Journal of Visual Language and Computing

journal homepage: www.ksiresearch.org/jvlc

A Semantic Comparative Analysis of Agile Teamwork Quality Instruments in Agile Software Development[⋆,⋆⋆]

Ramon Santos[a], Felipe Cunha[a], Thiago Rique[a], Mirko Perkusich[a], Ademar Neto[a], Danyllo Albuquerque[a], Hyggo Almeida[a] and Angelo Perkusich[a]

[a]Federal University of Campina Grande, Paraiba, Brazil

ARTICLE INFO

Article History:
Submitted 5.10.2023
Revised 8.7.2023
Accepted 8.8.2023

Keywords:
agile software development
teamwork quality
teamwork effectiveness
teamwork quality instruments
semantic comparative analysis

ABSTRACT

Multiple models (or instruments) for measuring Teamwork Quality (TWQ) and Teamwork Effectiveness (TWE) for Agile Software Development (ASD) have been created. Regardless, such models have different constructs and measures, with a limited understanding of how they are semantically related. [Objective] Our goal is to understand how specific instruments for ASD are related, considering the semantic relationship between them. [Method] We analyzed eight specific teamwork instruments for ASD (ASD instruments), comparing quantitative factors to identify which such instruments use most. Then, we compared them qualitatively from a semantic perspective, given that they are specific instruments in an agile context, considering the solid theories that support them. [Results] The results showed that Team Orientation and Coordination were identified among the top three rankings, both in the frequency of instrument questions and in the frequencies of literature-based Thematic Network themes. We found in our semantic analysis important themes associated a many instrument factors: Team Interaction associated with Communication factor, Acceptance of Goals associated with Coordination, etc. Qualitative concepts can be investigated considering the ASD factors from the knowledge of the identified parts of the agile instruments. [Conclusion] The semantic analysis brings new perspectives for researchers and practitioners to highlight more investigation about different teamwork aspects (new instruments themes) in ASD. We argue the need to add other ASD instruments to be compared to solidify the results found in this study, so we advocate further studies on this topic.

1. Introduction

The success of Agile Software Development (ASD) heavily relies on the competencies, interactions, and skills of its professionals [27, 31]. As software teams are the critical source of agility in ASD [32, 10], people are a crucial resource [23, 32, 3], and the quality of team interactions can significantly impact a project's outcome. Hence, Teamwork Quality (TWQ) is essential for agile projects' success [15, 6, 16]. The industry is rapidly adopting ASD [29], and the need for systematic team development [22] has compelled researchers to focus on teamwork aspects increasingly.

A team can be defined as a social system of two or more people which is embedded in an organization (context) whose members perceive themselves as such and are perceived as members by others (identity), collaborating on a common task (teamwork) [1, 12, 11]. The main focus of TWQ research is on the quality of interactions within teams rather than team members' (task) activities. Starting from the widespread fundamental proposition that the success of work conducted in teams depends (beyond the quantity and correctness of the task activities) on how well team members collaborate or interact.

The construct TWQ was proposed [13] as a comprehensive concept of the quality of team interactions. To capture the nature of team members working together, six facets of

ramon.santos@virtus.ufcg.edu.br (R. Santos); felipe.cunha@virtus.ufcg.edu.br (F. Cunha); thiago.rique@virtus.ufcg.edu.br (T. Rique); mirko@virtus.ufcg.edu.br (M. Perkusich); ademar.sousa@virtus.ufcg.edu.br (A. Neto); danyllo.albuquerque@virtus.ufcg.edu.br (D. Albuquerque); hyggo@virtus.ufcg.edu.br (H. Almeida); angelo.perkusich@virtus.ufcg.edu.br (A. Perkusich)

ORCID(s): 0000-0003-4864-0480 (R. Santos); 0000-0003-4864-0480 (F. Cunha); 0000-0003-0897-4953 (T. Rique); 0000-0002-9433-4962 (M. Perkusich); 0000-0002-1651-4159 (A. Neto); 0000-0001-5515-7812 (D. Albuquerque); 0000-0002-2808-8169 (H. Almeida); 0000-0002-7377-1258 (A. Perkusich)

DOI reference number: 10-18293/JVLC2023-N2-036

the collaborative team process integrate into the concept of TWQ: Communication, Coordination, Balance of Member Contribution, Mutual Support, Effort, and Cohesion. These facets capture both task-related and social interaction within teams. Research has shown that TWQ has a positive impact on team development [13], increasing the chances of succeeding with ASD. [13][20][22].

In this context, researchers have proposed instruments for assessing teamwork quality in the agile context, such as (i) a Radar Plot [21] that considers five dimensions for assessing TWQ: Shared Leadership, Orientation, Redundancy, Learning, and Autonomy; (ii) a Structural Equation Model [15] (TWQ-SEM), based on a differentiated replication from [13], which considered that the teamwork construct is comprised of six variables: Communication, Coordination, Balance of Member Contribution, Mutual Support, Effort, and Cohesion.

All the instruments mentioned are generic and cannot represent specific situations in the agile context. This was evidenced by the emergence of new instruments tailored for Agile Software Development (ASD). For instance, the aTWQ instrument [22] was developed based on the TWQ instrument [13], while the ATEM instrument [30] was developed based on the Big Five theory [25]. Additionally, a Bayesian networks-based model (TWQ-BN) [8] was developed based on the TWQ instrument [13]. Moreover, the TACT instrument [9] was developed based on the TCI instrument [2], and finally, the STEM instrument [33] was developed considering that some specific factors in Scrum.

Although the literature on TWQ has evolved, there was no unified understanding of what factors influence teamwork in ASD. Silva et al. [28] took a first step toward better understanding the relationship between agile TWQ instruments by performing a quantitative comparison between TWQ-SEM [15] and TWQ-BN [8]. However, the study is limited to only two instruments and focused on a high-level analysis (i.e., factors), not explicitly considering the instruments‘ questions.

Freire et al. [7] took a step further by developing a literature-based Thematic Network identifying the most frequent codes and themes in agile teamwork literature. Freire et al. [7] argued that researchers and practitioners can use the thematic network as a reference for understanding the factors and dimensions that comprise ASD Teamwork. With this, practitioners can, for example, define mechanisms to monitor such dimensions and use the collected data as a reference to drive actions toward improving the team's performance.

In our earlier research [26], we used Freire et al. [7]'s thematic network as a reference for analyzing three ASD teamwork instruments: ATEM, aTWQ, and TWQ-BN. However, we only performed a syntactic (i.e., quantitative) analysis, which brings many limitations, such as loss of information. This study complements our past research by considering eight ASD teamwork instruments and performing a semantic (i.e., qualitative) analysis. This paper provides a more comprehensive understanding of the interrelationships between factors and questions within the instruments, enhancing comprehension of their functioning.

Noteworthy enhancements and novel contributions in this paper, not covered in Santos et al. [26], include the following:

- **Expanded Scope of Comparison**: The quantity of compared ASD instruments has been increased to eight, all of which were identified in our systematic literature review (SLR) work, soon to be published in the 37th Brazilian Symposium on Software Engineering (SBES 2023).
- **Enhanced ASD Instruments Factors and Freire et al. [7]** Themes Comparison: The comparison now encompasses eight instruments, leading to more robust and dependable results.
- **Semantic Comparison**: A refined approach has been adopted for comparing instruments' questions based on a semantic analysis of their factors and questions.
- **Investigation of Teamwork Instruments Factors Evolution**: The association between the chronological evolution of the instruments and the evolution of subjects associated with the factors of these instruments has been thoroughly investigated, revealing discernible patterns and trends.

This paper is organized as follows: Section 2 presents teamwork theoretical concepts and general information on the ASD Teamwork instruments compared in this work. Section 3 describes the research questions. Section 4 presents the quantitative comparison. Section 5 presents the instruments' semantic comparison. Section 6 presents the Discussion of the results. Section 7 covers the study's limitations and threats to validity. Section 8 presents the study implications. Lastly, Section 9 presents our final remarks, discussing potential future work.

2. Background

The topic of TWQ assessment has garnered considerable attention in the ASD research community [8, 22, 9, 30, 33]. This section provides an overview of the main concepts related to this field of research relevant to our study. Section 2.1 defines what is a "teamwork instrument" in the scope of our research and elucidates the distinction between the concepts of "Team effectiveness" and "Team performance." Secondly, Section 2.2 presents a comprehensive overview of the eight ASD teamwork instruments objects of our study. We identified such instruments through a Systematic Literature Review (In press). Lastly, Section 2.3 discusses the theoretical evolution of ASD teamwork instruments.

2.1. Teamwork Models

This section defines what is "teamwork instrument" in the scope of our research and discusses fundamental concepts of teamwork models in software engineering.

Definition of a "teamwork instrument". A teamwork instrument is an assessment tool designed to capture and evaluate various factors pertaining to teamwork. Typically, it comprises questions or statements specifically crafted to gather information and assess specific aspects of team collaboration, communication, coordination, and other relevant

Table 1
Teamwork Quality Instruments used in Agile Software Development.

Instrument Number	Year	Title	Instrument
I1	2001	Teamwork Quality and the Success of Innovative Projects: A Theoretical Concept and Empirical Evidence	**TWQ** - Teamwork Quality
I2	2009	Putting Agile Teamwork to the Test – An Preliminary Instrument for Empirically Assessing and Improving Agile Software Development	**Radar Plot**
I3	2010	A teamwork model for understanding an agile team: A case study of a Scrum project	**ASTM-Agile Scrum Teamwork Model**
I4	2018	A Bayesian networks-based approach to assess and improve the teamwork quality of agile teams	**TWQ-BN** - Teamwork Quality - Bayesian network
I5	2020	Evaluation of Agile Team Work Quality	**aTWQ** - Agile Teamwork Quality
I6	2020	An Instrument to Assess the Organizational Climate of Agile Teams - A Preliminary Study	**TACT** - Assess the Organizational Climate of Agile Teams
I7	2022	A teamwork effectiveness model for agile software development	**ATEM** - Agile Team Effectiveness Model
I8	2022	A Theory of Scrum Team Effectiveness	**STEM** - Scrum Team Effectiveness Model

dimensions. Through administering such instruments, researchers or practitioners can systematically measure and evaluate different facets of teamwork, identify potential issues or barriers, and make informed decisions to enhance team performance and productivity. The prevalent technique employed in constructing these instruments is Structural Equation Modelling (SEM) [24], a large sample technique where a sample size of at least 200 is preferable [14]. For further guidance on building a teamwork instrument, readers can refer to the work of Marsicano et al. [17].

Team effectiveness x Team performance. The distinction between Team effectiveness and Team performance is highlighted in the work of Salas et al. [25]. Team performance is characterized as the outcome of a team's actions, irrespective of the approach employed to complete their task. In the context of software development, team performance encompasses meeting project goals, adhering to budgets and schedules, and delivering high-quality software. On the other hand, Team effectiveness is defined in a more comprehensive manner, encompassing how the team collaborates and interacts while accomplishing their tasks. This includes various team interactions, such as planning meetings, reviews, retrospectives, pair programming, and the use of coordination artifacts like iteration and product backlogs. In essence, team effectiveness considers not only the end result but also the dynamics and cooperation displayed during the task execution.

Team effectiveness models find frequent application in software engineering studies. Examples of such models include the Big Five model [25], which is utilized in various studies such as [6, 20, 30]), the Teamwork Quality model [13], featured in studies like [15, 22], and the Input-Process-Output (IPO) model [18], which is employed in studies like [19]. A comprehensive overview of these three models can be found in the work of Strode et al. [30].

In this work, we considered the TWQ instrument [13] as a comparative base because it has been extensively referenced in ASD [13, 8, 22, 15]. Also, we recognize that "teamwork quality" and "teamwork effectiveness" are closely related concepts that are commonly evaluated through measurable results [13, 33, 30, 22, 7]. Therefore, we refer to these concepts as "teamwork quality" or simply "teamwork".

2.2. Teamwork Instruments in ASD

This section summarizes the eight ASD teamwork instruments under study: TWQ instrument [13] (I1), Radar Plot instrument [21] (I2), ASTM [20] instrument (I3), TWQ-BN [8] instrument (I4), aTWQ) [22] instrument (I5), TACT [9] instrument (I6), ATEM [30] instrument (I7), and STEM [33] instrument (I8). Table 1 showcases a comprehensive list of all the teamwork instruments, along with the associated articles and their respective creation years.

TWQ - Teamwork Quality instrument (2001) [13]: Hoegl and Gemuenden [13] presented a comprehensive concept of collaboration in teams called Teamwork Quality (TWQ). This construct has six facets (i.e., Communication, Coordination, Balance of Member Contributions, Mutual Support, Effort, and Cohesion). Based on these facets and data collected in their study, the authors proposed a way for measuring the TWQ where the high order factor (i.e., TWQ) is the dependent variable, and the construct facets are the independent variable.

Radar Plot instrument (2009) [21]: Moe et al. [21] proposed an instrument that addresses key concerns and characteristics of agile teamwork and presents them along five dimensions: Shared Leadership, Team Orientation, Redundancy, Learning, and Autonomy. The instrument outputs a radar plot of the teamwork's status. To assess the teamwork's current status, it is necessary to answer a set of questions for each dimension and, based on these answers, assign a score on a scale from 0 to 10 for the dimension.

ASTM - A teamwork model for understanding an agile team: A case study of a Scrum project instrument (2010) [20]: Based on Dickinson and McIntyre's [4] teamwork model, Moe et al. [20] focused on the interrelations between essential teamwork components. Problems with team orientation, team leadership, and coordination, in addition to highly specialized skills and corresponding division of work, were important barriers to achieving team effectiveness.

TWQ-BN - Teamwork Quality Bayesian networks (2018) [8]: According to the agile principles and values, teamwork factors are critical to achieving success in agile projects. The TWQ-BN has a predicting and diagnosis purpose using Bayesian Networks. According to agile principles and values, teamwork factors are critical to achieving success in agile projects.

However, teamwork does not automatically arise. There are some existing instruments with the purpose of assessing the teamwork quality based on Structural Equation Modeling (i.e., empirically derived) and Radar Plot [21]. TWQ-BN instrument has 17 questions.

aTWQ - Agile Team Work Quality (2020) [22]: Based on Hoegl and Gemuenden's study [13] and a systematic literature review about challenges and success factors for large-scale agile transformations performed by Paasivaara et al. [5]. Poth et al. [22] derived the aTWQ at the initial team-level approach covering the following six factors: communication, coordination, balance of contribution, mutual support, effort, and cohesion. These six quality aspects lead to team performance [15], legitimating economically the effort for measurement and further TWQ improvement. They combined these aspects with those of TCI [2] and defined 19 related questions to develop a holistic team evaluation questionnaire for aTWQ [22].

TACT - An insTrument to Assess the organizational ClimaTe of agile teams (2020) [9]: TACT allowed for classifying the organizational climate of teams into the Communication, Collaboration, Leadership, Autonomy, Decision-Making, and Client Involvement dimensions. Some items were assessed negatively or neutrally, which represents points of attention. TACT captured the lack of agile ceremonies, the difficulty of the product owner in planning iterations, and the distance in leadership.

ATEM - Agile teamwork effectiveness model (2022) [30]: Teamwork is crucial in software development, particularly in agile development teams which are cross-functional and where team members work intensively together to develop a cohesive software solution. Effective teamwork is not easy; prior studies indicate challenges with communication, learning, prioritization, and leadership. Nevertheless, much advice is available for teams, from agile methods, practitioner literature, and general studies on teamwork to a growing body of empirical studies on teamwork in the specific context of ASD. The ATEM [30] model is based on evidence from focus groups, case studies, and multi-vocal literature and is a revision of a general Big Five [25] team effectiveness model. The ATEM [30] model comprises shared leadership, team mentoring, redundancy, adaptability, and peer feedback. Coordination mechanisms are needed to facilitate these components. Coordination mechanisms are shared mental models, communication, and mutual trust. ATEM instrument has 31 questions.

STEM - A Theory of Scrum Team Effectiveness Model (2022) [33]: The STEM model [33] proposes that the effectiveness of Scrum teams depends on five high-level factors - responsiveness, stakeholder concern, continuous improvement, team autonomy, and management support - and thirteen lower-level factors. The main finding is the interplay between stakeholder concern and responsiveness as drivers of agile team effectiveness. In turn, this requires a high degree of team autonomy, continuous improvement, and support from management.

2.3. Theoretical Evolution of Teamwork Instruments in ASD

Since the emergence of the TWQ instrument [13] in 2001, several other instruments have emerged in the literature. In this context, researchers have proposed instruments for assessing TWQ in the agile context, such as (i) a Radar Plot [21] that considers five dimensions for assessing TWQ (Shared Leadership, Orientation, Redundancy, Learning, and Autonomy); (ii) Moe et al. [20] used ASTM [20] that considers seven factors: Team orientation, Team leadership, Monitoring, Feedback, Backup, Coordination, and, Communication, (iii) a Structural Equation Model [15] (TWQ-SEM), based on a differentiated replication from Hoegl et al. [13], which considered that the teamwork construct is comprised of six variables: Communication, Coordination, Balance of Member Contribution, Mutual Support, Effort, and Cohesion. All the instruments mentioned are generic and cannot represent specific situations in the agile context because the instrument questions are not focused on agile terms.

Since 2018, specific instruments for ASD have emerged: a Bayesian networks-based model (TWQ-BN) [8] was developed based on the TWQ [13] instrument. The aTWQ instrument [22] was developed based on the TWQ [13] instrument. The ATEM instrument [30] was developed based on the Big Five theory [25]. The TACT instrument [9] was developed based on the TCI instrument [2]. The ATEM instrument [30] was developed to measure team effectiveness in the agile context. The STEM instrument [33] was developed considering some specific factors in Scrum. All the mentioned instruments have something in common: they have instrument questions directly associated with agile context situations. In this work, we named these instruments Specific Agile Teamwork Instruments because they are specific for ASD.

Based on this observation, we propose the classification of agile teamwork instruments into two groups: Generic teamwork instruments and Agile-based teamwork instruments. The generic ones were developed until 2018: TWQ, Radar Plot, and ASTM. The Agile-based ones were developed in 2018: TWQ-BN, aTWQ, TACT, ATEM, and STEM.

We found that, generally, the instruments are built and supported by a general theory in literature. Observing this, we created a *Level 1* in this architecture (Figure 1). As examples, we can cite the theories in Teamwork Literature: the Teamwork Quality Theory [13], the Team Climate Theory [2], the Big Five Theory [25] and The Group Development Theory [34]. In *Level 2*, the theories are combined with empirical research to build the instruments, as examples we have: the TWQ instrument [13], the TCI instrument [2], the GDQ instrument [34], the ATEM instrument [30], and the STEM instrument [33]. In *Level 3*, the theories and instruments are combined to build new specific ones. In the case of our study, for ASD, as examples we have: the TWQ-BN instrument [8], built taking as reference the TWQ instrument [13], the aTWQ instrument [22] taking as reference the instruments TWQ [13], TCI [2], and GDQ [34]. The TACT instrument [9] takes as reference the TCI instru-

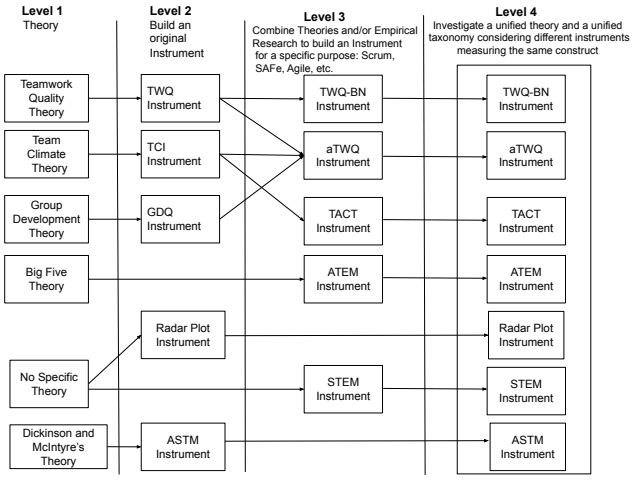

Figure 1: Evolution of Teamwork Instruments in ASD

ment [2]. Figure 1 illustrates the evolution of ASD Teamwork Instruments in ASD. In *Level 4*, we propose to investigate a unified theory and a unified taxonomy considering that we found seven ASD instruments measuring the same: the teamwork quality. Figure 1 depicts the Evolution of Agile Teamwork Instruments in ASD.

3. Study Configuration

This study presents a comprehensive study that aims to examine, compare, and synthesize eight specific instruments utilized for measuring Teamwork Quality (TWQ) in Agile Software Development (ASD). In what follows, we provide some details regarding research questions (Section 3.1). Then, we point out the research activities configuration (Section 3.2).

3.1. Research questions

We aimed to perform a quantitative and qualitative comparison between Teamwork Quality instruments in ASD, identifying trends in this comparison by focusing on the following research questions (RQs):

- **RQ1.** What is the quantitative relationship between Agile Teamwork instruments (factors and questions) and literature-based Agile Teamwork factors (themes)?
  - **RQ1.1** What are the factors used in teamwork instruments in ASD?
  - **RQ1.2** What are the most frequent factor names used in Teamwork instruments in ASD?
  - **RQ1.3** How are the frequencies of the instruments related to the work of Freire et al. [7]?
- **RQ2.** How can the Agile Teamwork instruments (factors and questions) be semantically compared in ASD?
  - **RQ2.1** What are the semantic relationships between the teamwork instrument factors in ASD?
  - **RQ2.2** What are the relationships between the evolution of teamwork instruments in ASD and the evolution of teamwork instrument factors' names and questions?

RQ1: The quantitative relationship between Agile Teamwork instruments and literature-based Agile Teamwork factors is a crucial aspect to explore in this study. Understanding how the factors and questions used in specific teamwork instruments align with established literature-based themes can shed light on the reliability and validity of these instruments. By answering RQ1.1, we can identify the factors commonly employed in teamwork instruments in Agile Software Development (ASD). This knowledge is essential as it provides a foundation for subsequent comparisons and allows researchers to focus on key aspects of teamwork assessment. RQ1.2 aims to pinpoint the most frequently utilized factor names in ASD instruments, which is valuable for understanding the prevalent themes and language employed by researchers in this field. Additionally, RQ1.3 investigates how the frequencies of instrument usage relate to the work of Freire et al. [7], a literature-based study. This comparison serves as an important validation step, enhancing the trustworthiness of the instruments' application in real-world contexts.

RQ2: The semantic comparison of Agile Teamwork instruments in ASD constitutes a fundamental aspect of this research. Semantic alignment between instrument factors and questions provides insights into the conceptual coherence and consistency of the instruments. RQ2.1 delves into the semantic relationships among teamwork instrument factors, revealing whether different instruments share common themes and concepts. This information helps researchers and practitioners in selecting the most appropriate instruments for specific assessment needs. Furthermore, RQ2.2 explores the connection between the evolution of teamwork instruments in ASD and the evolution of teamwork instrument factors' names and questions. This investigation offers valuable insights into how the instruments have evolved over time, potentially reflecting the changing nature of teamwork in the agile context. Understanding these relationships can inform future instrument development and enhance their relevance and effectiveness. By addressing RQ2, the study contributes to a deeper understanding of the nuances and intricacies of teamwork assessment, enabling researchers and practitioners to make informed decisions in their Agile Software Development projects.

3.2. Research Design

This study adopts a mixed-methods research design that combines both quantitative and qualitative approaches to achieve the research goals effectively. This dual approach enables a comprehensive exploration of the semantic relationships between specific TWQ instruments for ASD. By integrating quantitative and qualitative analyses, the study aims to provide a richer and more nuanced understanding of how these instruments are related and aligned within the agile context.

3.2.1. Data Collection

To ensure a comprehensive analysis, eight specific teamwork instruments tailored for ASD are selected (Table 1). The instruments are chosen based on their relevance and suitability to the agile context, considering their past usage and availability in the existing research literature. This rigorous selection process ensures that the chosen instruments represent the range of teamwork assessment tools applicable to ASD.

The next step involves data extraction from each selected teamwork instrument. Relevant data pertaining to the factors and questions used within each instrument is systematically gathered and organized for subsequent analysis. Additionally, literature-based Thematic Network themes identified by Freire et al. [7] are compiled to serve as a basis for comparison in the study. This addition enhances the study's depth by comparing instrument factors with established thematic themes.

3.2.2. Quantitative Analysis

In the quantitative phase, the factors present in each of the eight ASD instruments are mapped and systematically compared. This step aims to identify and highlight the most frequently utilized factors across the selected instruments. The quantitative analysis provides insights into the prevalence and significance of specific factors within the agile context.

The quantitative assessment delves further into examining instrument questions associated with each identified factor. Through frequency analysis, the study determines the prominence and prevalence of individual instrument questions for each factor. This in-depth examination helps ascertain the relative importance and weightage of different questions within the instruments.

Building on the literature-based Thematic Network themes identified by Freire et al. [7], the study compares with the factors extracted from the selected teamwork instruments. By aligning the identified factors with established thematic themes, the study seeks to identify potential overlaps, similarities, and divergences, providing a holistic view of the thematic representation within the instruments.

3.2.3. Qualitative Analysis

The qualitative phase adopts a semantic perspective to delve deeper into the context-specific characteristics of the analyzed teamwork instruments. This approach enables the identification of nuanced relationships and alignment among the instruments, considering their specificity within the agile context. The analysis also considers the theoretical underpinnings that support these instruments offering valuable insights into their semantic coherence and theoretical basis.

The qualitative analysis goes beyond descriptive exploration to identify emerging trends and patterns within the data. Drawing from the knowledge derived from the identified segments of agile instruments, the study investigates qualitative concepts that help uncover underlying themes and tendencies. This in-depth analysis supports the understanding of the semantic connections among the instruments and the contextual significance of specific factors.

3.2.4. Data Interpretation

The final phase of the study involves the interpretation of results obtained from both the quantitative and qualitative analyses. By collectively integrating the findings, the study gains a comprehensive understanding of the semantic relationships between the teamwork instruments in the context of Agile Software Development. This data interpretation stage provides valuable insights for researchers and practitioners, facilitating a coherent presentation of the findings in a manner that enhances their usability and applicability.

4. ASD instruments' factors (RQ1)

This section presents the factors we identified by analyzing the ASD instruments. It describes the computed frequencies of similar factors and discusses the results of a comparative analysis between the frequencies of ASD instrument factors and the teamwork thematic themes from the work of Freire et al. [7].

4.1. Factors of each ASD Instrument (RQ1.1)

This section provides a comprehensive description of the instrument factors identified in this study. Table 2 displays the relevant information, including the instrument name in the first column, the corresponding factor name in the second column, the symbol "#" denoting the number of questions associated with each factor in the third column, and the "Tot." representing the total number of questions for each instrument in the fourth column.

The **TWQ** instrument has six factors: Communication, Coordination, Balance of Member Contributions, Mutual Support, Effort, and Cohesion. The **Radar Plot** instrument has five factors: Shared Leadership, Team Orientation, Redundancy, Learning, and Autonomy. The **ASTM** instrument has seven factors: Team orientation, Team leadership, Monitoring, Feedback, Backup, Coordination, and Communication. The **TWQ-BN** has 17 factors: Teamwork, Team Autonomy, Cohesion, Collaboration, Self-Organizing, Coordination, Team Orientation, Communication, Daily Meetings, Team Distribution, Means of Communication, Monitoring, All Members Present, Personal Attributes, Expertise, Shared Leadership, and Team Learning. The **aTWQ** instrument has five factors: Participative safety, Support for Innovation, Vision, Task orientation, and Coordination. The **TACT** instrument has six factors: Communication, Collaboration, Leadership, Autonomy, Decision Making, and Client Involvement. The **ATEM** instrument has eight factors: Shared Mental Models, Mutual trust, Communication, Shared leadership, Peer feedback, Redundancy, Adaptability, and Team orientation. The **STEM** instrument has five factors and fourteen subfactors: Responsiveness (Refinement, Release Frequency), Stakeholder Concern (Stakeholder Collaboration, Shared Goals, Sprint Review Quality, Value Focus), Continuous Improvement (Shared Learning, Learning Environment, Psychologi-

Table 2
Teamwork Instrument Factors.

Instrum.	Factor	#	Tot.
TWQ	Communication	10	34
	Coordination	4	
	Bal.of Member Contribut.	3	
	Mutual Support	3	
	Effort	4	
	Cohesion	10	
Radar Plot	Shared Leadership	4	19
	Team Orientation	4	
	Redundancy	5	
	Learning	3	
	Autonomy	3	
ASTM	Team Orientation	2	14
	Team Leadership	2	
	Monitoring	2	
	Feedback	2	
	Backup	2	
	Coordination	2	
	Communication	2	
TWQ-BN	17 factors	17	17
aTWQ	Participative Safety	7	21
	Support for Innovation	5	
	Vision	4	
	Task Orientation	4	
	Coordination	1	
TACT	Communication	9	49
	Collaboration	7	
	Leadership	9	
	Autonomy	9	
	Decision Making	8	
	Client Involvement	7	
ATEM	Shared Mental Models	6	31
	Mutual Trust	3	
	Communication	3	
	Shared Leadership	8	
	Peer Feedback	2	
	Redundancy	3	
	Adaptability	3	
	Team Orientation	3	
STEM	Responsiveness	5	37
	Stakeholder Concern	10	
	Continuous Improvement	15	
	Team Autonomy	5	
	Management Support	2	

cal Safety, Quality, Sprint Retrospective Quality), Team Autonomy (Cross-Functionality, Self-Management), and Management Support (Management Support).

4.2. Frequency of similar factors in teamwork instruments (RQ1.2)

This section presents and analyzes the frequency of matches among the teamwork instruments. In the first step, we cross-referenced factors with identical names. For instance, Table 2 shows that both the Radar Plot instrument and the ASTM instrument have a factor named "Team Orientation." We calculated the frequency of matches for all instrument factors and presented this information in Table 3. In Column #F1 of Table 3, we listed the number of factors with the exact same name in each instrument. For example, in the "Team Autonomy" factor, a value of 1 in the TWQ-BN instrument indicates that it also has a factor named "Team Autonomy." On the other hand, Column #F2 represents cases where the factor names do not match exactly but convey the same meaning. For instance, the TACT instrument does not have an exact match for the "Team Autonomy" factor, but it does have a similar concept named "Autonomy." We accounted for this match in Column #F2. Finally, we combined the values from Column #F1 and Column #F2 into a Total column to determine that there were a total of 4 matches for the "Team Autonomy" factor.

As seen in Table 3, the "Communication" factor has 5 matches; the "Coordination", "Team Orientation", "Team Autonomy", and "Learning" factors have 4 matches; the "Collaboration", "Shared Leadership", and "Mutual Support" factors have 3 matches; the "Leadership" and "Redundancy" factors have 2 matches, and "Stakeholder Concern", "Continuous Improvement", "Team Autonomy", "Feedback", "Peer Feedback", and "Responsiveness" factors have only one match.

The factor that ranks highest with the most matches is "Communication." It secures the top position in the ranking and is present in five instruments: TWQ, ASTM, TWQ-BN, TACT, and ATEM, all of which include "Communication" within their variables. Following closely in the ranking is the "Coordination" factor with four matches. The instruments TWQ, ASTM, TWQ-BN, and aTWQ all feature a specific factor named "Coordination." As for the "Collaboration" factor, TWQ-BN, TACT, and STEM show varying degrees of matching. Specifically, TWQ-BN has one exact match, while TACT and STEM also exhibit matches. When it comes to the factors of "Shared Leadership", "Redundancy", "Feedback", and "Stakeholder Concern", STEM and ATEM present more matches than the other instruments.

When considering the frequency of occurrence among the instruments, it is worth noting that the STEM instrument stands out with the highest number of specific factors, including Team Autonomy, Continuous Improvement, Stakeholder Concern, and Responsiveness. On the other hand, the ATEM instrument features a unique factor, Peer Feedback. A comprehensive overview in Table 3 highlights the fact that only the ATEM and STEM instruments possess such distinctive factors. This observation suggests a trend towards employing more concrete factors aligned with agile-specific terminology. In contrast, the TWQ, ASTM, and TACT instruments exhibit a higher frequency of general factors, reflecting a prevalence of more generalized aspects.

We found instruments with different factor names but with the same meaning. In "Team Autonomy" factor, there is a "Autonomy" factor in the TACT instrument, a "Team Autonomy Cross-Functionality" and "Team Autonomy Self-Management" in the STEM instrument. All these questions are related to the "Team autonomy" concept but with different granularity. The STEM instrument has subfactors: Cross-Functionality and Self-Management for the "Team Autonomy" factor. In the "Learning" factor, the Radar-Plot in-

strument has a "Learning" factor. The TWQ-BN instrument has a "Team Learning" factor. The STEM instrument has the factors: "Continuous Improvement-Shared Learning" and "Continuous Improvement- Learning Environment". Note that in STEM instrument has specialized subfactors: "Shared Learning" and "Learning Environment", both related to "Continuous Improvement". In the "Collaboration" factor, the TWQ-BN and TACT instruments have the "Collaboration" factor. In STEM instruments, there is a "Stakeholder Concern-Stakeholder Collaboration" directly associated with Stakeholder Collaboration. In TWQ-BN and TACT instruments, "Collaboration" is associated with team collaboration. In the "Mutual Support" factor, The TWQ and ATEM instruments have the "Mutual Support" factor, both associated with team collaboration. In the STEM instrument, there is a factor named "Management Support" associated with the support from people in management positions.

4.3. Comparing ASD Instruments Frequencies with Freire et al. [7] teamwork thematic network (RQ1.3)

Freire et al. [7] presented a literature-based Thematic Network identifying themes and codes shown in Table 4. For example, the theme "Coordination" is related to the codes "Coordination", "Performance Monitoring", "Task Novelty" and "Familiarity". Table 4 shows that the most frequent theme in the agile teamwork literature is "Team Orientation" with 22 matches, followed by "Coordination" with 16 matches. The third most frequent is "Expertise" with 15 matches, and so on. By comparing the results shown in Table 4 and Table 3, we identified an important trend: "Team Orientation" and "Coordination" are in the top 3 ranking in both, Freire et al.'s work and our work, indicating that they are key factors for agile teamwork quality.

5. Semantic Comparison between ASD Instruments (RQ2)

This section discusses the results of the semantic comparison between the instruments (Section 5.1) and investigates the relationship between the evolution of teamwork instruments in ASD and the evolution of teamwork instruments factors names and questions (Section 5.2).

5.1. Semantic relationship between teamwork instruments factors in ASD (RQ2.1)

As previously discussed, Table 3 shows the frequency in which we identified each of Freire et al.'s factors in the agile teamwork instruments under study. For example, we identified the factor "Communication" in five instruments: TWQ, ASTM, TWQ-BN, TACT, and ATEM. This result indicates that such instruments are similar in terms of containing questions related to such a factor. To address RQ2.1, we went beyond and performed qualitative analysis on the questions of each instrument that were mapped to such a factor to consider semantic aspects and have a more in-depth comparison between the instruments under study.

Table 3
Frequency in each Instruments Factors.

Instrum. Factor	Instrum.	#F1	#F2	Tot.
Communication	TWQ	1	0	5
	ASTM	1	0	
	TWQ-BN	1	0	
	TACT	1	0	
	ATEM	1	0	
Coordination	TWQ	1	0	4
	ASTM	1	0	
	TWQ-BN	1	0	
	aTWQ	1	0	
	ATEM	0	0	
Team Orientation	Radar Plot	1	0	4
	ASTM	1	0	
	TWQ-BN	1	0	
	ATEM	1	0	
Team Autonomy	TWQ-BN	1	0	4
	TACT	0	1	
	STEM	2	0	
Learning	Radar Plot	1	0	4
	TWQ-BN	0	1	
	STEM	0	2	
Collaboration	TWQ-BN	1	0	3
	TACT	1	0	
	STEM	0	1	
Shared Leadership	Radar-Plot	1	0	3
	TWQ-BN	1	0	
	ATEM	1	0	
Mutual Support	TWQ	1	0	3
	ATEM	1	0	
	STEM	0	1	
Leadership	TACT	1	0	2
	ASTM	1	0	
Redundancy	Radar Plot	1	0	2
	ATEM	1	0	
Stakeholder Concern	STEM	1	0	1
Continuous Improvement	STEM	1	0	1
Feedback	ASTM	1	0	1
Peer Feedback	ATEM	1	0	1
Responsiveness	STEM	1	0	1

5.1.1. Communication

For the "Communication" factor, we compared TWQ, ASTM, TWQ-BN, TACT, and ATEM. Ten questions from Instrument 1 (TWQ) focus on team communication (Questions 1, 2, 3, 4, 5, 6, 7, 8, 9, 10), while Question 13 from Instrument 3 (ASTM) is related to verifying information before making a report. Effective communication and information exchange are also implicit in the questions from Instrument 6 (TACT) and Instrument 4 (TWQ-BN), as they inquire about freely talking, updating lists, listening to opinions, and knowing team members' skills. By analyzing these questions, we identified the following themes:

Openness and Transparency: Three Questions from Instrument 1 (TWQ) address the openness of communication (Questions 5, 6, 7), and Instrument 6 (TACT) emphasizes

Table 4
ASD Theme frequencies in Freire et al. work

ASD Theme	ASD Code	#	Tot.
Team Orientation	Orientation	7	22
	Value Diversity	1	
	Goals	2	
	Roles	2	
	Holistic Team Involvement	1	
	Experience in the Organi.	1	
	Trust	5	
	Motivation	1	
	Norms	2	
Coordination	Coordination	5	16
	Performance Monitoring	9	
	Task Novelty	1	
	Familiarity	1	
Expertise	Tools knowledge	2	15
	Collective Knowledge	4	
	Adequate Skills	1	
	Redundancy	7	
	Experience with Work	1	
Management Mechanisms	Management	4	10
	Planning	1	
	Discussion	1	
	Implementation	1	
	Evaluation	1	
	Information Radiators	1	
	Decision Making	1	
Shared Leadership	Shared Leadership	8	9
	Formal Leadership	1	
Communication	Communication	9	9
Organization Culture	Culture	4	8
	Structure	1	
	Team Size	2	
	Organization Support	1	
Collaboration	Interdependence	1	8
	Collaboration	7	
Learning	Learning	8	8
Members Personality	Individual Differences	1	5
	Heterogeneity	1	
	Personality	3	
Team Autonomy	Autonomy	4	5
	Task Control	1	
Feedback	Awareness	1	5
	Acceptance	1	
	Feedback	3	
Cohesion	Cohesion	3	3

openness in freely talking about difficulties (Question 1).

Team Interaction and Understanding: Question 6 from Instrument 1 (TWQ) and Question 6 from Instrument 6 (TACT) both relate to understanding team members' skills and expertise and using them appropriately.

Project Progress and Information: Questions from Instrument 7 (ATEM) focus on project progress and information visualization (Questions 10, 11, 12), while Instrument 6 (TACT) has a question related to knowing project problems and team difficulties through daily meetings (Question 7).

Information Accuracy and Precision: Question 9 from Instrument 1 (TWQ) and Question 9 from Instrument 6 (TACT) inquire about the precision and scope of information received.

5.1.2. Coordination

For the "Coordination" factor, we compared questions from the following instruments: TWQ, ASTM, TWQ-BN, aTWQ, and ATEM. Next, we present our results grouped by the main themes identified while analyzing the questions.

Task Coordination: Questions from Instrument 1 (TWQ) and Instrument 5 (aTWQ) both focus on task coordination and harmonization: Question 11 from TWQ: "The work done on subtasks within the project was closely harmonized". Question 21 from aTWQ: "Is there a common understanding when working on parallel subtasks and agreement on common work breakdown structures, schedules, budgets, and deliverables?".

Clarity and Acceptance of Goals: Instrument 1 (TWQ) and Instrument 3 (ASTM) include questions related to goal clarity and acceptance: Question 12 from TWQ: "There were clear and fully comprehended goals for subtasks within our team". Question 11 from ASTM: "Passing performance-relevant data to other members efficiently". Question 12 from ASTM: "Facilitating the performance of other members' jobs".

Synchronization and Integration of Tasks: Instrument 4 (TWQ-BN) has a question that relates to the synchronous and integrated execution of tasks: Question 6 from TWQ-BN: "The team executes its tasks in a synchronous and integrated manner".

Conflict and Diverging Interests: Instrument 1 (TWQ) includes a question about conflicting interests regarding subtasks/subgoals: Question 14 from TWQ: "There were conflicting interests in our team regarding subtasks/subgoals". These are some of the semantic similarities between the questions from the different instruments. The themes of task coordination, goal clarity and acceptance, task synchronization, and conflict are present in the questions.

5.1.3. Team Orientation

For the "Team Orientation" factor, we compared RadarPlot, ASTM, TWQ-BN, and ATEM. By analyzing these instruments questions, we identified the following themes:

Valuing and Considering Alternative Suggestions: Questions from Instrument 2 (Radar Plot - Team Orientation) and Instrument 7 (ATEM-TC-Team Orientation) focus on how the team values and considers alternative suggestions: Question 5 from Radar Plot: How does the team take into account alternative suggestions in team discussions? Question 6 from Radar Plot: How does the team value alternative suggestions? Question 29 from ATEM: "Taking into account alternative solutions provided by teammates and appraising that input to determine what is most correct".

Participation and Commitment to Team Goals: Instrument 3 (ASTM - Team Orientation) and Instrument 7 (ATEM-TC-Team Orientation) include questions related to team member participation and commitment to team goals: Question 1 from ASTM: assigning a high priority to team goals. Question 2 from ASTM: Participate willingly in all relevant as-

pects of the team. Question 30 from ATEM-TC: Increased task involvement, information sharing, strategizing, and participatory goal setting.

Trust and Collaboration: Question 7 from Instrument 4 (TWQ-BN - Team Orientation) and Question 31 from Instrument 7 (ATEM-TC-Team Orientation) touch on trust and collaboration within the team. Question 7 from TWQ-BN: The team members trust each other and feel motivated to work together to achieve the team's goals. Question 31 from ATEM-TC: The team sticks together and remains united.

Task and Individual Relations: Instrument 2 (Radar Plot - Team Orientation) includes questions that inquire about the relationship between team members and their tasks. Question 7 from Radar Plot: How do team members relate to the tasks of individuals? Question 8 from Radar Plot: What kind of ownership do the team members have to the project? These are some of the semantic similarities between the questions from the different instruments. The themes of valuing alternative suggestions, participation in team goals, trust, collaboration, and task relations are present in the questions.

5.1.4. Team Autonomy

For the "Team Autonomy" factor, we compared TWQ-BN, TACT, and STEM and identified the following themes:

Autonomy in Decision Making and Planning: Questions from Instrument 6 (TACT - Autonomy) and Instrument 8 (STEM - Team Autonomy) both focus on autonomy in decision-making and planning: Question 28 from TACT: In the current project, I can choose the tasks I want to execute in the iteration. Question 34 from TACT: My team has the decision authority and responsibility to plan the iteration. Question 36 from STEM: Most people in this team have the ability to solve the problems that come up in their work. Question 38 from STEM: This team has control over the scheduling of teamwork.

Autonomy in Technical Solutions: Instrument 6 (TACT - Autonomy) and Instrument 4 (TWQ-BN - Team Autonomy) have questions related to autonomy in technical solutions: Question 30 from TACT: In this organization, we can suggest changing the team's software process development. Question 33 from TACT: My team can communicate with the product owner and other relevant stakeholders. Question 2 from TWQ-BN: No external agent is interfering with how the team executes its tasks. The external agent collaborates with them to define what will be.

Protection of Team Autonomy: Instrument 6 (TACT - Autonomy) includes a question about the team facilitator protecting the team's autonomy from external interferences: Question 29 from TACT: "In the current project, the team facilitator protects the team autonomy from external interferences". These are some of the semantic similarities between the questions from the different instruments. The themes of autonomy in decision-making, planning, technical solutions, communication, and protection of team autonomy are present in the questions but keep in mind that this analysis is based on the questions provided, and there may be other connections and interpretations depending on the specific usage and context of these instruments.

5.1.5. Learning

For the "Learning" factor, we compared Radar-Plot, TWQ-BN, and STEM and identified the following themes:

Learning and Improvement in Software Development: Questions from Instrument 2 (Radar Plot - Learning) directly relate to learning and improvement in software development: [14] from Radar Plot: What are the arenas where you give feedback on each other's work? [15] from Radar Plot: "How are software development problems identified, and do you improve the development method?" [16] from Radar Plot: Do you keep what works well in your development process? [17] from Radar Plot: "How are artifacts in the development process (burndown chart, backlog, daily meetings, sprint reviews, and retrospectives) used to learn?"

Team Learning and Adaptation: Instrument 4 (TWQ-BN - Team Learning) has a question related to team learning and adaptation: [17] from TWQ-BN: The team adapts itself to changes in the team environment and adjusts the strategies as needed.

Shared Learning and Collaboration: Instrument 8 (STEM - Continuous Improvement - Shared Learning) includes questions related to shared learning and collaboration with other teams: [21] from STEM (Continuous Improvement - Shared Learning): This team frequently works with other groups or teams to solve shared problems; [22] from STEM (Continuous Improvement - Shared Learning): Teams in this organization share what they learn with other teams; [23] from STEM (Continuous Improvement - Shared Learning): Members of this team frequently meet with other teams to identify improvements.

Learning Environment and Support for Learning: Instrument 8 (STEM - Continuous Improvement - Learning Environment) also has questions related to the learning environment and support for learning: [24] from STEM (Continuous Improvement - Learning Environment): In and around this team, people are given time to support learning; [25] from STEM (Continuous Improvement - Learning Environment): In and around this team, people are rewarded for learning.

The Radar Plot questions focus on aspects of software development processes and feedback mechanisms, while the STEM questions explore how teams collaborate, share knowledge, and support learning. The TWQ-BN question touches on the team's adaptability and strategy adjustments in response to changes in the team environment.

5.1.6. Collaboration

For the "Collaboration" factor, we compared TWQ-BN, TACT, and STEM. Question 4 (TWQ) and Questions 10, 11, 12, 13, 14, 15, and 16 (TACT) all revolve around teamwork, collaboration, and how team members work together to achieve common goals. As a result of analyzing such questions, we identified the following themes:

Project Development: Question 4 (TWQ) talks about success on project development, and some questions from

TACT (e.g., Questions 14, 15, 16) mention specific aspects related to projects, such as project-related decisions, problem analysis, and software design.

Team Support: Questions 4 (TWQ) and Questions 11, 12, and 13 (TACT) highlight the aspect of team members helping each other and providing support whenever needed.

Knowledge Sharing: Question 10 (TACT) indicates team members' consideration of sharing know-how with each other, which might be related to the collaboration and success mentioned in Question 4 (TWQ).

Semantic Similarities for Collaboration: The questions from TACT (10, 11, 12, 13, 14, 15, and 16) are all related to different aspects of team collaboration. They cover topics like knowledge sharing, mutual support, efficient teamwork, consistent decision-making, problem analysis, and software design based on user stories. These factors indicate a strong emphasis on collaboration and teamwork within the team.

Semantic Similarities for Stakeholder Concern - Stakeholder Collaboration: The questions from STEM (11, 12, and 13) all revolve around the team's interactions with stakeholders, users, and customers. They suggest a high level of engagement and collaboration between the team and external parties. These factors indicate that the team is attentive to stakeholder needs and actively seeks their input and collaboration.

Overall, the semantic similarities between the questions can be summarized as follows: TWQ-BN and TACT instruments focus on collaboration within the team. TWQ-BN specifically mentions "a high degree of collaboration", while TACT addresses various collaboration aspects like knowledge sharing, support, efficient teamwork, and decision-making. The STEM instrument, on the other hand, emphasizes stakeholder concern and collaboration. It highlights the team's interactions with stakeholders, customers, and users, indicating a strong focus on understanding and meeting their needs. In conclusion, the instruments TWQ-BN, TACT, and STEM all share the theme of collaboration, but they approach it from different angles. TWQ-BN emphasizes collaboration within the team, while TACT covers various aspects of team collaboration. STEM, on the other hand, emphasizes stakeholder concern and the team's collaboration with external parties.

5.1.7. Shared Leadership

For the "Shared Leadership" factor, we compared Radar-Plot, TWQ-BN, and ATEM. After analyzing these instruments' questions, we identified the following themes:

Decision-Making and Empowerment: Questions from Instrument 1 (Radar Plot - Shared Leadership) and Instrument 2 (ASTM - Team Leadership) focus on decision-making and empowerment within the team: Question 1 from Radar Plot: Is everyone involved in the decision-making process? Question 2 from Radar Plot: "Do team members make important decisions without consulting other team members?" Question 3 from ASTM: explaining to other team members exactly what is needed from them during an assignment. Question 4 from ASTM: listening to the concerns of other team members. Shared Decision Authority and Leadership: Instrument 3 (TWQ-BN - Shared Leadership) has a question related to shared decision authority and leadership: Question 16 from TWQ-BN: The decision authority and leadership are shared.

Team Facilitation: Instrument 4 (TACT - Leadership) focuses on team facilitation and the role of a team facilitator: Questions 17 to 25 from TACT include various aspects of team facilitation, such as providing helpful feedback, eliminating barriers, listening to team ideas and concerns, discussing team problems, protecting the team from outside interference, helping the team acknowledge and solve disagreements, and assisting in understanding iteration objectives.

Agile Team Practices: Instrument 5 (ATEM-TC-Shared Leadership) is centered around agile team practices and servant leadership: Questions 13 to 20 from ATEM-TC focus on various aspects of agile team practices, such as team problem-solving, determining performance expectations, and interaction patterns, synchronizing and combining individual contributions using agile practices and automated tools, seeking and evaluating information affecting team functioning, determining team member roles based on agile values and methodologies, determining the frequency and type of preparatory meetings and feedback sessions, and the role of a servant leader in facilitating a boundary-spanning function. These are some of the semantic similarities between the questions from the different instruments. The themes of decision-making, empowerment, shared leadership, team facilitation, and agile practices are present in the questions.

5.1.8. Mutual Support

For the "Mutual Support" factor, we compared TWQ, ATEM, and STEM. We identified the following questions related to this factor: [18] TWQ: "The team members helped and supported each other as best they could."; [19] TWQ: "If conflicts came up, they were easily and quickly resolved."; [20] TWQ: "Discussions and controversies were conducted constructively."; [7] ATEM-TCM: "Mutual trust - Information sharing."; [8] ATEM-TCM: "Mutual trust - Willingness to admit mistakes and accept feedback."; [9] ATEM-TCM: "Mutual trust - Supportive team social climate."

Mutual support and Trust: Questions [18], [19], and [20] from TWQ and Questions [7], [8], and [9] from ATEM, all address different aspects of mutual support and trust within the team. TWQ focuses on supporting each other, resolving conflicts, and constructive discussions, while ATEM highlights mutual trust through information sharing, feedback acceptance, and a supportive social climate.

Management support: STEM contains questions related to management support: [41] STEM: "People in a management position generally understand why this team works with Scrum."; [42] STEM: "People in a management position help this team work with Scrum.". Questions [41] and [42] from STEM, both pertain to management support in the context of the team working with Scrum. They suggest that people in a management position know the team's utilization of Scrum

and provide assistance in this regard.

Overall, TWQ (Instrument 1) and ATEM (Instrument 7) emphasize aspects of mutual support and trust within the team. While TWQ addresses support, conflict resolution, and constructive discussions, ATEM focuses on information sharing, feedback acceptance, and a supportive team social climate. STEM (Instrument 8) questions center around management support, particularly regarding the team's use of Scrum.

5.1.9. Leadership

In the "Leadership" factor, we compared TACT and ASTM. Based on the questions provided by Instrument 3 (ASTM) and Instrument 6 (TACT), we identified s emantic similarities in the next questions. Team Leadership: [3] ASTM: "Explaining to other team members exactly what is needed from them during an assignment."[4] ASTM: "Listening to the concerns of other team members." [17] TACT: "In the current project, the team, the product owner, and the team facilitator work excellently together to plan the iteration." [18] TACT: "The team facilitator gives me helpful feedback on how to be more effective." [19] TACT: "The team facilitator eliminates barriers, encourages, and facilitates the use of agile methods." [20] TACT: "The team facilitator listens to my ideas and concerns." [21] TACT: "The team facilitator discusses the problems of the team." [22] TACT: "The team facilitator protects the team from outside interference." [23] TACT: "The team facilitator helps my team to acknowledge and solve our disagreements." [24] TACT: "The team facilitator assists in understanding whether the iteration objectives are clear and whether the team agrees with these objectives." [25] TACT: "The team facilitator gives the team helpful feedback on how to be more agile."

Both ASTM (3 and 4 questions) and TACT (17 to 25 question) instruments include questions related to team leadership. ASTM focuses on team leadership involving explaining assignments clearly and listening to team members' concerns. TACT addresses leadership in the context of the team facilitator's role and their collaboration with the team and product owner. The TACT questions highlight various aspects of effective l eadership, s uch a s p roviding feedback, encouraging agile methods, protecting the team, resolving disagreements, and promoting agility.

Overall, ASTM (Instrument 3) and TACT (Instrument 6) have questions related to team leadership. ASTM focuses on leadership involving task explanation and listening to concerns, while TACT addresses leadership in the context of the team facilitator's role and their impact on the team's performance, collaboration, and agile practices.

In conclusion, the instruments ASTM and TACT touch on different a spects o f t eam l eadership. A STM addresses leadership in terms of task communication and listening, while TACT emphasizes the team facilitator's role and their influence on team dynamics, problem-solving, and agile practices.

5.1.10. Redundancy

For the "Redundancy" factor, we compared Radar Plot and ATEM. Based on the questions provided by Instrument 2 (Radar-Plot) and Instrument 7 (ATEM), we seek to identify the semantic similarities between them: Redundancy: [9] Radar-Plot: "How easy is it to complete someone else's task?"; [10] Radar-Plot: "If you are stuck, do you get help?"; [11] Radar-Plot: "Do you help others when they have problems?"; [12] Radar-Plot: "How are tasks allocated?"; [13] Radar-Plot: "If someone leaves the team, is it easy to substitute this person?"; [23] ATEM-TC: "Recognition by potential backup providers that there is a workload distribution problem in their team."; [24] ATEM-TC: "Shifting of work responsibilities to underutilized team members."; [25] ATEM-TC: "Completion of the whole task or parts of tasks by other team members." Semantic Similarities for Redundancy: The questions from both Radar-Plot (9 to 13) and ATEM-TC (23 to 25) instruments touch on the concept of redundancy within the team. Radar-Plot questions focus on how easy it is to complete each other's tasks, provide help, and allocate tasks. They also inquire about the ease of substituting team members if needed. On the other hand, ATEM questions address redundancy in terms of recognizing workload distribution issues, shifting work responsibilities, and task completion by other team members.

Overall, both Radar-Plot (Instrument 2) and ATEM (Instrument 7) have questions related to redundancy within the team. Radar-Plot addresses the ease of completing tasks, providing help, task allocation, and substitution of team members. ATEM questions highlight redundancy in terms of recognizing workload issues, shifting responsibilities, and task completion by other team members. In conclusion, the instruments Radar-Plot and ATEM touch on different aspects of redundancy within the team. Radar-Plot addresses the ease of task completion and support, while ATEM emphasizes workload distribution, task shifting, and task completion by various team members.

5.1.11. Stakeholder Concern

For the "Stakeholder Concern" factor, we investigated the STEM instrument. Based on the questions provided by Instrument 8 (STEM), we identified the semantic similarities between the following questions: Stakeholder Collaboration: [11] STEM: "Members of this team frequently meet with users or customers of what this team creates."; [12] STEM: "People from this team often invite or visit people that use what this team works on."; [13] STEM: "People in this team closely collaborate with users, customers, and other stakeholders."; Shared Goals: [14] STEM: "This team generally has clear Sprint Goals."; [15] STEM: "During Sprint Planning, this team formulates a clear goal for the Sprint."; Sprint Review Quality: [16] STEM: "The Product Owner of this team uses the Sprint Review to collect feedback from stakeholders."; [17] STEM: "During Sprint Reviews, stakeholders frequently try out what this team has been working on during the Sprint." Value Focus: [18] STEM: "The Product Owner of this team has a clear vision for the product.";

[19] STEM: "The Product Backlog of this team is ordered with a strategy in mind."; [20] STEM: "Everyone in this team is familiar with the vision for the product.";

Further, STEM's questions [16] and [17] are related to the quality of Sprint Reviews. They discuss the involvement of stakeholders in providing feedback and trying out the team's work during the Sprint Review, indicating a focus on gathering valuable input from stakeholders. Value Focus: Questions [18], [19], and [20] relate to the team's value focus. They touch on aspects such as the Product Owner having a clear vision for the product, the strategic ordering of the Product Backlog, and everyone in the team being familiar with the product's vision. These questions suggest a strong orientation toward delivering value to stakeholders.

The questions from Instrument 8 (STEM) can be grouped into several categories based on their similarities: Stakeholder Collaboration: Questions [11], [12], and [13] all pertain to stakeholder collaboration. They highlight the team's frequent interactions with users, customers, and other stakeholders, focusing on engaging and working closely with them. Shared Goals: Questions [14] and [15] revolve around shared goals. They address the team's clarity on Sprint Goals and the formulation of clear goals during Sprint Planning, which indicates a strong emphasis on having well-defined objectives.

Overall, STEM (Instrument 8) questions address stakeholder collaboration, shared goals, sprint review quality, and value focus. The instrument focuses on actively involving stakeholders, defining clear goals, obtaining valuable feedback during reviews, and delivering value through a well-defined product vision and ordered b acklog. In conclusion, the instrument STEM (Instrument 8) focuses on various aspects of stakeholder engagement, goal-setting, review quality, and value-driven development, all contributing to effective project execution and successful product delivery.

5.1.12. Continuous Improvement

For the "Continuous Improvement" factor, we investigated the STEM instrument. Based on the questions provided by Instrument 8 (STEM), we identified the following themes:

Shared Learning: [20] STEM: "This team frequently works with other groups or teams to solve shared problems."; [21] STEM: "Teams in this organization share what they learn with other teams."; [22] STEM: "Members from this team frequently meet with other teams to identify improvements."; Continuous Improvement - Learning Environment: [23] STEM: "In and around this team, people are given time to support learning."; [24] STEM: "In and around this team, people are rewarded for learning.";

Psychological Safety: [25] STEM: "In and around this team, people give open and honest feedback to each other."; [26] STEM: "In and around this team, people listen to others' views before speaking."; [27] STEM: "In and around this team, whenever people state their view, they also ask what others think."; [28] STEM: "In and around this team, people openly discuss mistakes to learn from them."; [29] STEM: "In and around this team, people help each other learn.".

Quality: [30] STEM: "Members of this team have a shared understanding of what quality means to them."; [31] STEM: "People in this team frequently talk about quality and how to improve it.";

Sprint Retrospective Quality: [32] STEM: "The Sprint Retrospectives of this team generally result in at least one useful improvement."; [33] STEM: "During Sprint Retrospectives, this team openly discusses improvements." Semantic Similarities: The questions from Instrument 8 (STEM) can be grouped into several categories based on their similarities.

Shared Learning: Questions [20], [21], and [22] all focus on shared learning and collaboration. They highlight how the team works with other groups or teams, shares knowledge within the organization, and engages in cross-team meetings to identify improvements. Continuous Improvement - Learning Environment: Questions [23] and [24] pertain to the learning environment. They address the provision of time and rewards for supporting learning, which fosters a culture of continuous improvement.

Psychological Safety: Questions [25] to [29] all relate to psychological safety. They emphasize the importance of open and honest feedback, active listening, inviting others' views, openly discussing mistakes, and helping each other learn. Continuous Improvement - Quality: Questions [30] and [31] are related to the team's understanding of quality and how they frequently discuss it and work to improve it.

Sprint Retrospective Quality: Questions [32] and [33] focus on the quality of Sprint Retrospectives. They mention the usefulness of improvements resulting from these retrospectives and the team's open discussions during them.

Overall, STEM (Instrument 8) questions address various aspects of continuous improvement. They cover shared learning and collaboration with other teams, creating a supportive learning environment, fostering psychological safety for open communication, discussing quality improvements, and the effectiveness of Sprint Retrospectives in generating useful insights.

In conclusion, the instrument STEM (Instrument 8) highlights different dimensions of continuous improvement within the team, encompassing shared learning, learning environment, psychological safety, quality discussions, and Sprint Retrospective effectiveness. These factors collectively contribute to the team's ability to continuously learn, evolve, and deliver value.

5.1.13. Feedback and Peer Feedback

For the "Feedback" factor, we investigated the ASTM instrument. For the "Peer Feedback" factor, we investigated the ATEM instrument. Based on the questions provided by Instrument 3 (ASTM) and Instrument 8 (STEM), we identified the following questions: [7] ASTM: "Responding to other members' requests for information about their performance."; [8] ASTM: "Accepting time-saving suggestions offered by other team members."; [21] ATEM-TC: "Identifying mistakes and lapses in other team members' actions.";

[22] ATEM-TC: "Regular feedback regarding team member actions to facilitate self-correction."

The questions from ASTM (7 and 8) and ATEM-TC (21 and 22) instruments focus on different aspects of feedback within the team: ASTM questions emphasize the exchange of feedback between team members. Question 7 addresses how team members respond to requests for performance-related information, while Question 8 focuses on their receptiveness to time-saving suggestions provided by others. ATEM-TC questions focus on peer feedback within the team. Question 21 mentions identifying mistakes and lapses in other team members' actions, indicating a form of feedback that helps in recognizing areas for improvement. Question 22 highlights the importance of regular feedback to facilitate self-correction, suggesting an ongoing feedback loop to enhance team performance.

Overall, ASTM (Instrument 3) and ATEM (Instrument 8) have questions related to feedback within the team. ASTM focuses on responding to information requests and accepting suggestions, while ATEM emphasizes the identification of mistakes, providing regular feedback, and facilitating self-correction.

In conclusion, the instruments ASTM and ATEM address different aspects of feedback within the t eam. ASTM highlights feedback exchange and acceptance of suggestions, while ATEM focuses on peer feedback for recognizing errors and supporting ongoing improvement through regular feedback.

5.1.14. Responsiveness

In the "Responsiveness" factor, we analyzed the STEM instrument (Instrument 8). We intended to identify the semantic similarities within the provided questions. In the domain of Responsiveness and Refinement, the questions were: [6] STEM: "The team's Sprint Backlog typically comprises numerous small items."; [7] STEM: "This team allocates time during the Sprint to elaborate on the work slated for the succeeding Sprints."; and [8] STEM: "Throughout the Sprint, this team commits time to decompose work for upcoming Sprints."

Regarding Responsiveness and Release Frequency, the questions were: [9] STEM: "The bulk of this team's Sprints lead to software that is prepared for deployment to production."; [10] STEM: "For this team, the majority of Sprints culminate in an increment ready for user release."

For Responsiveness and Refinement, questions [6], [7], and [8] collectively denote the team's adaptability and responsiveness. They underline the team's approach of maintaining a Sprint Backlog with numerous smaller items and dedicating time within the Sprint to clarify and decompose work for future Sprints. Responsiveness and Release Frequency are addressed in questions [9] and [10], where the focus is on the team's aptitude to regularly produce software or increments that can be released to users, showcasing the team's capacity to deliver value frequently.

In essence, the STEM instrument (Instrument 8) queries examine various facets of responsiveness. Questions related to refinement underscore the team's competency in decomposing and elucidating work during the Sprint, enabling adaptability. Questions associated with release frequency emphasize the team's consistent delivery of software or increments prepared for deployment or user release.

In summary, Instrument 8 (STEM) emphasizes distinct aspects of responsiveness, including refinement practices that foster adaptability and the team's ability to deliver valuable software or increments regularly. These factors collectively enhance the team's agility and capacity to deliver user value.

5.2. Relationship between the evolution of teamwork instruments in ASD and evolution of teamwork instruments factors names and questions (RQ2.2)

Given the results presented in Section 5, we found that the instruments ATEM, STEM, aTWQ and TWQ-BN brought new concepts directly associated with the agile context, among them: daily meetings, retrospective meetings, and Sprint Review. STEM brought other concepts like Cross-Functionality and Self-Management associated with Team Autonomy.

We suggest classifying agile teamwork instruments into two groups: Generic teamwork instruments and Agile-based teamwork instruments. The generic ones were developed until 2018: TWQ, Radar Plot, and ASTM. The Agile-based ones were developed later: TWQ-BN, aTWQ, TACT, ATEM, and STEM. We noted that the factors and questions from the Agile-based one included a terminology closely related to agile concepts. Further, ATEM (with seven factors) and STEM (with five factors and 14 subfactors) present a trend toward increasing the number of factors and subfactors compared to the older instruments.

6. Discussion and Findings

This section discusses this study's research questions and the trends observed. In summary, we mapped the factors of the eight teamwork instruments, then we compared them with the Themes found by Freire et al. [7]. The objective was to understand how the themes and instrument questions are quantitatively related. Then, we intended to identify trends in these factors. Considering the Themes analysis in Section 4.3. The results showed that Team Orientation and Coordination were identified among the top three rankings, both in the frequency of instrument questions and the frequencies of literature-based Thematic Network developed in Freire et.al [7].

We found in our semantic analysis important themes associated a many instrument factors. In Communication we found the themes: Openness and Transparency, Team Interaction and Understanding, Project Progress and Information, Information Accuracy and Precision. In Coordination factor, we found: Task Coordination, Clarity and Acceptance of Goals, Synchronization and Integration of Tasks, etc. In Team Orientation we found: Valuing and Considering Alternative Suggestions, Participation and Commitment to Team Goals, Trust and Collaboration, Task and Individ-

ual Relations, etc. In Collaboration we found: Team Support, Knowledge Sharing, etc. In Mutual Support, we found: Trust, and Management support. The present study can be a starting point for the development of new studies exploring the relationships between the instruments' factors and the themes identified in this study.

The researchers could investigate whether lower frequencies are, in fact, less important for teamwork quality. In this way, researchers will already know which subparts of the instruments to use. It was found the frequency of appearance of each factor related to the teamwork quality and the number of corresponding questions for each instrument. With this knowledge, this work can support other works that need to use ASD teamwork instruments for a specific p urpose. For example, if a researcher needs to investigate the relationship between Communication and Shared Leadership in a company, he can choose specific A SD i nstruments: F or Communication (TWQ, ASTM, TWQ-BN, TACT, and ATEM) and Shared Leadership (Radar Plot, TWQ-BN, and ATEM) in the investigation based on the requirements. Qualitative concepts can be investigated in future works focusing on investigating the ASD factors from the knowledge of the identified parts of the agile instruments.

This study can support using a Teamwork Instrument for a specific p urpose. For example, if a researcher needs to investigate the relationship between Feedback and Team Autonomy, he can choose what parts of the instruments to use. This work highlights that the ASD literature codes: Task Control, Communication, Coordination, and Team Autonomy are the most used in ASD Teamwork Instruments. This is an important result, as it confirms t hat t he factors identified by Freire et al. [7] are, in fact, those that are being used more frequently in specific A SD i nstruments, w hich were developed based on strong literature theories and empirical studies. Additionally, we identified a nd c ompared t he referred questions in the eight ASD instruments analyzed in this work. We noted that finding a standard terminology for ASD Teamwork factors remains challenging, and there is a need for further investigation into this area. Finally, practitioners can benefit from the study's findings by better understanding the importance of Teamwork instruments in ASD.

7. Limitations and threats to validity

In this study, we explored various validity threats that may arise during the realization of our research, encompassing internal, external, construct, and conclusion validity.

Regarding internal validity, potential issues may arise from selection bias, history effects, instrumentation, and maturation. To mitigate these threats, we employed random sampling techniques and defined clear inclusion criteria for selecting ASD instruments and research articles. Additionally, we carefully control external events and changes by collecting data over a consistent time period and conducting longitudinal studies. Standardization and pilot testing of instrument administration and interpretation help address potential instrumentation concerns.

Construct validity threats may arise from conceptual clarity, instrument validity, and measurement errors. We take measures to address these concerns by providing a clear definition of the constructs of interest and employing a conceptual framework. Established and validated teamwork instruments are used to ensure accurate measurement of constructs. Additionally, we employed reliable data collection methods and appropriate techniques to minimize measurement errors.

External validity threats revolve around generalizability and timeframe relevance. To address these concerns, we clearly define the target population and context of our study. Efforts are made to replicate real-world conditions in the study design to enhance validity. We ensure that data collection and analysis are up-to-date and reflect current practices in the field. Moreover, the study relies on solid theories that support the analyzed teamwork instruments. The results may not fully capture the variability or applicability of other theoretical frameworks, potentially limiting the external validity of the conclusions to different theoretical perspectives.

Regarding conclusion validity, the study analyzes eight specific teamwork instruments for Agile Software Development (ASD). The findings may not fully represent the entire population of ASD instruments, potentially limiting the generalizability of the results to other instruments that were not included in the analysis. Moreover, the study focuses on teamwork instruments specifically designed for an agile context. The results may not directly apply to teamwork instruments used in non-agile contexts, reducing the external validity of the findings for broader applications.

By proactively addressing these validity threats and implementing appropriate actions, we aim to enhance the quality and reliability of our study, providing more robust and meaningful findings for the scientific community.

8. Implications

In light of the findings from this study, we have identified several implications for both research and practice in the context of measuring TWQ in ASD.

Implications for research. This study sheds light on the evolution of TWQ instruments, providing valuable insights for further research. The findings highlight the existence of multiple models with different constructs and measures for assessing TWQ and TWE. This prompts researchers to delve deeper into understanding the relationships between these instruments and how they have evolved over time. The study also emphasizes the need for standardization of terminology, as semantically similar factors are often labeled differently across instruments. This calls for future research to focus on developing a conceptual framework that integrates instrument factors within the agile context, facilitating better alignment and comparison of results. Moreover, the identified gaps and specialized factors specific to the agile context present opportunities for researchers to develop new instruments and further advance the understanding of teamwork in ASD.

Implications for practice. The findings of this study hold practical implications for organizations engaged in ASD. Classifying teamwork instruments into generic and specific agile instruments guides practitioners in selecting appropriate instruments based on their specific context and requirements. The evolution of instruments with specialized factors underscores the importance of considering these factors when evaluating and managing teamwork in agile projects. Furthermore, the identified need for terminology standardization emphasizes the importance of consistent and clear team communication. Organizations can benefit from adopting a unified taxonomy derived from this research to ensure consistent understanding and usage of teamwork concepts. The study also emphasizes the value of developing new instruments that align with the agile context, allowing organizations to assess and improve their teamwork practices effectively. Overall, the insights gained from this study can inform and guide practitioners in selecting and implementing appropriate teamwork instruments and strategies to enhance collaboration and team performance in ASD projects.

9. Final Remarks

Our study significantly contributes to the teamwork literature by exploring the relationship between ASD literature-based codes identified by Freire et al. [7] and Agile Instruments factors in ASD. By comparing eight specific ASD instruments and showcasing the frequency of matches, we offer insights that can inform further research. Moreover, our identification of ASD instrument questions through semantic analysis enables broader coverage for future studies, potentially leading to new discoveries and advancements in research. Moreover, our findings demonstrate that researchers have employed numerous factors to measure Teamwork Quality (TWQ) and Teamwork Effectiveness (TWE) in ASD. Additionally, the observed similarity in questions across different instruments suggests the need for standardizing terminology. By highlighting the most frequent questions of each instrument, our results support the development of a unified Teamwork instrument in ASD.

The presented results offer valuable insights for both practitioners and researchers. For practitioners, this paper serves as a practical guide in utilizing the presented teamwork instruments, as it provides detailed information about their characteristics. This facilitates their practical application in Agile Software Development (ASD) projects. For researchers, this work highlights identified gaps and specialized factors specific to the agile context, offering opportunities to develop new instruments and advance the understanding of teamwork in ASD.

Future research endeavors should focus on establishing a unified taxonomy for teamwork instrument factors in ASD, creating a standardized framework to categorize and organize these factors consistently. Conducting longitudinal research can provide valuable insights into the evolution and effectiveness of teamwork instruments over time, enhancing our understanding of their performance in various contexts and identifying opportunities for adaptation and improvement. Additionally, investigating the relationship between specific teamwork instruments and project outcomes in ASD can shed light on how effective teamwork, as measured by these instruments, influences project success, productivity, and overall performance.

Supplementary Material

To ensure the study's transparency and completeness, we have provided a Supplementary Material [1] that contains the eight teamwork instruments factors, and questions. The additional methodological details and a comprehensive presentation of the results.

CRediT authorship contribution statement

Ramon Santos: Conceptualization of this study, Methodology, Software. **Felipe Cunha:** Data analysis, Writing - Review & Editing. **Thiago Rique:** Data collection, Visualization. **Mirko Perkusich:** Writing - Review & Editing. **Ademar Neto:** Supervision. **Danyllo Albuquerque:** Review & Editing. **Hyggo Almeida:** Funding acquisition, Resources. **Angelo Perkusich:** Project administration.

References

[1] Alderfer, C.P., 1983. An intergroup perspective on group dynamics. Technical Report. Yale Univ New Haven CT School of Organization and Management.

[2] Anderson, N., West, M.A., 1996. The team climate inventory: Development of the tci and its applications in teambuilding for innovativeness. European Journal of work and organizational psychology 5, 53–66.

[3] Armour, P.G., 2002. The spiritual life of projects. Communications of the ACM 45, 11–14.

[4] Dickinson, T.L., McIntyre, R.M., 1997. A conceptual framework for teamwork measurement, in: Team performance assessment and measurement. Psychology Press, pp. 31–56.

[5] Dikert, K., Paasivaara, M., Lassenius, C., 2016. Challenges and success factors for large-scale agile transformations: A systematic literature review. Journal of Systems and Software 119, 87–108.

[6] Dingsøyr, T., Fægri, T.E., Dybå, T., Haugset, B., Lindsjørn, Y., 2016. Team performance in software development: research results versus agile principles. IEEE software 33, 106–110.

[7] Freire, A., Neto, M., Perkusich, M., Gorgônio, K., Almeida, H., Perkusich, A., 2021. Towards a comprehensive understanding of agile teamwork: A literature-based thematic network, SEKE.

[8] Freire, A., Perkusich, M., Saraiva, R., Almeida, H., Perkusich, A., 2018. A bayesian networks-based approach to assess and improve the teamwork quality of agile teams. Information and Software Technology 100, 119–132.

[9] Goncalves, E., Lima, P., Cerdeiral, C., Diirr, B., Santos, G., 2021. Tact: An instrument to assess the organizational climate of agile teams-a preliminary study. Journal of Software Engineering Research and Development 9, 18–1.

[10] Gren, L., Torkar, R., Feldt, R., 2015. Group maturity and agility, are they connected?–a survey study, in: 2015 41st Euromicro Conference on Software Engineering and Advanced Applications, IEEE. pp. 1–8.

[11] Guzzo, R.A., Shea, G.P., 1992. Group performance and intergroup relations in organizations. .

[1] https://figshare.com/s/4b7ad201c314512f372b

[12] Hackman, J., 1987. The design of work teams. in jw lorsch (ed.), handbook of organizational behavior, englewood cliffs, nj: Prentice-hall .

[13] Hoegl, M., Gemuenden, H.G., 2001. Teamwork quality and the success of innovative projects: A theoretical concept and empirical evidence. Organization science 12, 435–449.

[14] Lei, P.W., Wu, Q., 2007. Introduction to structural equation modeling: Issues and practical considerations. Educational Measurement: issues and practice 26, 33–43.

[15] Lindsjørn, Y., Sjøberg, D.I., Dingsøyr, T., Bergersen, G.R., Dybå, T., 2016. Teamwork quality and project success in software development: A survey of agile development teams. Journal of Systems and Software 122, 274–286.

[16] Lukusa, L., Geeling, S., Lusinga, S., Rivett, U., 2020. Teamwork and project success in agile software development methods: A case study in higher education, in: Eighth International Conference on Technological Ecosystems for Enhancing Multiculturality, pp. 885–891.

[17] Marsicano, G., da Silva, F.Q., Seaman, C.B., Adaid-Castro, B.G., 2020. The teamwork process antecedents (tpa) questionnaire: developing and validating a comprehensive measure for assessing antecedents of teamwork process quality. Empirical Software Engineering 25, 3928–3976.

[18] Mathieu, J., Maynard, M.T., Rapp, T., Gilson, L., 2008. Team effectiveness 1997-2007: A review of recent advancements and a glimpse into the future. Journal of management 34, 410–476.

[19] Melo, C.d.O., Cruzes, D.S., Kon, F., Conradi, R., 2013. Interpretative case studies on agile team productivity and management. Information and Software Technology 55, 412–427.

[20] Moe, N.B., Dingsøyr, T., Dybå, T., 2010. A teamwork model for understanding an agile team: A case study of a scrum project. Information and software technology 52, 480–491.

[21] Moe, N.B., Dingsøyr, T., Røyrvik, E.A., 2009. Putting agile teamwork to the test–an preliminary instrument for empirically assessing and improving agile software development, in: Agile Processes in Software Engineering and Extreme Programming: 10th International Conference, XP 2009, Pula, Sardinia, Italy, May 25-29, 2009. Proceedings 10, Springer. pp. 114–123.

[22] Poth, A., Kottke, M., Riel, A., 2020. Evaluation of agile team work quality, in: Agile Processes in Software Engineering and Extreme Programming–Workshops: XP 2020 Workshops, Copenhagen, Denmark, June 8–12, 2020, Revised Selected Papers 21, Springer. pp. 101–110.

[23] Radhakrishnan, A., Zaveri, J., David, D., Davis, J.S., 2022. The impact of project team characteristics and client collaboration on project agility and project success: An empirical study. European Management Journal 40, 758–777.

[24] Russo, D., Stol, K.J., 2021. Pls-sem for software engineering research: An introduction and survey. ACM Computing Surveys (CSUR) 54, 1–38.

[25] Salas, E., Sims, D.E., Burke, C.S., 2005. Is there a "big five" in teamwork? Small group research 36, 555–599.

[26] Santos, R., Cunha, F., Rique, T., Perkusich, M., Almeida, H., Perkusich, A., Costa, Í., . A comparative analysis of agile teamwork quality instruments in agile software development: A qualitative approach .

[27] Serrador, P., Pinto, J.K., 2015. Does agile work?—a quantitative analysis of agile project success. International journal of project management 33, 1040–1051.

[28] Silva, M., Perkusich, M., Freire, A., Albuquerque, D., Gorgônio, K.C., Almeida, H., Perkusich, A., Guimaraes, E., 2022. A comparative analysis of agile teamwork quality measurement models. Journal of Communications Software and Systems 18, 153–164.

[29] Stavru, S., 2014. A critical examination of recent industrial surveys on agile method usage. Journal of Systems and Software 94, 87–97.

[30] Strode, D., Dingsøyr, T., Lindsjorn, Y., 2022. A teamwork effectiveness model for agile software development. Empirical Software Engineering 27, 56.

[31] Truong, D., Jitbaipoon, T., 2016. How can agile methodologies be used to enhance the success of information technology projects? International Journal of Information Technology Project Management (IJITPM) 7, 1–16.

[32] Van Assen, M.F., 2000. Agile-based competence management: the relation between agile manufacturing and time-based competence management. International Journal of Agile Management Systems 2, 142–155.

[33] Verwijs, C., Russo, D., 2023. A theory of scrum team effectiveness. ACM Transactions on Software Engineering and Methodology .

[34] Wheelan, S.A., Hochberger, J.M., 1996. Validation studies of the group development questionnaire. Small group research 27, 143–170.