1 Introduction

As a way to keep being competitive in a globalized world, software industries worldwide have been moving into a global team approach by hiring from regions with an abundance of developers at a low cost (Alqahtani et al., 2013). Global virtual teams, or GVTs, can be described as electronically connected workgroups formed by people from different cultures and geographies (Daim et al., 2012) that function as a team by working on shared goals (Scott, 2013).

Those teams can be purely virtual, relying solely on computer-mediated communication to connect their members, or they can be partially distributed, when some collocated collaborators (working together from the same space) are joined by remote peers through virtual means (Bos et al., 2006; Webster & Wong, 2008). Those different formats result in new types of work patterns, decision-making approaches, and relationships, which in turn open new challenges to theory, practice, and research when it comes to understanding team effectiveness (Alsharo et al., 2017).

Parallel to the increase in GVTs in the software development industry, there is also an increase in the adoption of Agile methodologies by those teams, so that Agile methods are now applied to a distributed context (Alqahtani et al., 2013; Pries-Heje and Pries-Heje, 2011; Persson et al., 2012). Agile is built around the Agile manifesto, which "emphasizes individuals and interactions over processes and tools, working software over comprehensive documentation, customer collaboration over contract negotiation, and responding to change over following a plan" (Agile Alliance, 2001). Amongst the frameworks originated by Agile, Scrum is one of the most widely used (Jalali and Wohlin, 2010). Scrum is an iterative approach to development that values transparency, inspection, and adaptation (Agile Alliance, 2001). According to Scrum Alliance's "State of Scrum" reports, the percentage of respondents applying Scrum as their project methodology has increased from 61% in 2013 to 94% in 2018, with the IT industry leading the way with the highest rate of its adoption (Scrum Alliance, 2013; Scrum Alliance, 2018). The values reinforced by Agile methods such as Scrum change a team in terms of structural features, spirit, amount and nature of interactions, creating a unique culture (Ashmore, 2012). The changes in methodology also impact how subgroups get formed within a team (Pflügler et al., 2018).

A subgroup is a subset of at least two members with some degree of interdependence from the rest of the team (Carton and Cummings, 2012). It is formed based on faultlines, a concept introduced by Lau and Murnighan (1998) as an alignment of characteristics that could split members into smaller groups based on one or more attributes. Due to the geographic dispersion of their members, GVTs are likely to experience the formation of subgroups (Gilson et al., 2015). Even if subgroup formation can lead to positive outcomes for the overall team, the majority of past studies focused on the negative effects of it, which can include ethnocentrism, asymmetry in the perception of fairness, disturbs to decision making process, and interruption of knowledge flow within the team (Pflügler et al., 2018).

Past literature reviews focused intensely on challenges faced by GVTs. Examples are studies by Zahedi et al. (2016) that synthesized knowledge sharing challenges in practices in Global Software Development; Morrison-Smith and Ruiz (2020) that focused on how technology can be better put to use to address GVTs challenges; Scott and Wildman (2015), who showed how the challenges brought by culture could influence GVTs effectiveness; Jain and Suman (2015) that looked at the various phases of software development to understand challenges and best practices applicable to each; and, finally, Jalali and Wohlin (2010) who included the Agile methods perspective in a literature review by investigating when Agile had been applied efficiently. Although virtual teams have been a frequent research topic across

several areas such as Information Systems, Human Resources, Management, Computer Science, Psychology, and other disciplines for the past two decades (Hacker et al., 2019), there is still a lack of understanding of how subgroups are formed and how they impact virtual Agile teams. We choose the software development industry in this study since "today working as an IT developer means being able to collaborate in globally distributed work, and this makes global software development an obvious domain for investigating distributed work" (Bjørn et al., 2019).

The limited research focuses on subgroups and their impact on GVTs (Gilson et al., 2015), especially within software development projects using Agile approaches (Pflügler et al., 2018), is the gap to be filled out by this paper. When Gilson et al. (2015) presented opportunities to drive the following ten years of research in virtual teams (VTs), they asked for more studies around subgroups: "if conditions are ripe for subgroups to form in VTs, then research must start to examine how, when, and the implications of such groupings on VT processes, emergent states, and outcomes" (Gilson et al., 2015, p.16). Morrison-Smith and Ruiz (2020) also suggested an opportunity for future research to ramify on the differences of trust and conflict in subgroups in balanced and unbalanced teams. Pflügler et al. (2018) asked for more research on how Agile methods could reduce location-based faultlines in distributed teams and how Agile teams resolve subgroup-related issues. Przybilla referred to this gap as a critical void, whose "effects may have far-reaching implications for team dynamics and thus project success" (Przybilla et al., 2018, pg. 1).

The Covid-19 outbreak brings a new urgency to understand better how VTs work and perform since working from home is no longer a privilege but a necessity (Banjo et al., 2020). Furthermore, several companies such as Facebook, Twitter, Nielsen Research, Nationwide Research, Tata Consultancy Services, Infosys, and HCL Technologies already reported plans to, post-pandemic, continue to transition their workforce to working remotely (Neeley, 2021).

2 Method

We conducted this paper through a literature review, which "... is an essential feature of any academic project. An effective review creates a firm foundation for advancing knowledge" (Webster and Watson, 2002, p. xiii). It is also a helpful way to summarize past research findings and uncover areas where more research could be needed (Snyder, 2019).

Snyder (2019) suggests putting together a search strategy to identify relevant literature, including search terms, and inclusion criteria to recognize what is relevant. For this review, we used the query *(("global" or "international" or "geographically dispersed" or "virtual") and "teams")) and ("agile" OR "scrum") and ("software development" or "it projects").* We specifically included "Scrum" since it is one of the most widely used frameworks for Agile (Jalali and Wohlin, 2010). Although the term "subgroup" is important for this literature review, it was excluded from the query since we understood that adding it could pose a limitation and prevent relevant records from being returned as part of the results.

We ran the criteria in November 2020 across three different databases: Google Scholar, Scopus, and Web of Science (WoS). We filtered by articles published since 2010, then removed patents, citations, and proceedings books from past conferences. We only considered articles in English. Google Scholar returned 935 results for the parameter informed above. Those search results were ranked by relevance, and the first 65 articles returned were cataloged. For Scopus, we ran the query against title, keywords, and abstract, resulting in 31 documents found, all of them cataloged. WoS returned 26 results, which were also cataloged. The total records cataloged from the three databases represented a total of 122.

Once the search results had been cataloged, we began detecting and removing duplicates across the three databases. That led to the selection of 100 papers for further refinement. Snyder (2019) mentions that one can conduct the review in stages by reading the abstract and making selections and later by doing the full-text reading for the final selections. We began by reading the title, abstract, and keyword to understand if an article fits our context. We excluded articles that lacked quality or did not fit into the context of GVTs working with Agile to develop software. We also removed findings from past conferences that had two or fewer pages.

After this check, the number of articles to analyze was reduced to 77, to which we read the full text. If the entire paper was not accessible, we requested it from the authors through academic social networks such as ResearchGate. At the analysis step of this study, 11 requests for articles remain unanswered, and those articles were, therefore, excluded. After reading the full text of the 66 remaining articles, there was a new round of refinement, where we removed articles that did not relate to the research objectives, leading to 43 articles.

With the final list of selected articles, Webster and Watson (2002) recommended doing a backward and forward search, reviewing the references of the selected studies, and looking forward to other articles that quote the selected papers. We then cataloged new papers found through citations of the past articles selected or by searching for more work on authors relevant to the goal of this review. The total number of articles selected for this literature review, once adding backward and forward search, is 60.

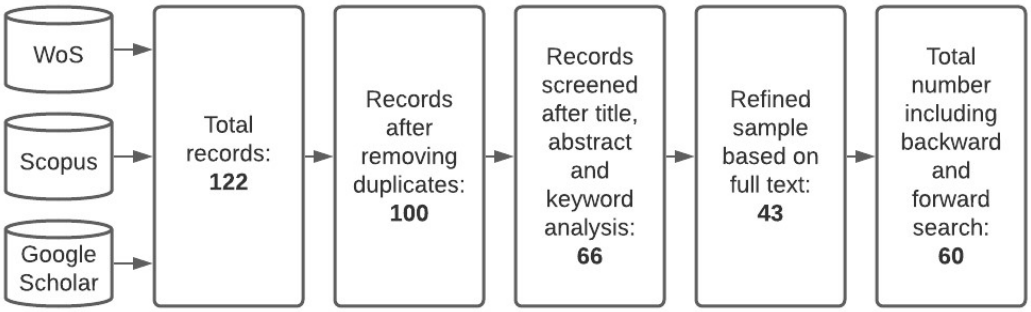

Figure 1 - Steps for article selection
Source: Author

All the 60 final selected articles were read and categorized using content analysis techniques described by Moser and Korstjens (2018): initially, an inductive coding scheme was developed using open coding and abstraction. We created labels as we read through the papers. The next step was to cluster the labels identified into preliminary categories, which were then ordered according to similarities, forming broader high-level categories. Those broader categories are the three sections presented below: the first group refers to the impact of diversity on a GVT. The second group is about exploring the combination of Agile with distributed development. Finally, the third group had a more significant focus on subgroup formation, strength, and impacts. Figure 2 shows the conceptual map elaborated through different categories identified during Content Analysis.

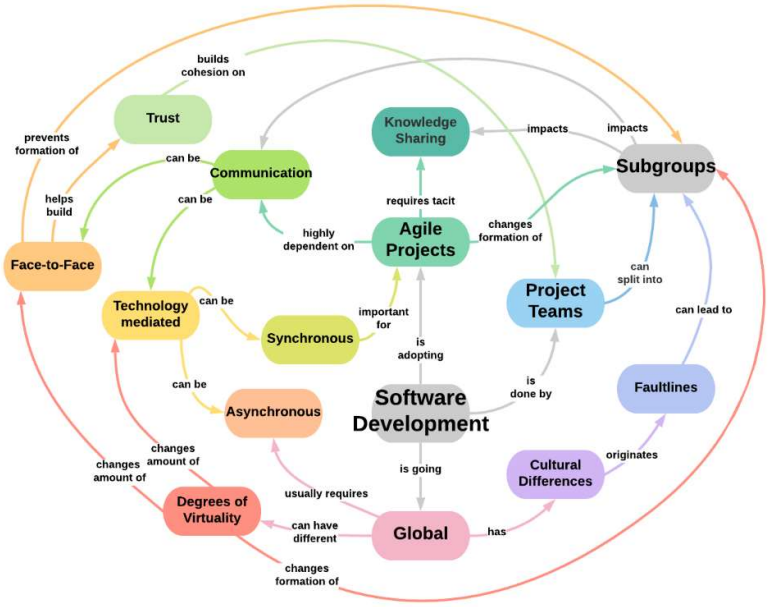

Figure 2 - Conceptual map derived from categories identified.
Source: Author

3 Analysis of the Results

This session describes the main topics identified during our analysis. We divided the content as follows: Section 3.1 addresses what is being discussed regarding multicultural teams. Section 3.2 explores the impacts of combining Agile with distributed development. Section 3.3 takes a deeper look into subgroup formation for GVTs.

3.1 Global Virtual Teams

Global Software Development (GSD), which has become a standard in the software development industry today, happens when the teams developing software collaborate across different locations, time zones, cultural and organizational backgrounds (Bjørn et al., 2019). Distributed, Multisite, Dispersed, Offshoring, and Outsourcing Development are other terms found through the literature to describe GSD (Shrivastava and Rathod, 2015).

This form of development relies on GVTs, which are often composed of individuals across the globe who carry different cultural values and can speak different languages (Scott and Wildman, 2015). A GVT can be described as "culturally diverse, geographically dispersed, electronically-connected workgroups" (Daim et al., 2012, p. 202) that, to function as a team, need to work on shared goals (Scott, 2013). GVTs are also occasionally labeled as multicultural, multinational, transnational, cross-cultural, or international teams (Scott and Wildman, 2015). Many adopt the GSD/GVT strategy to take advantage of worldwide talent, around-the-clockwork, and cheaper resources (Ashmore, 2012).

Several studies on the GVT literature focus on its demographic differences. Based on the concept of culture as national differences, Hofstede's cultural dimension theory is a popular way to measure culture within a GVT (Scott and Wildman, 2015). Ende et al. (2013) used Hofstede's framework to categorize and analyze certain behaviors from German and Mongolian students working together. Nordbäck and Espinosa (2019) used Hofstede's power distance score to understand different formality levels in leadership expectations on a multicultural team. Nevertheless, there is a need to look at culture from a more complex perspective, perceiving

culture with a more multifaceted and dynamic approach than the one offered by Hofstede (Scott and Wildman, 2015). This is done by Bjørn et al. (2019), who adopted a "translocal" practice to their studies, detaching culture from national constraints by thinking of it not as a noun but as verbs and doing. Matthiesen et al. (2014) mentioned that understanding culture based on nationalities could create stereotypes and opted for "localizing the global" by studying global engagement through local practices. Nordbäck and Espinosa (2019) decided not to automatically assign a value of high or low power distance to their interviewed country of origin but assess its values at the individual level. Still, while reporting their results, the authors often mentioned how one nationality versus another would behave.

Regardless of being a regional or national culture, past studies often bring the cultural diversity of a GVT as a drawback. Issues originated from it can lead to diverse leadership expectations making shared leadership less effective (Nordbäck and Espinosa, 2019), lack of transparency, honesty, trust and team's responsibility understanding (Alqahtani et al., 2013), and different perceptions around punctuality, scheduling and urgency causing misunderstandings and negatively impacting relationships (Jain and Suman, 2015). Language barriers and linguistic challenges can lead to problems (Alqahtani et al., 2013; Daim et al., 2012), such as influencing a team's activity regarding the use of instant messages to communicate (Stray et al., 2019) and increasing complexity in knowledge sharing (Anwar et al., 2019; Wendling et al.; 2013; Zahedi et al., 2016). According to Stadler et al. (2019), since distributed teams use various channels to communicate, language differences have a more substantial impact on them than on collocated ones, making it harder for non-native speakers to follow conversations and argue with colleagues.

However, the same diversity that can cause issues can also lead to positive outcomes: Badiale (2020) mentioned that different communication styles and understanding of each other could pose a setback but, on the other hand, the diversity of thoughts, knowledge, and capabilities can be an advantage. Wendling et al. (2013) mentioned that cultural differences, distance, and multiple time zones could be barriers and enablers to knowledge sharing by providing a wealth of different professional profiles. Scott and Wildman (2015) concluded that a relationship between a GVT and cultural differences could be harmful, positive, or neutral, depending on situational and contextual variables. One way to explain how those outcomes can change is by looking at the subgroup formation within those teams, which will be done as part of the final chapter of this literature review.

Another essential factor to consider when discussing a GVT is its virtuality. Although past literature considered a team's virtuality as a dichotomy, where a team was either virtual or not, virtuality should be seen as a continuum based on the concept of degree of virtuality (Schweitzer and Duxbury, 2010; Bos et al., 2006; Webster & Wong, 2008). Bos et al. (2006) and Webster & Wong (2008) referred to a team's as being "purely virtual" when all team members need to use technology to communicate, and "partially distributed" when part of a team is collocated (communicating face-to-face) and part is remote (using technology to interact). Tran and Nguyen (2018) calculated the virtuality of a team as a percentage of the proportion of time spent communicating through various means versus time spent face-to-face. Schweitzer and Duxbury (2010) added that a team could be placed across a scale, wherein one end would be a team that conducts all its interaction face-to-face, while on the other end would be the team that never meets face-to-face. The authors also noted that geographic dispersion and working asynchronously are conditions for a team to be virtual, and the dependence on electronic means to communicate is a consequence of virtuality. Because of that, it is essential to address the impact that information and telecommunication technologies have on those teams (Scott and Wildman, 2015).

Technology can play a significant role in supporting the execution of Agile practices (Persson et al., 2012). However, it can also bring some challenges to communication, affecting the team's effectiveness (Cruzes et al., 2016). Humans communicate mainly through tone of voice and body language, which are naturally present in face-to-face communication, but lacking when communication occurs via computer (Daim et al., 2012). The absence of this type of communication might not only delay trust between team members (Morrison-Smith and Ruiz, 2020), but when combined with cultural differences, lead to anxiety, conflict, confusion, and miscommunication, negatively affecting relationships (Scott, 2013; Daim et al., 2012). Morrison-Smith and Ruiz (2020) observed that the lack of face-to-face communication also affects informal communication, which is vital to foster the feeling of being part of a team. Not having information situations such lunches, coffee breaks, and after-work activities makes distributed teams focus more on professional aspects (Stadler et al., 2019) and can also decrease the awareness about what is happening across each site, creating misunderstandings and harming trust and relationships (Jain and Suman, 2015).

Although face-to-face communication is not possible for a GVT, several different mediums are available (Ashmore, 2012). Some authors recommend the use of high media synchronicity, such as video conferences, since those are the richer form of online communication available (Borrego et al., 2017), being a low-cost alternative to "being there" (Daim et al., 2012). Amongst its benefits, the literature highlights the increased perception of social presence, better communication performance (Nevo and Chengalur-Smith, 2011), and lower collaboration barriers across sites (Clerc et al., 2011). However, video communication also has its drawbacks, such as its equipment requiring a complicated setup (Sharp et al., 2014), being difficult to use (Liukkunen et al., 2010), and demanding high-speed connection (Daim et al., 2012).

Since temporal distance reduces the overlapping hours of a GVT, there could be constraints to using synchronous interactions (Li and Maedche, 2012), making asynchronous communication quite essential for those teams as well. Stawnicza (2015) observed a trend towards using asynchronous communication that allows for quicker and more instant feedback, such as instant messaging. In general, combining different communication mediums will allow the correct message to flow (Green et al., 2010).

Another essential item to consider when studying GVTs is the methodology the team chooses to follow. According to Aarseth et al. (2014), the organizational challenges brought by GVTs are still underestimated and not fully taken into account by task-oriented approaches, such as the PMI. That, combined with the increase in combining Agile methods with distributed development (Kaur and Haddad, 2015; Lous et al., 2018), calls out for a better understanding of the impacts of Agile methods in GVTs.

3.2 Agile Methodology and Distributed Development

Issues with traditional software development methods, such as bureaucratic and slow processes and final deliverables that did not meet the original expectations (Van Hillegersberg et al., 2011), led to Agile Methods' adoption in the past decade by the software industry (Borrego et al., 2017). When compared to traditional methods, such as Waterfall, Agile differs in how it adapts to change and uncertainty (Green et al., 2010), its response to improve delivery speed (Paasivaara et al., 2018), and its closer relationship with customers and stakeholders (Ashmore, 2012). The most common Agile practices are extreme programming, pair programming, lean software development, and Scrum (Jalali and Wohlin, 2010). Scrum is a framework that iteratively approaches development, and values transparency, inspection, and adaptation (Agile Alliance, 2001; Pries-Heje and Pries-Heje, 2011).

Agile values change a team in terms of structural features, spirit, amount, and nature of interactions, creating a unique team culture (Ashmore, 2012). We can observe those changes concerning improved trust and shared cognition (Li and Maedche, 2012), the role of the testing team (Cruzes et al., 2016), motivation (Noll et al., 2017), communication frequency and formality level (Ashmore, 2012), communication tools and skills needed by the team (Wendling et al., 2013), and dominant factors that can lead to subgroup formation within the team (Pflügler et al., 2018).

One fundamental change brought by Agile is the relevance it gives to communication and collaboration over documentation (Alqahtani et al., 2013), leading to a strong belief that the collocation of the entire team is required (Ashmore, 2012; Green et al., 2010; Jalali and Wohlin, 2010). However, entirely collocated teams are no longer the real-world scenario for several organizations, creating a demand to combine Agile with distributed development (Kaur and Haddad, 2015; Lous et al., 2018).

Applying Agile to a GVT is a potentially rewarding but challenging and risky task (Van Hillegersberg et al., 2011). Issues in this combination arise primarily because Agile practices and distributed development differ from each other in critical principles (Shrivastava and Rathod, 2015): one of the core strengths of Agile lies in team members needing to interact and communicate daily, and since GVTs relies on technology to do, they need to work harder to synchronize (Vallon et al., 2013). The distances faced by the distributed setup lead to communication, coordination, and collaboration challenges, increasing the complexity for project teams (Zahedi et al., 2016).

The distributed-Agile structure creates a strong antagonism: on the one hand, Agile Software Teams prefer to convey in face-to-face communication valuing tacit over explicit knowledge (Almeida et al., 2019). On the other hand, GSD has a more significant need for explicit communication, preferring an approach based on codification (Borrego et al., 2017). It is as if a way for GVTs to compensate for being distributed is to come up with more documents and plans, but using heavy documentation goes against what Agile values in the first place (Alqahtani et al., 2013).

Because of that, knowledge management is one of the critical areas of distributed Agile, negatively affected by distance, time zones, cultural differences, and fewer opportunities to meet face-to-face (Borrego et al., 2017; Wendling et al., 2013; Zahedi et al., 2016). Some authors argue that with a distributed team, members must see themselves as responsible for deciding what and with whom to share information (Sharp et al., 2012). Sharing knowledge across each site demands the proactiveness of each individual to initiate and promote it (Clerc et al., 2011) and the availability and willingness of people to ask and share (Borrego et al., 2016). Since people are scattered apart, they might not identify what is missing elsewhere and are likely to neglect sharing relevant information to help remote colleagues (Zahedi et al., 2016).

Past research brings some suggestions to address those challenges. Borrego et al. (2017) recommended the adoption of Communities of Practices. Clerc et al. (2011) and Persson et al. (2012) noted success cases with a hybrid knowledge sharing approach that combined codification and personalization practices and created some formal control elements to compensate for the distribution. Kaur and Haddad (2015) suggested that all team members send out status notes to each other at the end of their workday to increase team awareness. Sharp et al. (2012) and Stadler et al. (2019) encouraged pair programming. Zahedi et al. (2016) recommended choosing a team member who is an expert in some business domain or technical skill and having that person acting as a human bridge, facilitating knowledge in different locations.

Time zone differences can also become a sensitive area, especially for teams adopting Agile Scrum. Ashmore (2012) and Cruzes et al. (2016) observed that temporal distance could prevent team members from joining Sprint Ceremonies. Bjørn et al. (2019) described a stressful environment for a team due to the need to join the Daily Ceremony outside of their working hours. Because of that, it is essential to have at least some overlap in business hours between different sites collaborating (Sharp et al., 2014).

The iterative approach suggested by Scrum sprints was also reported by Bjørn et al. (2019) as something that negatively impacted the working conditions by adding pressure for shorter sprints and tiger deadlines. Other factors that could negatively affect the team are continuous testing highlighting inferior and low-quality code, making developers feel exposed in their deficiencies (Kaur and Haddad, 2015); and issues caused by cultural differences, such as language barriers causing offshore teams to be quieter during the ceremonies (Alqahtani et al., 2013).

However, many success cases found in the literature also highlight the benefits of combining distributed development with Agile: Pries-Heje and Pries-Heje (2011) mentioned that Scrum could help a GVT build relationships, trust, coordinate work, communicate better, and give the team more energy and motivation. Techniques such as the Ceremonies can improve communication and collaboration (Badiale, 2020; Paasivaara et al., 2018), acting as a facilitator for knowledge sharing (Almeida et al., 2019), helping teams to manage cross-boundary dependencies (Li and Maedche, 2012) and maintaining group awareness (Zahedi et al., 2016). The short iterations can help reveal issues very swiftly (Stadler et al., 2019). The repetition of ceremonies and rituals provides the team with familiarity, encouraging communication and collaboration (Scott, 2013). The predefined processes in Scrum also allow for better mechanistic leadership coordination, facilitating transparency and preventing redundant leadership (Nordbäck and Espinosa, 2019). Because of that, authors such as Nevo and Chengalur-Smith (2011) and Mudumba and Lee (2010) believe that Agile methods are, in fact, recommended and valuable for distributed teams.

It is not usual to modify Agile to make it work in the distributed context. According to Van Hillegersberg et al. (2011), Agile is about trust and empowerment; therefore, freedom should be allowed to local teams to tailor the process and documentation style. Lous et al. (2018) can be used as example – the author studied a team that crafted an entire work environment to embrace the distributed setup. However, not all modification scenarios have a happy ending: Vallon et al. (2013) reported a case where the customizations made to apply Scrum to distributed development focused mainly on the primary supplier, leading to lower transparency, efficiency, and stress increase across the teams. How much change should be allowed to a GVT trying to go Agile is still an item open for debate (Jalali and Wohlin, 2010).

Overall, any software development team, especially VTs transitioning to Agile, need to consider the changes in culture, tolling, and attitude required to make the shift (Ashmore, 2012). Introducing Agile in a big bang fashion is hard, so it is recommended that companies take an experimental approach first and then expand it gradually (Paasivaara et al., 2018; Van Hillegersberg et al., 2011).

3.3 Teams and Subgroups

GVTs, due to their multicultural and geographically dispersed characteristics, are often linked to the formation of subgroups (Pflügler et al., 2018). Because of that, subgroups are a common thread of interest not only in cultural diversity but also in VT literature (Tran and Nguyen, 2018).

A subgroup is a subset of at least two team members with some form or degree of interdependence (Carton and Cummings, 2012). It is formed based on faultlines, a concept introduced by Lau and Murnighan (1998) as an alignment of characteristics that could split members into subgroups based on one or more attributes. These hypothetical lines can be related to several topics, such as demography, personality, employer relation, and knowledge orientation (Pflügler et al., 2018). Carton and Cummings (2012) grouped faultlines into three categories: separation-based, related to the values that team members possess; disparity-based, related to resources team members might have; and variety-based, related to how team members process information.

Even though different faultlines can exist in a team, not all of them necessarily lead to the formation of subgroups (Pflügler et al., 2018; Hinds et al., 2014). Only when a dormant faultline is activated, and team members perceive a group's division, is a subgroup formed (Jehn and Bezrukova, 2010). Originally, Lau and Murnighan (1998) proposed that faultline activation would be dependent on the similarities and salience of group members' attributes. However, it is not faultline strength that necessarily leads to its activation: this can happen due to team members carrying certain personality traits such as entitlement (Jehn and Bezrukova (2010), or when a power contest is triggered along geographic-nationality and language-based faultlines (Hinds et al., 2014).

Carton and Cummings (2012) classified subgroups formed by activated faultlines in three types: identity-based, about people who share common values; knowledge-based, based on technical language; and resource-based, grouping members based on the power they have. The authors believe that identity-based subgroups have a stronger relationship with separation-based faultlines, resource-based groups are likely formed by disparity-based faultlines, and variety-based faultlines often form knowledge-based subgroups. Even though some studies show a positive effect of subgroups on team members, especially those based on knowledge, the negative consequences of identity and resource-based subgroups are majorly reported (Pflügler et al., 2018). Also, not all subgroups impact a team the same way - Panteli and Davison (2005), for instance, mentioned that the influence a subgroup can have on the overall team could be low, moderate, or high.

Pflügler et al. (2018) observed that due to its dispersed and multicultural characteristics, GVTs are often linked to identity-based subgroups, with the faultlines locations and languages being the trigger. We found examples while analyzing GVT literature, with research showing how distributed teams can lead to local clusters (Manteli et al., 2014), fostering a "us versus them" mindset (Scott, 2013), creating disconnections and absence of togetherness (Stawnicza, 2015; Kaur and Haddad, 2015; Jalal and Wohlin, 2010), impeding the implementation of standard work practices (Li and Maedche, 2012), reducing collaboration and communication (Scott and Wildman, 2015) and creating silo-type structures that lead to barriers in knowledge sharing (Wendling et al., 2013). Not so common but still reported are also examples of resource-based subgroups within GVTs (Bjørn et al., 2019), with the imbalance in powers causing team members less involved and motivated (Clerc et al., 2011; de Farias et al., 2012) or even exacerbating the impacts of subgroups in collaboration (Matthiesen et al., 2014).

Past studies focused on understanding how demographic factors, such as alignment of age, gender, and ethnicity, could impact subgroup formation and affect the team. Lau and Murnighan (1998) proposed that demographic-related faultlines become stronger as more attributes are highly correlated and weaker when multiple fragment subgroups can form. Cramton and Hinds (2004) noted that when demographic attributions align with geographical location, subgroups are likely to become more salient and can lead to subgroup ethnocentrism, impacting effectiveness. Hinds et al. (2014) observed how language asymmetries aligned with

distance and nationality and discovered that it was not faultline strength, but the emotion-regulation strategies adopted by team members that dictated how powerful a subgroup would be: the more emphatic the team, the lower were the impacts the subgroups had on them. Jehn and Bezrukova (2010) noticed that coalition formation and conflict mediated the relationship between activated faultlines and group outcomes, such as decreased satisfaction and performance. Besides faultline strength and activation, other factors must be considered when discussing subgroups for GVTs: how distant the faultlines are from each other (Bezrukova et al., 2009), the size and balance of each subgroup, the relationship among subgroups (Carton and Cummings, 2012), and the degree of virtuality a team is exposed to (Webster & Wong, 2008).

For faultline distance, Bezrukova et al. (2009) studied a collocated team and reported that the adverse effects of social category faultlines were directly associated with faultline distance. Relative group size and disparity in subgroup power must also be taken into consideration, since those can create different group dynamics, with, for example, smaller-sized subgroups having more significant difficulty in being accepted by stronger and larger subgroups, or groups that contain subgroups of comparable power and size experiencing more intense conflict (Lau and Murnighan, 1998). For a GVT, the number of locations and the number of people at each location are likely to affect subgroup dynamics (Cramton and Hinds, 2004), with geographical proximity of some team members enabling subgroup formation (Panteli and Davison, 2005) and team virtuality and perception of subgroup formation displaying a positive relationship (Tran and Nguyen, 2018).

Several studies noted how a team's distribution could affect subgroup formation. Bos et al. (2006) discovered that the collocated members interacted less with their remote colleagues, creating what was referred to as a "collocation blindness." Bond-Barnard et al. (2018) observed a high measure of association between collaboration and collocation, concluding that when the project team works physically close, they collaborate more. O'Leary and Mortensen (2010) noted that teams with geographically-based subgroups performed poorlier regarding identification, transactive memory, conflict, and coordination, issues that were intensified with an unbalanced configuration of the teams across locations. Curiously, the authors found that teams with a single isolated member experienced positive outcomes. Webster & Wong (2008) noted that the collocated part of a team experienced even higher local group perceptions on a semi-virtual team than on a fully collocated one, affecting the in-group identity, communication, and trust. The authors suggested that it is best to avoid semi-virtual teams, keeping all team members local or all remote. Polzer et al. (2006) discovered that there is less conflict and more trust among collocated team members than among distant ones, and compared three different types of team configuration: (a) fully dispersed, (b) two subgroups of three members each, and (c) three subgroups of two members each. Scenario (b) reported most trust issues and conflicts. They also observed that geographic distance between subgroups evoked more conflict and less trust when subgroup members were from the same nationality.

Because of the reported issues, it seems only natural that leaders try to reduce interdependence amongst distributed teams - but, by doing so, the motivation to engage across differences is also being limited (Cramton and Hinds, 2004). Instead, the path forward should be to highlight and celebrate the diversity - research suggests cultural training and workshops (Anwar et al., 2019; Jain and Suman, 2015), and having local leaders or parties highlighting the importance of institutionalizing the cultural context of each team (de Farias et al., 2012). Another healthy practice is relocation, even if temporary, allowing members to experience and understand another's culture (Daim et al., 2012). Those visits also help mitigate asymmetries between different locations (Paasivaara et al., 2018). The face-to-face conversations can

facilitate building trust and motivation (Jain and Suman, 2015), increasing information exchange (Cramton and Hinds, 2004), and creating a social relationship that stimulates knowledge sharing in the team (Anwar et al., 2019; Zahedi et al., 2016). With the right encouragement to engage across differences, ethnocentrism can become cross-national learning, and demographic faultlines can result in more resilient teams and team members (Cramton and Hinds, 2004).

Literature also highlights that frequent and informal communication can create proximity, reduce the perceived distance between team members, and foster the feeling of being part of a team (Stadler et al., 2019; Morrison-Smith and Ruiz, 2020). The extent to which group members identify with the team can help prevent adverse effects of subgroup formation: Bezrukova et al. (2009) noted that high levels of team identification led to higher levels of group performance, even for groups with strong and distance faultlines. Since Agile changes the communication frequency and formality level (Ashmore, 2012), it is vital to understand how it affects faultline strength and subgroup formation.

Some studies are beginning to do so: Scott (2013) studied a team split across two countries that extended the values promoted by Agile Scrum and invented their own, creating a united team with equal power status between locations. Pflügler et al. (2018) researched the differences in subgroup formation with GVTs using Agile and traditional approaches. They discovered that although geographical distance and language faultlines can be activated under both methodologies, task-based subgroups are dominant in traditional method projects, whether previous ties are more relevant for subgroup formation in Agile teams (Pflügler et al., 2018). Overall, Pflügler et al. (2018) concluded that there are fewer solid and severe subgroups in projects with Agile methods than in traditional approached ones

Contrary to that, Przybilla et al. (2018) found out that Agile practices act as a double-edged sword regarding the negative impacts of subgroups. The improved information flow of those teams meliorates the effects of subgroups on conflict and satisfaction. However, ceremonies, such as the retrospective, increase reflexivity, add conflict, and reduce satisfaction. The authors concluded that intense communication could highlight differences and potentially divide colleagues even further.

As Agile adoption grows, the understanding of issues such as 1. how teams with different virtuality degrees lead up to different faultlines, 2. what makes those faultlines strong and weak, 3. what activated them, and 4. how subgroups that come out of the impact the overall team, will allow better development of further research on virtual software development teams.

4 Final Remarks

This paper explored the challenges and benefits of GSD teams adopting Agile methods, studying how subgroups are formed and their impacts on those teams. Like diversity, which can act as a setback and a success ingredient for GVTs, Agile methods can lead to both negative and positive outcomes for distributed development. If, on one hand, the reliance on technology to communicate leads to complexity in knowledge management, on the other hand, ceremonies, such as the ones proposed by Scrum, can help bring team awareness. We showed that the methodology a team will work with, as well as characteristics of team members and how people are distributed across different locations can affect faultline activation and subgroup formation and must therefore be considered by GVTs leaders. Most importantly, the more team members trust and emphasize with one another, the greater are the chances for success.

This literature review has a two-fold contribution to the Project Management and Software Development communities. Theoretical contributions lie in answering the calls from Gilson et al. (2015) and Morrison-Smith and Ruiz (2020) for more studies of subgroups in

GVTs, as well as the ones from Przybilla et al. (2018) and Pflügler et al. (2018) on how Agile can affect faultline and subgroup formations. With that in mind, we present a state-of-the-art picture of what is currently known regarding how distributed Agile and different degrees of virtuality interplay with faultline and subgroup formation within a GVT. We extended past reviews on subgroup formation for GVTs to show that work methodology and virtuality are essential components that GVT research must consider.

This study also has practical implications by offering project managers and leaders of GVTs adopting Agile tools to create more cohesive and performative teams. Some suggestions include: allowing for time overlaps among teams in different locations (Sharp et al., 2014), adapt Agile so it will not rely exclusively on personalization for knowledge management (Van Hillegersberg et al., 2011; Lous et al., 2018), invest in different communication methods, balancing out asynchronous and synchronous communication (Green et al., 2010), and plan for a balanced team distribution across different locations, avoiding unbalanced subgroup formation (Carton and Cummings, 2012). Most importantly, leaders must work with team members to create an environment of trust and understanding, planning for periodic on-site visits (Jain and Suman, 2015) and cultural training and workshops (Anwar et al., 2019; Jain and Suman, 2015).

Due to time constraints, while conducting this review, we adopted a search criteria and established specific rules, such as filtering by articles published only after 2010. Future research can extend the search criteria to include research before 2010, especially around subgroup studies for more traditional teams. We also suggest future research to include terms such as "subgroups" and synonyms "in-group" and "outgroup" in the criteria. Our query only considered "dispersed" - a new search could also include "distributed." We opt to specifically call out "Scrum" in the query besides "Agile," but future research can also include other Agile frameworks to understand better how those affect subgroup formation. Finally, we believe that more studies are needed around the impacts that different degrees of virtuality, especially entirely virtual teams, have on subgroup formation for Agile teams and how those subgroups influence the team as a whole. We ask that more empirical studies are developed considering those elements.

References

Aarseth, W., Rolstadås, A., & Andersen, B. (2014). Managing organizational challenges in global projects. International Journal of Managing Projects in Business.

Agile Alliance. (2001). Scrum. Available at: <https://www.agilealliance.org/glossary/scrum/>. Accessed on Dec 29th, 2020.

Almeida, F., Miranda, E., & Falcão, J. (2019). Challenges and facilitators practices for knowledge management in large-scale scrum teams. Journal of Information Technology Case and Application Research, 21(2), 90-102.

Alqahtani, A. S., Moore, J. D., Harrison, D. K., & Wood, B. M. (2013). The challenges of applying distributed Agile software development: A systematic review. International Journal of Advances in Engineering & Technology, 5(2), 23.

Alsharo, M., Gregg, D., & Ramirez, R. (2017). Virtual team effectiveness: The role of knowledge sharing and trust. Information & Management, 54(4), 479-490.

Anwar, R., Rehman, M., Wang, K. S., & Hashmani, M. A. (2019). Systematic literature review of knowledge sharing barriers and facilitators in global software development organizations using concept maps. IEEE Access, 7, 24231-24247.

Ashmore, S. (2012). The impact of process on virtual teams: A comparative analysis of waterfall and agile software development teams. Graduate Theses and Dissertations. Iowa State University.

Badiale, M. E. (2020). The dynamics of communication in global virtual software development teams: A case study in the agile context during the Covid-19 pandemic. Master's Thesis, Uppsala University.

Banjo, S., Yap, L., Murphy, C., Chan, V. (2020). The Coronavirus Outbreak Has Become the World's Largest Work-From-Home Experiment. Available at: <https://time.com/5776660/coronavirus-work-from-home>. Accessed on Nov 17th, 2020;

Bezrukova, K., Jehn, K.A., Zanutto, E.L., and Thatcher, S.M.B., "Do Workgroup Faultlines Help or Hurt? A Moderated Model of Faultlines, Team Identification, and Group Performance", Organization Science, 20(1), 2009, p. 35-50.

Bjørn, P., Søderberg, A. M., & Krishna, S. (2019). Translocality in global software development: The dark side of global agile. Human-Computer Interaction, 34(2), 174-203.

Bond-Barnard, T. J., Fletcher, L., & Steyn, H. (2018). Linking trust and collaboration in project teams to project management success. International Journal of Managing Projects in Business.

Borrego, G., Morán, A. L., Cinco, R. R. P., Rodríguez-Elias, O. M., & García-Canseco, E. (2017). Review of approaches to manage architectural knowledge in Agile Global Software Development. IET Software, 11(3), 77-88.

Borrego, G., Morán, A. L., Palacio, R., & Rodríguez, O. M. (2016, August). Understanding architectural knowledge sharing in AGSD teams: An empirical study. In 2016 IEEE 11th International Conference on Global Software Engineering (ICGSE) (pp. 109-118). IEEE.

Bos, N., Olson, J., Nan, N., Shami, N. S., Hoch, S., & Johnston, E. (2006, April). Collocation blindness in partially distributed groups: is there a downside to being collocated?. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 1313-1321).

Carton, A. M., & Cummings, J. N. (2012). A theory of subgroups in work teams. Academy of management review, 37(3), 441-470.

Clerc, V., Lago, P., & van Vliet, H. (2011, August). Architectural knowledge management practices in agile global software development. In 2011 IEEE Sixth International Conference on Global Software Engineering Workshop (pp. 1-8). IEEE.

Cramton, C.D. and Hinds, P.J., "Subgroup Dynamics in Internationally Distributed Teams: Ethnocentrism or Cross-National Learning?", Research in Organizational Behavior, 26, 2004, p. 231-263.

Cruzes, D. S., Moe, N. B., & Dybå, T. (2016, August). Communication between developers and testers in distributed continuous agile testing. In 2016 IEEE 11th International Conference on Global Software Engineering (ICGSE) (pp. 59-68). IEEE.

Daim, T. U., Ha, A., Reutiman, S., Hughes, B., Pathak, U., Bynum, W., & Bhatla, A. (2012). Exploring the communication breakdown in global virtual teams. International Journal of Project Management, 30(2), 199-212.

de Farias Junior, I. H., Duarte, L., De Oliveira, J. P. N., Ari'dnes, R. N., Barbosa, J. F., & de Moura, H. P. (2012, August). Motivational factors for distributed software development teams. In 2012 IEEE Seventh International Conference on Global Software Engineering Workshops (pp. 49-54). IEEE.

Ende, M., Lammermann, R., Brockmann, P., Ayurzana, G. (2013). A virtual, global classroom to teach global software engineering: A Mongolian-German team-teaching project.

2013 2nd International Conference on E-Learning and E-Technologies in Education, ICEEE 2013, 6644379, pp. 229-233.

Gilson, L. L., Maynard, M. T., Jones Young, N. C., Vartiainen, M., & Hakonen, M. (2015). Virtual teams research: 10 years, 10 themes, and 10 opportunities. Journal of Management, 41(5), 1313-1337.

Green, R., Mazzuchi, T. H. O. M. A. S., & Sarkani, S. H. A. H. R. A. M. (2010). Understanding the role of synchronous & asynchronous communication in agile software development and its effects on quality. Journal of Information Technology Management, 21(2), 8-23.

Hacker, J. V., Johnson, M., Saunders, C., & Thayer, A. L. (2019). Trust in virtual teams: A multidisciplinary review and integration. Australasian Journal of Information Systems, 23.

Hinds, P.J., Neeley, T.B., and Cramton, C.D., "Language as a Lightning Rod: Power Contests, Emotion Regulation, and Subgroup Dynamics in Global Teams", Journal of International Business Studies, 45(5), 2014, p. 536-561.

Jain, R., & Suman, U. (2015). A systematic literature review on global software development life cycle. ACM SIGSOFT Software Engineering Notes, 40(2), 1-14.

Jalali, S., & Wohlin, C. (2010, August). Agile practices in global software engineering-A systematic map. In 2010 5th IEEE International Conference on Global Software Engineering (pp. 45-54). IEEE.

Jehn, K. A., & Bezrukova, K. (2010). The faultline activation process and the effects of activated faultlines on coalition formation, conflict, and group outcomes. Organizational Behavior and Human Decision Processes, 112(1), 24-42.

Kaur, H., & Haddad, H. M. (2015). Distributed agile development: A survey of challenges and solutions. In Proceedings of the International Conference on Software Engineering Research and Practice (SERP) (p. 42). The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp).

Lau, D. C., & Murnighan, J. K. (1998). Demographic diversity and faultlines: The compositional dynamics of organizational groups. Academy of Management Review, 23(2), 325-340.

Li, Y., & Maedche, A. (2012). Formulating effective coordination strategies in agile global software development teams. Thirty Third International Conference on Information Systems, Orlando, USA. 1-12.

Liukkunen, K., Lindberg, K., Hyysalo, J., & Markkula, J. (2010, August). Supporting collaboration in the geographically distributed work with communication tools in the remote district SME's. In 2010 5th IEEE International Conference on Global Software Engineering (pp. 155-164). IEEE.

Lous, P., Tell, P., Michelsen, C. B., Dittrich, Y., Kuhrmann, M., & Ebdrup, A. (2018, May). Virtual by design: How a work environment can support agile distributed software development. In 2018 IEEE/ACM 13th International Conference on Global Software Engineering (ICGSE) (pp. 97-106). IEEE.

Manteli, C., Van Den Hooff, B., & Van Vliet, H. (2014). The effect of governance on global software development: An empirical research in transactive memory systems. Information and Software Technology, 56(10), 1309-1321.

Matthiesen, S., Bjørn, P., & Petersen, L. M. (2014, February). " Figure out how to code with the hands of others" recognizing cultural blind spots in global software development. In Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing (pp. 1107-1119).

Morrison-Smith, S., & Ruiz, J. (2020). Challenges and barriers in virtual teams: a literature review. SN Applied Sciences, 2, 1-33.

Moser, A., & Korstjens, I. (2018). Series: Practical guidance to qualitative research. Part 3: Sampling, data collection and analysis. European journal of general practice, 24(1), 9-18.

Mudumba, V., & Lee, O. K. (2010, August). A new perspective on GDSD risk management: agile risk management. In 2010 5th IEEE International Conference on Global Software Engineering (pp. 219-227). IEEE.

Neeley, T. (2021). Remote work revolution: succeeding from anywhere. 1st ed. Harper Business

Nevo, S., & Chengalur-Smith, I. (2011, January). Enhancing the performance of software development virtual teams through the use of agile methods: a pilot study. In 2011 44th Hawaii International Conference on System Sciences (pp. 1-10). IEEE.

Noll, J., Beecham, S., Razzak, A., Richardson, B., Barcomb, A., & Richardson, I. (2017, February). Motivation and Autonomy in Global Software Development. In International Workshop on Global Sourcing of Information Technology and Business Processes (pp. 19-38). Springer, Cham.

Nordbäck, E. S., & Espinosa, J. A. (2019). Effective coordination of shared leadership in global virtual teams. Journal of Management Information Systems, 36(1), 321-350.

O'Leary, M. B., & Mortensen, M. (2010). Go (con) figure: Subgroups, imbalance, and isolates in geographically dispersed teams. Organization Science, 21(1), 115-131.

Paasivaara, M., Behm, B., Lassenius, C., & Hallikainen, M. (2018). Large-scale agile transformation at Ericsson: a case study. Empirical Software Engineering, 23(5), 2550-2596.

Panteli, N., & Davison, R. M. (2005). The role of subgroups in the communication patterns of global virtual teams. IEEE Transactions on Professional Communication, 48(2), 191-200.

Persson, J. S., Mathiassen, L., & Aaen, I. (2012). Agile distributed software development: enacting control through media and context. Information Systems Journal, 22(6), 411-433.

Pflügler, C., Wiesche, M., & Krcmar, H. (2018, January). Subgroups in agile and traditional IT project teams. In Proceedings of the 51st Hawaii International Conference on System Sciences.

Polzer, J. T., Crisp, C. B., Jarvenpaa, S. L., & Kim, J. W. (2006). Extending the faultline model to geographically dispersed teams: How colocated subgroups can impair group functioning. Academy of Management Journal, 49(4), 679-692.

Pries-Heje, L., & Pries-Heje, J. (2011, August). Why Scrum works: A case study from an agile distributed project in Denmark and India. In 2011 Agile Conference (pp. 20-28). IEEE.

Przybilla, L., Wiesche, M., & Krcmar, H. (2018, June). The influence of agile practices on performance in software engineering teams: A subgroup perspective. In Proceedings of the 2018 ACM SIGMIS Conference on Computers and People Research (pp. 33-40).

Schweitzer, L., & Duxbury, L. (2010). Conceptualizing and measuring the virtuality of teams. Information systems journal, 20(3), 267-295.

Scott, C. P., & Wildman, J. L. (2015). Culture, communication, and conflict: A review of the global virtual team literature. In Leading global teams (pp. 13-32). Springer, New York, NY.

Scott, M. E. (2013). “Communicate through the roof”: A case study analysis of the communicative rules and resources of an effective global virtual team. Communication Quarterly, 61(3), 301-318.

Scrum Alliance (2013). The State of Scrum: Benchmarks and Guidelines. Available at <https://www.scrumalliance.org/ScrumRedesignDEVSite/media/ScrumAllianceMedia/Files%

20and%20PDFs/State%20of%20Scrum/2013-State-of-Scrum-Report_062713_final.pdf> Accessed on Nov 17th, 2020.

Scrum Alliance (2018). State of Scrum 2017-2018: Scaling and Agile transformation. Available at <https://resources.scrumalliance.org/Article/state-scrum-2017-2018-report> Accessed on Nov 17th, 2020.

Sharp, H., Giuffrida, R., & Melnik, G. (2012, May). Information flow within a dispersed agile team: a distributed cognition perspective. In International Conference on Agile Software Development (pp. 62-76). Springer, Berlin, Heidelberg.

Sharp, J. H., Ryan, S. D., & Prybutok, V. R. (2014). Global Agile Team Design: An Informing Science Perspective. Informing Sci. Int. J. an Emerg. Transdiscipl., 17, 175-187.

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. Journal of Business Research, 104, 333-339.

Shrivastava, S. V., & Rathod, U. (2015). Categorization of risk factors for distributed agile projects. Information and Software Technology, 58, 373-387.

Stadler, M., Vallon, R., Pazderka, M., & Grechenig, T. (2019). Agile Distributed Software Development in Nine Central European Teams: Challenges, Benefits, and Recommendations. International Journal of Computer Science & Information Technology (IJCSIT) Vol, 11.

Stawnicza, O. (2015). Distributed team cohesion–not an oxymoron. The impact of information and communications technologies on teamness in globally distributed IT projects. International Journal of Information Systems and Project Management, 3(2), 23-39.

Stray, V., Moe, N. B., & Noroozi, M. (2019, May). Slack me if you can! using enterprise social networking tools in virtual agile teams. In 2019 ACM/IEEE 14th International Conference on Global Software Engineering (ICGSE) (pp. 111-121). IEEE.

Tran, M., & Nguyen, H. T. M. (2018). Investigating the interaction of team virtuality, cultural diversity and team member adaptivity in relation to perceived subgroup formation and how it affects team p formation and how it affects team (Master's thesis, Handelshøyskolen BI)

Vallon, R., Strobl, S., Bernhart, M., & Grechenig, T. (2013, June). Inter-organizational co-development with scrum: experiences and lessons learned from a distributed corporate development environment. In International Conference on Agile Software Development (pp. 150-164). Springer, Berlin, Heidelberg.

Van Hillegersberg, J., Ligtenberg, G., & Aydin, M. N. (2011, March). Getting agile methods to work for cordys global software product development. In International Workshop on Global Sourcing of Information Technology and Business Processes (pp. 133-152). Springer, Berlin, Heidelberg.

Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. MIS quarterly, xiii-xxiii.

Webster, J. & Wong, W.K.P. (2008). Comparing traditional and virtual group forms: Identity, communication and trust in naturally occurring project teams. International Journal of Human Resource Management, 19, 41-62.

Wendling, M., Oliveira, M., & Maçada, A. C. G. (2013). Knowledge sharing barriers in global teams. Journal of Systems and Information Technology. 15(3), 239-253.

Zahedi, M., Shahin, M., & Babar, M. A. (2016). A systematic review of knowledge sharing challenges and practices in global software development. International Journal of Information Management, 36(6), 995-1019.