118, S119, S121, S122, S127, S129, S131, S132, S137, S138, S139, S145, S153, S154, S155, S159, S160, S161, S165, S166, S178, S179, S181, S186, S187, S189, S191, S195	33
tool	The study used a machine learning/computational intelligence, some kind of algorithm, gamecard, or CASE tools	S91, S94, S97, S98, S100, S101, S107, S111, S124, S143, S148, S156, S175, S185, S192	15
software repository	The study used a software repository that stores some kind of software artifacts, (e.g. code, documents, or tasks)	S116, S134, S149, S152, S164, S180, S188, S194	8
given software project	The study uses software (or requirements of it) that has not yet been developed, but it was during the conduction of the study	S105, S108, S142, S157, S183, S184, S190	7
crowdsourcing projects	The study has crowdsourcing projects as a software system	S162, S167, S168	3
implementation of a dataware house system	The study used data from implementation of a dataware house system	S109	1
programming contest	The study used what is developed in a programming contest	S128	1
industry projects	The study used data regarding industrial projects	S130	1

C
Relation two by two of coded constructs related to SE

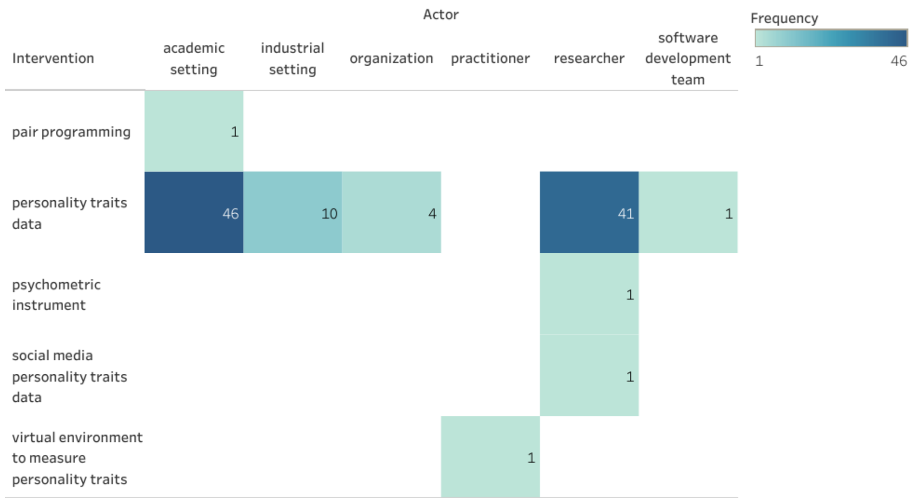

Figure C.1: Matrix representation of the coded archetypal classes **Actor** x **Intervention**

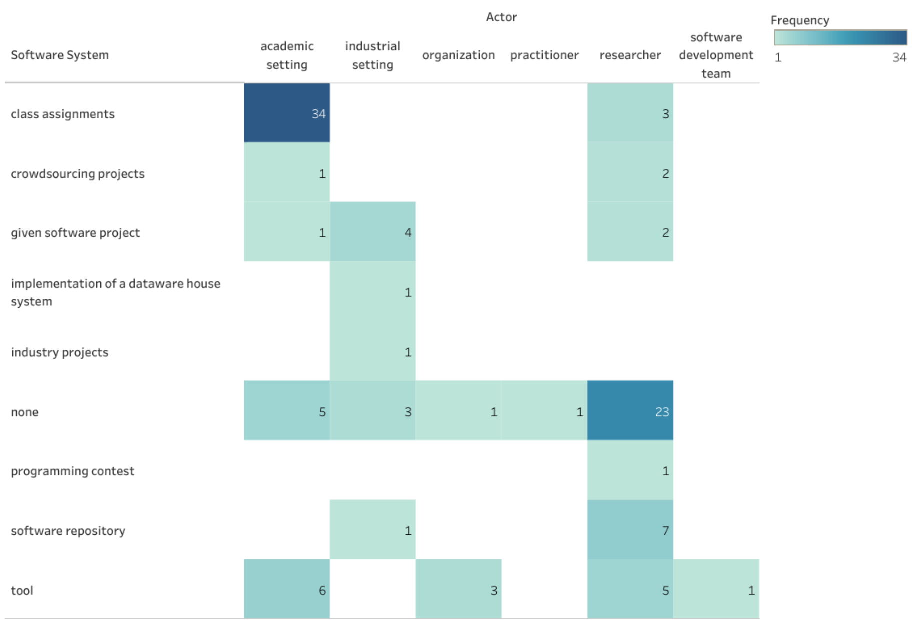

Figure C.3: Matrix representation of the coded archetypal classes **Actor** x **Software System**

Frequency: 1 – 14

Activity	Actor: academic setting	industrial setting	organization	practitioner	researcher	software development team
academic resilience	1					
assessment of activities execution	8				2	
assessment of impact	1	1				
boosting on personality traits	1					
characterization of software engineer	2	2	1		5	
check software engineer wellbeing					2	
develop software design	1					
inquiry on programmer performance	2				2	
inquiry on project success		1				
intention to use a technique					1	
investigate programmer creativity					1	
investigate programming performance					1	
investigate programming styles	2					
investigate quality of sotware	1					
investigate team performance	3	1			1	
investigate the quality of software		1				
investigate turnover intention					1	
mapping of suitable roles	1				4	
measure user experience				1		
pair programming	7					
perform the validity of a measure					1	
prediction of preferences	1				3	
project management activities		1	1		1	
propose a pyschometric instrument					1	
requiments elicitation activities		1				
software engineer characterization	2	1			14	
software engineer preferences	1					
software testing activities	1					
task selection	1				2	
team building	11	1	2		1	1

Figure C.2: Matrix representation of the coded archetypal classes **Actor** x **Activity**

Figure C.4: Matrix representation of the coded archetypal classes **Intervention** x **Activity**

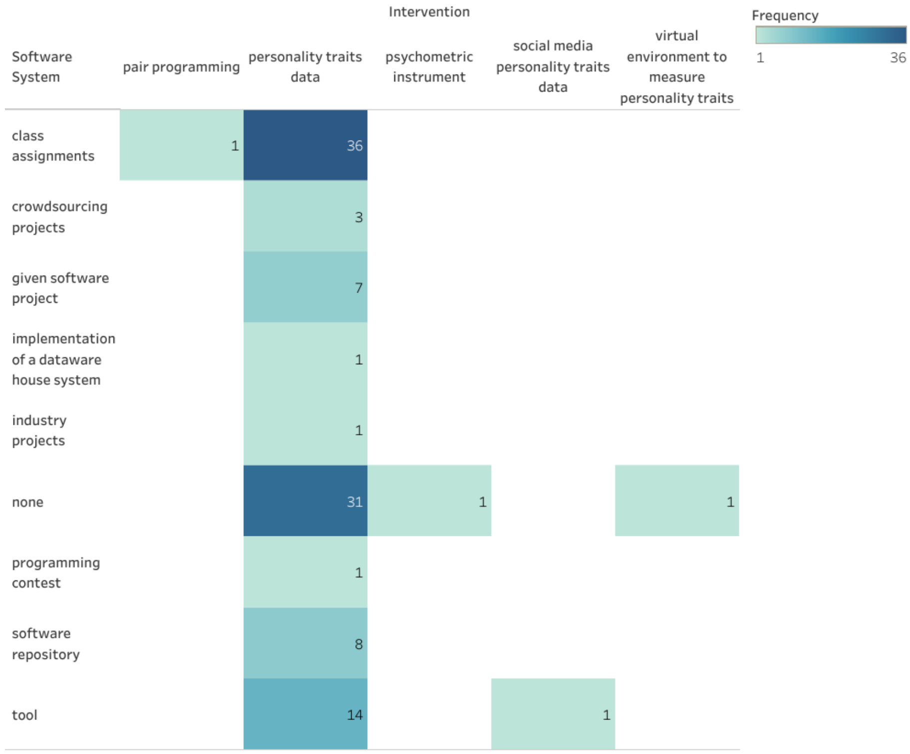

Figure C.5: Matrix representation of the coded archetypal classes **Intervention** x **Software System**

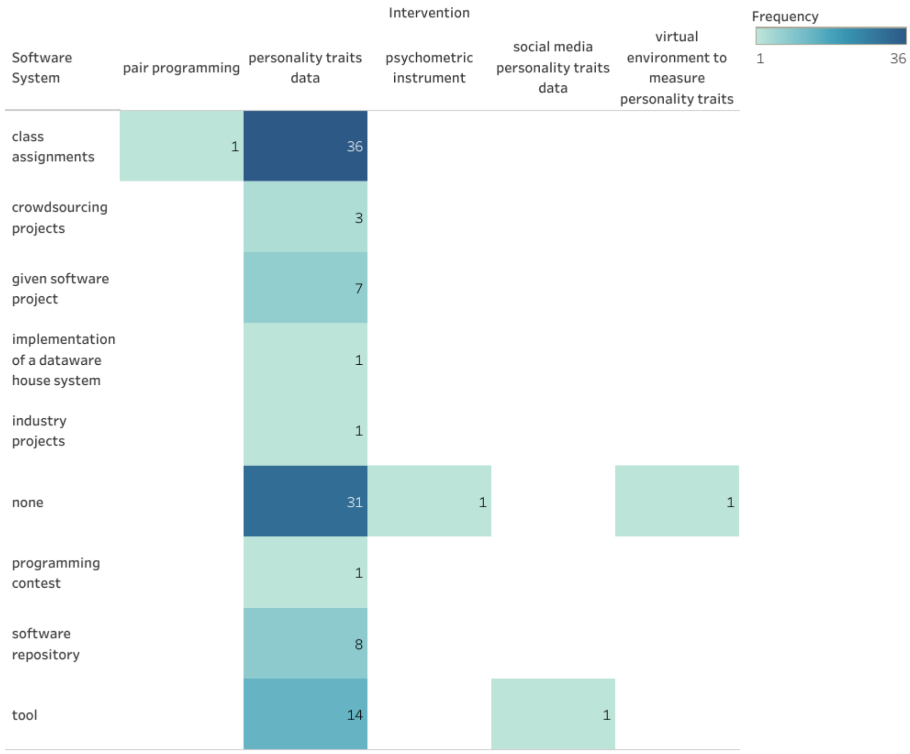

Figure C.6: Matrix representation of the coded archetypal classes **Activity** x **Software System**