What makes a successful team?

Theoretical background and applicability of a team interaction questionnaire in two different contexts

Eero Antero Mikkelä

Master's Thesis

Psychology

Faculty of Medicine

October 2019

Supervisor: Silja Martikainen

Research project: HUMEX

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

Tiedekunta – Fakultet – Faculty	Koulutusohjelma – Utbildingsprogram – Degree Programme
Faculty of Medicine / Department of Psychology and Logopedics	Psychology

Tekijä – Författare – Author
Eero Mikkelä

Työn nimi – Arbetets titel – Title
What makes a successful team? Theoretical background and applicability of a team interaction questionnaire in two different contexts

Oppiaine/Opintosuunta – Läroämne/Studieinriktning – Subject/Study track
Psychology

Työn laji – Arbetets art – Level	Aika – Datum – Month and year	Sivumäärä – Sidoantal – Number of pages
Master's Thesis	October 2019	59 + 1 appendix (5 p.)

Tiivistelmä – Referat – Abstract

Objective. In today's workplaces an increasing number of tasks is completed in teams. Hence, to understand the psychological processes affecting interaction within and success of teams in different environments is of great importance. To study these processes, a new TT10 questionnaire (84 questions, 10 separate subscales) was created. The aim of this thesis was to study the validity and reliability of the TT10 by studying multiple literature-based hypotheses about connections between the subscales of the TT10 in two separate studies.

Methods. In study 1 (n=49) 10 teams from two Finnish technology companies filled the TT10 and basic demographic information online. In study 2 (n=124) there were 62 pairs consisting of an employee of a Finnish insurance company and a customer. The counterparts in each pair were anonymously in contact with each other in an online chat for 20 minutes during which their task was to solve puzzles together. After the experiment they were asked to review the interaction of their pair with a shortened version (27 questions) of the TT10.

Results. Almost all of the hypotheses gained support. Different subscales were in connection with each other mostly in the hypothesized ways.

Conclusions. The most promising subscales of the TT10 were psychological safety, social cohesion, collaboration, and co-flow. However, limitations such as a very limited number of participants in study 1 made some of the results a bit unclear. However, according to these preliminary results, the TT10 seems to be a promising questionnaire that still needs fine-graining.

Avainsanat – Nyckelord – Keywords
Work teams, TT10, construct validity, organizational psychology, collaboration, co-flow, social cohesion, mentalizing, emotion contagion, social presence/attention, psychological safety, positive orientation, equality, shared reflection

Ohjaaja tai ohjaajat – Handledare – Supervisor or supervisors
Silja Martikainen

Säilytyspaikka – Förvaringställe – Where deposited
Helsingin yliopiston kirjasto – Helda / E-thesis (opinnäytteet) ethesis.helsinki.fi

Muita tietoja – Övriga uppgifter – Additional information
Research project: HUMEX

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

Tiedekunta – Fakultet – Faculty	Koulutusohjelma – Utbildingsprogram – Degree Programme	
Lääketieteellinen tiedekunta	Psykologia	
Tekijä – Författare – Author		
Eero Mikkelä		
Työn nimi – Arbetets titel – Title		
What makes a successful team? Theoretical background and applicability of a team interaction questionnaire in two different contexts		
Oppiaine/Opintosuunta – Läroämne/Studieinriktning – Subject/Study track		
Psykologia		
Työn laji – Arbetets art – Level	Aika – Datum – Month and year	Sivumäärä – Sidoantal – Number of pages
Pro gradu -tutkielma	Lokakuu 2019	59 + 1 liite (5 ls.)

Tiivistelmä – Referat – Abstract

Tavoitteet. Nykyajan työympäristössä työ tehdään yhä enenevissä määrin työryhmissä. Tästä syystä ryhmien vuorovaikutukseen ja menestymiseen vaikuttavien psykologisten tekijöiden ymmärtäminen eri ympäristöissä on tärkeää. Näiden tekijöiden tutkimiseksi kehitettiin TT10–kysely, joka koostuu 10 alaskaalaa muodostavasta 84 kysymyksestä. Tämän tutkielman tavoite oli kahden erillisen tutkimuksen avulla tarkastella TT10–kyselyn reliabiliteettia ja validiteettia muodostamalla useita tutkimustietoon perustuvia teoreettisia hypoteeseja TT10:n alaskaalojen yhteyksistä.

Menetelmät. Ensimmäisessä tutkimuksessa (n=49) kahden suomalaisen teknologiayrityksen työntekijät täyttivät TT10:n ja taustatietoja koskeneen kyselyn internetissä. Toisessa (n=124) 62 paria, jotka koostuivat suomalaisen vakuutusyhtiön työntekijästä ja asiakkaasta, olivat toisiinsa yhteydessä anonyymisti 20 minuutin ajan internet-pohjaisessa keskustelusovelluksessa. Työparin tehtävänä oli ratkaista yhdessä päättelytehtäviä. Tehtävien teon jälkeen he arvioivat parinsa työskentelyä TT10:n lyhennetyllä versiolla, joka sisälsi 27 kysymystä.

Tulokset. Lähes kaikki hypoteesit saivat tukea. TT10:n alaskaalat olivat yhteydessä toisiinsa pääosin oletetuilla tavoilla.

Johtopäätökset. Kaikkein lupaavimmat alaskaalat olivat psykologinen turvallisuus, sosiaalinen koheesio, yhteistyö ja co-flow. Tutkimusten rajoitukset, kuten ensimmäisen tutkimuksen erittäin pieni otoskoko, jättivät osan tuloksista kyseenalaisiksi. Näiden alustavien tulosten perusteella TT10 on kuitenkin lupaava kysely, joka vaatii vielä hienosäätöä.

Avainsanat – Nyckelord – Keywords
Työryhmät, TT10, rakennevaliditeetti, organisaatiopsykologia, yhteistyö, co-flow, sosiaalinen koheesio, mentalisaatio, emootioiden tarttuminen, sosiaalinen läsnäolo/tarkkaavuus, psykologinen turvallisuus, Positiivinen suhtautuminen, Tasa-arvo, Reflektointi

Ohjaaja tai ohjaajat – Handledare – Supervisor or supervisors
Silja Martikainen

Säilytyspaikka – Förvaringställe – Where deposited
Helsingin yliopiston kirjasto – Helda / E-thesis (opinnäytteet) ethesis.helsinki.fi

Muita tietoja – Övriga uppgifter – Additional information
Tutkimusprojekti: HUMEX

Table of Contents

1. Introduction

Teams are a nonseparable part of working culture in today's world (e.g. Edmundson, 1999; West, Patera, & Carsten, 2009). Teams have become one of the corner stones for learning and accomplishing work in organizations (Knapp, 2010), and team-based collaboration has become an important ingredient of organizational performance (Edmondson & Lei, 2014; Goleman, 1998; West et al., 2009). This is partly due to the fact that people have increasingly specific skill sets and working environments have become complex and changeable (Edmondson & Lei, 2014). One individual cannot hold all the information necessary for accomplishing the task of a team (Goleman, 1998). Thus, it could be argued that only a few skills are more essential for the modern organization than the ability to work effectively in groups (Goleman, 1998; Johnson & Johnson, 2003).

In this regard, it is not a surprise that teams and teamwork have stirred up a lot of different kinds of psychological theoretical models and study lines. Many aspects of teamwork have been considered. These include, among others, psychological safety (e.g. Carmeli, Brueller, & Dutton, 2009; Carmeli & Gittell, 2009; Edmundson, 1999), collaboration and cooperativeness (e.g. Edmondson & Lei, 2014; Goleman, 1998; West et al., 2009), and empathy in its different forms like resonance, emotion contagion, and mentalizing (e.g. Biocca, Harms, & Gregg, 2001; de Waal, & Preston, 2017; Luca & Cowan, 2001).

Combining these different perspectives, it has been said that successful teamwork 'is a synergistic process that relies upon all team members to contribute and participate in order to promote and nurture a positive, effective team environment' (Luca & Cowan, 2001, p. 4). To be able to be a good team member, individuals have to possess a wide variety of skills and certain amount of flexibility in order to adapt to different and always changing social situations. To be able to see others as persons rather than just working partners is an important factor in this process. As Luca and Cowan argue in their article, teams that 'care about each other at a personal and professional level are more likely to be successful than teams that ignore the importance of the relationship between positive interpersonal relationships, professional relationships and goal achievement' (p. 5).

To address this multifaceted nature of successful team interaction, a brand new TT10 (short for team-work questionnaire 10, "työtiimikysely-10" in Finnish) questionnaire was created. The TT10 in its current and original form includes ten separate subscales that are hypothesized to cover different psychological factors related to team-interaction. The aim of the TT10 questionnaire was 1. to combine multiple perspectives to successful interaction of work teams into the same questionnaire, and 2., when sufficient amount of data is available, to study how these different constructs are related to each other, and 3. to study which of these constructs are most relevant to team success in different kinds of communication pathways (face-to-face and computer mediated communication, e.g. chat, video calls and so on).

Because the TT10 is a new questionnaire that is used in this thesis for the first time, its subscales could still be altered. Thus, the main aim of the two studies presented in this thesis is to study the initial usefulness, reliability, and validity of the TT10 questionnaire and its different subscales and come up with suggestions for its further development if necessary. Do the proposed dimensions truly address the constructs they are supposed to address? These questions are addressed by forming hypotheses based on the previous research. Thus, main focus in this thesis is in theoretical validation of the TT10. This is especially the case in the first study of this thesis, later called as study 1, in which the TT10 is studied in the context of work teams from two Finnish technology companies.

At the same time some of the concepts used in the questionnaire are initially established in Finnish organizational research. To the author's knowledge, there has not previously been this comprehensive a research concerning the constructs addressed in this thesis in Finland. Moreover, some of the questionnaires and theories addressing these constructs are over two decades old. Therefore, it is not clear how well they are applicable to the modern-day working environment, in which computer mediated communication plays an increasingly important role.

In addition to teamwork, computer mediated communication systems are also a vital part of modern customer service situations. Thus, in the second study presented in this thesis, later called as study 2, there is also an interest in the applicability of the TT10 questionnaire in a customer service situation and understanding how the factors supporting successful team interaction might translate into the computer mediated customer interface. These insights are hoped to be gained even though some of the constructs that address teamwork take time to develop. For example, psychological safety requires trust between team members in order to create a feeling of safety within a team (Edmondson & Lei, 2014). In a brief customer interface situation these qualities could be hard to achieve. Moreover, some constructs like social reflection, that means the extent people reflect their actions with their team members (West, 1996), are not well suited for short customer situations. However, good customer service professionals should be able to create a supportive and confidence creating environment. This holds especially in situations where there is a clear goal to be achieved, as is the case in study 2 (see 6. Methods). Thus, the TT10 should be applicable to this situation also.

1.1. Theoretical background of the TT10 questionnaire

Social interaction and the mechanisms guiding our social relationships are mostly evolution based and are among of the most basic elements of everyday life (Klein, Shepherd, & Platt, 2009). There is often a need or a possibility to interact with other people (Hessels, Holleman, Kingstone, Hooge, & Kemner, 2019). We also tend to assess the quality of our lives partly by social relationships (Klein et al., 2009). We orient towards important individuals and tend to focus on the same objects in our environment as them.

Taking into account the importance of social interactions, it is not a surprise that there are many different ways to analyze these social processes, and one subscale hardly can cover all of the variation people encounter every day. Thus, ten different, but not necessarily totally distinct, elements of social interaction in groups were taken into the TT10 questionnaire. The selection was based on relevant research on successful collaboration. These ten subscales are supposed to reflect different aspects of social interaction vital to the success of teams. In the following section, these ten subscales are introduced; the specific questions pertaining to each subscale can be found in the Appendix 1. Theoretical backgrounds of these constructs are discussed, and original questionnaires from which some of the questions were taken are mentioned. The hypotheses concerning the subscale in question are also introduced for the two studies presented in this thesis. Included here is also an eleventh subscale of what is achieved through interaction (later referred as the achievement scale), in which the teams' self-evaluations of the successfulness of their interaction are addressed. Even though it's not counted as one of the principal subscales of the TT10, it is worth mentioning due to its importance in the studies presented in this thesis.

1.1.1. Collaboration

Team-based collaboration is an important factor in organizational performance (Edmondson & Lei, 2014; Goleman, 1998; West et al., 2009). Collaboration is based on good cooperation (Johnson & Johnson, 2003). Cooperation is defined as an interdependence and shared goals between individuals (Deutsch, 1949). In this sense, cooperative groups are internally more consistent, and group members need each other more to achieve their goals when compared to competitive groups.

In their review of multiple studies of collaboration, Johnson and Johnson (2003) argue that cooperative groups tend to perform higher, make better decisions, and solve problems better than individuals or competitive groups. This holds especially for tasks that have specific goals or procedures, like mathematical or verbal tasks. On the other hand, on more abstract tasks, like brainstorming, individuals may perform just as well as groups.

According to Johnson and Johnson (2003) similar results are also found when studying individual performance. Those individuals who work in either cooperative groups or competitive groups perform better than persons who work just individually. Thus, working in a group is also beneficial for individuals themselves and leads for example to better knowledge learning and higher individual proficiency.

Moreover, combining the results of multiple studies, Johnson and Johnson (2003) state that working in cooperative teams is connected to a better liking of other members, to a greater interpersonal attraction within team members, and to a greater amount of social support than within groups that do not work cooperatively. Working in cooperative groups is also connected to a better psychological health and a greater self-esteem than competitive or individual work.

Johnson and Johnson (2003) argue that successful cooperation is born from five different sources: positive interdependence, individual accountability/personal responsibility, promotive (face-to-face) interaction, social skills and group processing. However, in the TT10 questionnaire these domains were not specifically addressed. Instead, there were questions concerning the happiness of the individual when their team members succeeded, feelings of individuals when getting help from the team members when needed, and the competitive nature of the working habits in the teams (see appendix 1).

Based on the argumentation of the review of Johnson and Johnson (2003), it can be hypothesized that those teams that rate their collaboration higher tend also to evaluate the successfulness of their interactions higher. In addition, it should be pointed out that according to Johnson and Johnson, collaborative teams outperform competitive teams especially in tasks that have specific goals or procedures. Thus, it would be presumable that correlations between collaboration and the achievement scale are higher in study 2, in which the task is clearly defined, unlike in study 1 in which the tasks of the teams are not specified.

1.1.2. Co-flow

Flow is a term most usually connected to psychology of individuals. Its development started in 1970's by Csikszentmihalyi (as recalled in Csikszentmihalyi, Abuhamdeh, & Nakamura, 2014), and it came to full fruition in the book Flow: The psychology of optimal experience (1990) by Csikszentmihalyi. Flow is described as a subjective state in which people forget about everything else than the task they are involved in right at the time (Csikszentmihalyi et al., 2014). Their attention is fully focused on the task and they do not feel fatigue or notice the passing of time. The state of flow can be achieved in various situations and in doing various tasks ranging from reading a book to competing in a highly demanding situation as a professional athlete. Individually, in many sources flow is connected to creativity and person's intrinsic motivation (e.g. Csikszentmihalyi et al., 2014; MacDonald, Byrne, & Carlton, 2006).

There are three conditions recognized essential for a flow experience: a clear set of goals, a balance between perceived skills and perceived challenges, and clear and immediate feedback (Csikszentmihalyi, et al., 2014). Goals help to give structure to actions and guide attention in the right tasks. A balance between perceived skills and perceived challenges is achieved when the task at hand is just sufficient for the person's skills. In that way, accomplishing the task requires all the effort and attention one has but does not exceed his capabilities. Immediate feedback helps to guide and orientate the future actions. When given exact information on what the individual did right or wrong, it is easier to correct or continue those actions. Thus it becomes clearer, in which way one should continue with the task.

Similar concepts have also been described in the context of groups (Cross & Parker, 2004). Even though not talking about the term flow per se, Cross and Parker proposed that highly effective and immersive, in their terms 'energetic', groups need a clear goal, a possibility for an individual to contribute and use one's

ideas, and a perception of progress, partly in a form of clear feedback. However, Sawyer (2007) is one of the first to suggest that group flow is also a valid construct. In his book Group Genius, Sawyer proposed that group flow is ‘a peak experience, a group performing at its top level of ability’ and while a group is in the state of flow, its ‘activity becomes spontaneous, and the group acts without thinking about it first’ (p. 49).

Sawyer (2007) continues his conceptualization of group flow by defining 10 subcategories that lead to it. These are: having a goal, close listening, complete concentration, blending egos, equal participation, familiarity, communication, moving it forward, and the potential for failure. Duncan and West (2018) argue that, because of the overlapping of these 10 components, they could be summarized in three groups: vision, ownership and contribution, and communication. Duncan and West propose that using these three groups, it is easier to study and talk about flow as a quality of groups. These three groups roughly estimate the three groups proposed for individual flow by Csikszentmihalyi (Csikszentmihalyi et al., 2014). Vision includes clear goals and the potential to failure, ownership and contribution equal to a balance between perceived skills and perceived challenges, and communication is a natural counterpart of clear feedback.

Partly based on these three categories, co-flow was addressed in the TT10 questionnaire by questions about the immersion of a group’s interaction, clarity of goals of a team and an enjoyment that interaction with the team creates. Questions focused especially on the emotional side of collaboration. There were no specific questions about feedback given in groups, because this is separately addressed in shared reflection subscale of the TT10.

All in all, three hypotheses were formed for this subscale. Firstly, co-flow is hypothesized to be connected to a better self-rated interaction and higher rates of achievement within a team. Secondly, if co-flow really is an immersive experience and based on the nature of questions concentrating on the emotional side of interaction, it’s hypothesized that co-flow correlates with the subscales of social cohesion (which emphasizes connectedness to other team members and enjoyment of the interaction, see 1.1.3. Social cohesion in this thesis) and social presence/attention, as if the person is immersed in the task at hand, they are more likely to focus also to the others who are needed to accomplish it. Thirdly, because clear feedback and good communication are proposed to be important factors leading to the emergence of co-flow (Sawyer, 2007), shared reflection subscale of the TT10 and co-flow are hypothesized to correlate positively with each other.

1.1.3. Social Cohesion

Cohesion is defined as a bond that is based on a combination of social and motivational forces existing between group members (Beal, Cohen, Burke, & McLendon, 2003). These forces include, among others, attraction, task commitment, and group pride. Cohesion is thought to be one of the forces that increase group performance. It is presumed that the stronger the cohesion between group members, the greater the

productivity and successfulness of the group. That is because in cohesive groups the members know each other better, which creates an urge to perform better while completing the task.

According to the results of a meta-analysis by Beal and colleagues (2003), cohesion is an important factor when the main requirement of the group is efficiency (doing something as well as possible with the available resources, i.e. using as little time and money as possible), as opposed to effectiveness (achieving something no matter the costs, i.e. completing a task, winning a championship). Commitment, on the other hand, is the most important factor for both types of teams. Moreover, cohesion is more connected to the performance of a group conceptualized as a behavior instead of an outcome. The flip side of this is that cohesion is not so much connected to the results of the team-work itself but, instead, it is more connected to the behavior and working habits that can partly lead to better results. However, it should be noted that there are more aspects that modify the result than just the actions of the group on their own. Cohesion is thus not an optimal variable when predicting the group's potential for success. This is still not a problem for the first study of this thesis, where the successfulness of the teams was addressed as a behavior instead of a measurable performance outcome (see 1.1.11 What is achieved through interaction).

In the TT10 questionnaire Social cohesion was addressed with questions concerning enjoyment of the interaction with and company of the other team members, and feeling of connectedness towards the other members. Two of the questions have been used previously in the Psychological Involvement subscale of a social presence questionnaire (de Kort, Ijsselsteijn, & Poels, 2007) and two of the questions were based on a General Satisfaction Index from a study by Wageman, Hackman and Lehman (2005). In line with the results of Beal and colleagues (2003), it is hypothesized that higher social cohesion is connected to higher self-evaluated effectiveness of a group. That is, teams that rate their social cohesion higher tend to perform better and have better ways of working.

1.1.4. Resonance/emotion contagion

Emotion contagion is a term used to describe people's actions of mimicking or otherwise empathizing other people's feelings (de Waal & Preston, 2017). It does not necessarily lead to an action like helping (Preston & de Waal, 2002). Emotion contagion is seen to be a part of empathy (de Waal & Preston, 2017; Preston & de Waal, 2002), and is a comprehensive way to address the mechanisms of empathy in a questionnaire-based study. It is easier for people to rate the empathy related behavior of their team members, and thus emotion contagion, than to guess the inner states of the others.

Empathy can be defined as any process that reflects the individual's ability to understand and share feelings and mental states of the others (de Waal & Preston, 2017). Empathy is generated by different mental, neural, and personal representations. It helps to adopt the perspective of other people and, hence, to understand their actions and reasons behind their inner states. In any given moment, empathy is dependent

upon the observer's attention to the situation, their motivation to understand the actions of others, and their relevant personal past experiences.

Emotion contagion is a subclass under the hypernym of empathy (Preston & de Waal, 2002). In their model of empathy, Preston and de Waal propose that empathy should be defined as a process rather than a response. In this way, several approaches concerning empathy could be taken into account. These include both cognitive and affective mechanisms. Thus, empathy includes both individual's inner state and feelings aroused by the situation, and their possible reactions to it. Emotion contagion is classified under the affective mechanisms of this model.

Empathy is also seen as a part of emotional intelligence (Goleman, 1998). Emotional intelligence refers to "the capacity for recognizing our own feelings and those of others, for motivating ourselves, and for managing emotions well in ourselves and in our relationships" (pp. 317, Goleman, 1998). Other subscales of emotional intelligence are: self-awareness, self-regulation, motivation, and social skills. For a more throughout discussion of these other categories the reader is advised to familiarize with Goleman's book (1998) or an article by Luca and Cowan (2001). For this discussion, it is relevant to notice that higher emotional intelligence has been connected to more successful teamwork and lower emotional intelligence to dysfunctional teamwork (Luca & Cowan, 2001). Thus, high empathy, and higher emotion contagion, among other components of emotional intelligence, arguably improve teamwork.

The importance of empathy in successful teamwork is demonstrated by the observations of Luca and Cowan (2001). In their case study of university students working on a group assignment they showed that in functional teams the team members showed more appreciation to the diversity of other members and felt themselves more supported. On the other hand, in dysfunctional teams the team members didn't try to understand the feelings of the other members and the reasons behind them, the members saw the inadequacies of the others as deficiencies that cannot be helped, and some of the members felt alienated from their teams.

Emotion contagion was enquired in the TT10 with four questions adapted from the Empathy scale of The Networked Minds measure (Biocca et al., 2001). The Networked Minds measure is a questionnaire developed to address interconnectedness people experience with each other while using virtual environments. It is a good fit for today's working environments, in which communication is not solely face-to-face based but happens in growing amounts also via internet, phone calls, social media, and e-mails. In addition to these four questions, there were two questions about motoric reflectiveness, or resonance, between the group members in the TT10.

Based on previous results (Luca & Cowan, 2001), it is hypothesized that emotion contagion is related to greater group achievement and better group functioning. Regarding the remark that empathy, and thus emotion contagion, is dependent on individual's current attention to the situation (de Waal & Preston,

2017), it was also hypothesized that emotion contagion is correlated with the scale of social presence/attention.

In addition to the TT10 questionnaire, there were also other questionnaires used in study 2 of this thesis that measured constructs related to emotion contagion and are worth mentioning here: The Reading the Mind in the Eyes (RMIE) test (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001) and the Interpersonal Reactivity Index (IRI) (Davis, 1980). They were used to gain more insight in the theoretical validity of the emotion contagion scale. These questionnaires are better discussed in the methods section. Here it is sufficient to say that especially emotion contagion is hypothesized to be connected to the RMIE questionnaire as emotion contagion refers to people's skills in mimicking or otherwise feeling what others are feeling (de Waal & Preston, 2017). On the other hand, the RMIE is a questionnaire where the participants have to detect the right word for an emotion depicted in pictures of different individuals' eyes. Thus, emotion contagion should be closely related to this questionnaire. Moreover, the empathic concern scale of the IRI addresses the ability to feel warmth, compassion and concern when observing other individuals and their actions. This scale is hypothesized to be related to the emotion contagion scale.

1.1.5. Mentalizing

Mentalizing, like resonance/emotion contagion above, is thought to be a part of empathy (de Waal & Preston, 2017). As described in the previous section, there are two mechanisms included in empathy: cognitive and affective mechanisms. While resonance/emotion contagion belonged to the realm of affective mechanisms, mentalizing, on the other hand, is a cognitive activity. Simply stated, mentalizing is the ability "of conceptualizing other creatures (and oneself) as loci of mental life" (pp. 2, chapter 1; Goldman, 2006). Other way to describe mentalizing is that it is an ability to see the thoughts, feelings and beliefs of our own and those of the others as distinct from reality (Frith & Frith, 2003). Other terms used in scientific literature for mentalizing are mind reading (e.g. Gallese & Goldman, 1998) and theory of mind (e.g. Singer, 2006).

Implicit mentalizing skills develop during the first years of life (Frith & Frith, 2003). Despite of its early development, and thus quite an essential part in a human life from early on, mentalizing does not always affect people's actions. Sometimes people may temporarily lose awareness of the minds of others and can treat each other in an objectifying manner (Bateman & Fonagy, 2013). Problems with mentalizing can lead for example to psychopathology like borderline personality disorder (Bateman & Fonagy, 2013). Taking these remarks into account, mentalizing seems to be an important part of human life and problems in it could lead to severe consequences.

Despite its importance in human life in general (e.g. Bateman & Fonagy, 2013, Sharp & Fonagy, 2008), mentalizing has not stirred a lot of research in the psychology of work teams per se. One of the few results has suggested that better skills in mind reading is connected to greater collective intelligence in groups

(Woolley, Chabris, Pentland, Hashmi, & Malone, 2010). Collective intelligence was defined 'as the general ability of the group to perform a wide variety of tasks' (pp. 687) in this study. Women also scored higher in measures of mentalizing than men, and according to Woolley and colleagues, this was in line with previous results. However, they did not elucidate the studies they were referring to.

Another result in this area comes from Engel and colleagues (Engel, Woolley, Jing, Chabris, & Malone, 2014). They studied the role of collective intelligence in online-based teams versus face-to-face teams. They examined the role mentalizing skills play in the performance of face-to-face and online groups, and showed that mentalizing skills played an equally important role in both conditions. Thus, mentalizing skills seem to be important also in situations where subtle nonverbal cues are not as well available as they are in face-to-face situations. The conclusion of Engel and colleagues was that mentalizing abilities are 'a significant determinant of group collective intelligence even when, as in many online groups, the group has extremely limited communication channels'.

Therefore, based on these two studies (Woolley et al., 2010; Engel et al., 2014), it is hypothesized that teams having greater mentalizing scores perform better. This is equal to saying that collective intelligence (as defined by Woolley and colleagues (2010)) is higher in those teams. Furthermore, it is also hypothesized, based on the results by Woolley and colleagues (2010) that women have higher scores in mentalizing than men. This hypothesis could be proposed because the questions in the mentalizing scale of the TT10 are subjective in their nature. Thus, each participant's answers reflect just their own self-reviewed qualities rather than those of others. Some differences between groups could also arise from this same remark, as each team member had their own unique scores on this scale. Thus, it is hypothesized that teams that have more women than men score higher on the achievement scale than other groups.

In addition, as was the case with the emotion contagion scale, hypotheses about the connections between the mentalizing scale and the RMIE and the IRI questionnaires were formed. Mentalizing scores are hypothesized to correlate to the RMIE scores, because if participants made interferences on each other's emotions in pictures, they should be able to tell that other people have their unique perspectives and feelings, as was stated in a definition of mentalizing by Goldman (2006). Moreover, the perspective-taking scale of the IRI questionnaire addresses the ability to see other people's perspectives in real life situations and is thus closely related to the mentalizing scale by its definition. Hence, mentalizing and perspective-taking scores are hypothesized to correlate with each other.

1.1.6. Social presence/attention

One of the most usual ways to address social attention in scientific literature has been to study the direction of gaze. After all, looking at the others is a naturally emergent feature in our everyday life (Klein et al., 2009). In addition, when the others are not looking at us, it may confuse us and create a sense that something is wrong.

According to studies, humans have an innate expectation that other people are looking at them (Mareschal, Calder, & Clifford, 2013). This holds especially in situations where people don't see the eyes of the other, for example when someone is using sunglasses (Mareschal et al., 2013). In these moments of obscurity, people are prone to suppose that the other person is looking at them. Moreover, when people see the others looking at some direction, they are prone to gaze at the same direction (Deaner & Platt, 2003). Deaner and Platt argued that this is most likely an evolution-based feature, as monkeys are known to act in a same manner. It is also noteworthy that when people are listening to the others, they tend to look at them more than while they are speaking themselves (Hessels et al., 2019).

However, despite the importance of the effect of gaze in the social attention, it was a difficult aspect to be properly taken into account in this study. People may not consciously pay attention to the direction of the gaze of the others. Moreover, gaze does not necessarily imply a true focus on what the other is saying. Thus, in the TT10 questionnaire social presence and attention were attended with questions about the other's listening of the individual's speech. There were also questions about one's own attention to the others. After all, social interactions between people create a constant flow of information that needs to be perceived and acted upon (Hessels et al., 2019). Without paying attention to what the others are saying, the quality of interaction cannot be very satisfactory.

However, people are not always automatically paying attention to the others. For example, a very common finding in social attention studies is that those who have power tend to overlook lower ranked individuals and use stereotypes when thinking of them (e.g. Overbeck & Park, 2006; Rodriguez-Bailon, Moya, & Yzerbyt, 2000). This one-sided point of view has been challenged for example by Overbeck and Park (2006). Their proposition is that those who have power pay attention to the others flexibly. According to their argumentation, the level of social attention depends on the situation. When people are useful for a task at hand, powerful persons will attend to them carefully and systematically. On the other hand, when the others are not so useful or are irrelevant to accomplishing the goals, those with power may overlook them and pay more close attention to matters that are more instrumental for the task at hand at the time.

Thus, when addressing social attention in groups, at least two different aspects should be taken into account: 1. What is the purpose and the specific task of the group? How important is each member for accomplishing the goal the group has? And 2. How equally power is distributed among the members of the group? It should yet be pointed out that when group members are equal, there might still be some degree of variance between attention given to the others (Overbeck & Park, 2006).

In spite of this, one of the hypotheses, for the two studies of this theses, was that those groups that score the lowest in equality show low levels of social attention, too. This is in line with findings that people with power tend to overlook lower ranked individuals (e.g. Overbeck & Park, 2006; Rodriguez-Bailon et al., 2000). This seems to be the most promising way to approach the subject in this thesis, even though in study 1 it was

not addressed, how much every team member could affect their team's success and decision making. Thus, hierarchical structures of the teams were not clear. It is still presumable that in those teams in which self-rated equality is the lowest, attention is not divided to all the team members in the same ways, and not all the members are seen as equally important for accomplishing the team's tasks.

The second hypothesis about social presence was that social presence and mentalizing scores are correlated. This is due to the fact that without listening to the others, people cannot truly understand them, and their message and values behind that message. They might fail to see the others as individuals if they do not pay attention to them.

1.1.7. Psychological safety

Psychological safety refers to the perception that taking interpersonal risks is allowed within a team (Edmondson, 1999). While team members feel safe in their team, they may admit mistakes, seek help, and discuss about each other's accomplishments more openly (Edmondson & Lei, 2014). Psychological safety may also enhance the possibilities for individuals to speak up more often. Thus, psychological safety can create possibilities for organizations to improve their actions and solve problems through open interaction.

However, psychological safety requires confidence that the other members won't judge the member who speaks openly (Edmondson, 1999). Confidence in turn requires trust and mutual respect among team members. In other words, trust is needed to establish an environment where people can discuss openly. That in turn helps to develop and maintain psychological safety.

The exact mechanisms by which psychological safety is built or destroyed are not very well known or studied (Edmondson & Lei, 2014). One proposition is that high-quality relationships, as described in terms of shared goals, shared knowledge, and mutual respect between team members, create a safe environment where individuals feel free to express also their weaknesses (Carmeli et al., 2009; Carmeli & Gittell, 2009). This in turn creates psychological safety. However, both Carmeli and colleagues (2009) and Carmeli and Gittell (2009) used cross-sectional research methods. Thus, it is not clear whether high-quality relationships created a feeling of psychological safety or if the high-quality relationships were created by the already existing climate of psychological safety that could have, for example, aroused from the predispositions of those participating to express themselves openly.

According to Edmondson (1999), psychological safety is more like a tacit feature of a team than an openly shared norm. It is noteworthy that according to Edmondson's argumentation, previous findings have shown that psychological safety is not the same thing as team cohesiveness, which is also considered in this thesis separately (see 1.1.3.). Cohesiveness can make it harder for individual members to speak out loud their concerns about the team, and thus the sense of psychological safety could decrease.

The benefit of psychological safety is that it can lead to improvement and understanding of the learning process of the teams, when members can freely talk about their errors without being embarrassed (Carmeli et al., 2009; Edmondson, 1999; Edmondson & Lei, 2014; Kostopoulos & Bozionelosa, 2011). This way, they allow the teams to learn from their mistakes and improve their performance by avoiding the same mistakes as a group in the future (Carmeli & Gittell, 2009; Edmondson, 1999). Thus, psychological safety doesn't improve performance in a straight-forward fashion, but rather through organizational learning (Edmondson, 1999). Psychological safety is especially effective in settings where solution is not clear and team needs to experiment and use their innovation (Kostopoulos & Bozionelosa, 2011). In these settings, psychological safety is an eligible way to speedily enhance the team's motivation to engage in learning activities.

In their meta-analytical review of the studies of psychological safety, Edmondson and Lei (2014) confirmed that team's performance and learning behavior are connected to psychological safety. They note that according to multiple studies this connection is theoretically logical, especially in situations where the task at hand is not clearly defined or the team needs to make some adjustments or collaborate to accomplish it. This in turn leads to learning and in the need for new innovations. Thus, while the importance of innovation and learning is growing in today's organizations, psychological safety is an increasingly important and significant phenomenon for both science and organizations themselves.

Finally it should be noted that psychological safety is most appropriately studied as a feature of a group, not that of an organization (Edmondson & Lei, 2014). According to Edmondson and Lei, it is a common finding in different studies that groups vary in their level of psychological safety even within the same organizations. Thus, those who work more closely together tend to rate psychological safety in a similar fashion when compared to those whose work relationships are not so close.

In the TT10 questionnaire psychological safety was addressed with 5 of 7 questions of the psychological safety scale of Edmondson (1999). Edmondson's scale is the most used scale in the research of psychological safety (Edmondson & Lei, 2014). In addition, 5 new items were added to the TT10's psychological safety scale. These questions consider the individual's sense of safety to propose new but unfinished ideas, to compliment the ideas of the other team members, and to admit their lack of knowledge of something to the others. Thus, these new items were addressing the tangible features of the psychological safety that are part of everyday work environment.

The first hypothesis for psychological safety is based on the notion of Edmondson (1999): teams with the greatest cohesiveness are not necessarily the best rated in terms of psychological safety. Thus, to study whether this statement holds in this study, it was hypothesized that the scores on psychological safety of the teams and the social cohesion scales are not strongly correlated. The second hypothesis is that teams with the

highest scores on psychological safety also rate their performance better. This is in line with previous results (Edmondson, 1999; Edmondson & Lei, 2014).

1.1.8. Positive Orientation

In the positive orientation subscale of the TT10 questionnaire, there are seven questions concerning trust, appreciation, and encouragement that an individual feels towards the others and feels gaining as a team member. In addition, one of the questions asks the individuals whether they should help other members in need even if that would put themselves in a disadvantaged position. Trust and mutual respect among team members has been described as a vital element to psychological safety (Carmeli et al., 2009; Carmeli & Gittell, 2009; Edmondson, 1999). Moreover, in their study Singh, Winkel and Selvarajan (2013) argument that positive orientation towards the other team members is positively associated with psychological safety of a team. This connection shows itself especially in helping behaviors directed towards the other team members and the organization.

All in all, taking these results into consideration, this subscale is hypothesized to be closely connected to the subscale of psychological safety. However, it gives some closer insights into the working habits and interpersonal relationships of the team, and can thus be interesting for the team members themselves and have its value as a descriptive entity. However, no specific hypotheses were formed for the connections between the positive orientation and team's results and achievements. Thus, it will only be hypothesized that positive orientation and psychological safety are closely connected.

1.1.9. Equality

Heterogeneity in the characteristics of the team members is an important factor when addressing the performance and decision making of the team (Bunderson & Van der Vegt, 2018). Differences between team members can be both horizontal and vertical. Horizontal differences are those that are based on the specializations, gender, social categories, or other background features of the team members. Vertical differences are, on the other hand, based on hierarchies of power, status, prestige, and privilege. It does not matter whether these hierarchies are formal or informal in nature, they still affect the decision making of the teams.

Bunderson and Van der Vegt (2018) argued that in every study addressing properties of teams, both horizontal and vertical differences should be taken into account. However, in studies presented in this thesis, main role is given to the vertical differences. That is because in equality scale of the TT10, there are 10 questions concerning how individuals perceive hierarchical structures within their teams and how broadly decision making responsibilities are shared between team members. Thus, this subscale does not reflect horizontal differences between team members at all. However, team members are asked to report their

gender, age, income, and work position and experience as background factors before filling in the actual questionnaire.

Overall, the hierarchical structure, or vertical differences, of a team has been linked to an improved performance when the task requires high levels of procedural interdependence, i.e. team members are not able to complete the task individually but need each other a lot (Halevy, Chou, Galinsky, & Murnighan, 2012). Halevy and colleagues studied National Basketball Association teams. According to their results higher differences in pay and participation were associated with better performance on the basketball field, and with better intragroup coordination and cooperation. Similar results were also found in a setting where groups and their hierarchies were experimentally controlled (Ronay, Greenaway, Anicich, & Galinsky, 2012). Among groups of three people, the ones that had a variety of different hierarchies presented (high, baseline and low), were more productive in tasks that required high levels of procedural interdependence than groups in which there was not such a diversity in power.

On the other hand, in other studies, equality in distribution of conversational turn-taking within a group has been connected to better collective intelligence (Engel et al., 2014; Woolley et al, 2010). As discussed earlier, collective intelligence is defined as 'the general ability of the group to perform a wide variety of tasks' (Woolley et al., 2010, p. 687). Thus those groups, in which discussion was dominated by a few people, performed worse than groups where communication was more evenly distributed. This finding held both for face-to-face groups and for online-based groups (Engel et al., 2014).

However, in these studies (Engel et al., 2014; Woolley et al, 2010) goals of the groups were not so strictly given as in the realm of professional sports. In a more strictly tasked environment of online gaming, collective intelligence of League of Legends online gaming teams was negatively correlated with a perceived equality of leadership within teams (Kim, Engel, Woolley, Lin, McArthur, & Malone, 2015). This is in line with the observations that when the task of the team requires high levels of procedural interdependence, inequality in leadership is an advantage (Halevy et al., 2012; Ronay et al., 2012).

In a study of online-based and experimentally created teams, the perceived equality-based respect of the group members affected their motivation and performance positively in a group task (Renger & Simon, 2011). In this study, the task was to develop ways to improve young people's participation in politics. Everyone had a possibility to propose their ideas individually. Experimenters had, however, manipulated the answers the group members received from their fellows. Some participants got artificially created negative answers while the others got encouraging answers. These answers gave the participant an image that their team member was either interested or not interested in their suggestion. However, no real communication happened between the participants. The results showed that those who got positive and respecting feedback, i.e. those who were respected as equals, performed better in the task. It is noteworthy that in this setting, the tasks of the groups were not interdependent to the same degree as for example in basketball. For this reason, the

results are in line with other findings (Engel et al., 2014; Woolley et al, 2010) showing the advantages of equality in groups where interdependence is not so strong.

Hypotheses for equality are in line with the studies showing the positive effects of equality on groups (Engel et al., 2014; Renger & Simon, 2011; Woolley et al, 2010). In study 1, the tasks of the teams were not clearly defined. However, taking into account that teams were from modern organizations, it would be expected that their work is not solely accomplished together, and everyone has their unique individual contribution. Thus, the interdependence of the groups should not be as high as for example in the sporting world. Based on this reasoning, equality is assumed to correlate with better self-rated interaction and performance of the group.

1.1.10. Shared reflection

Team reflexivity is defined as 'the extent to which group members overtly reflect upon the group's objectives, strategies and processes, and adapt them to current or anticipated endogenous or environmental circumstances' (West, 1996, p. 559). By this West means that reflexive teams discuss their objectives, their plans for achieving their goals, and the exact ways to achieve those goals (e.g. communication, feedback, effectiveness) in terms of their values and the team's commitment to them. According to West, non-reflexive groups tend to focus on the situation that exists at the moment, but reflexive groups also take into account the future. In addition, reflexive groups will have more comprehensive picture about their current situation and they can alter their actions according to multiple cues. Via these actions, these groups can again and again reflect on their performance, and thus improve it due the continuous cycles of reflection.

Reflexivity within a team has been shown to lead to improved performance (Schippers, Homan, & Knippenberg, 2013). This held especially when team's initial performance was poorer. In those situations, reflexivity on previous accomplishment and working habits helped to change the future performance. The same did not hold when teams were already performing well, for then there was not so much to develop. Thus, reflexivity might have been more energy and time consuming than useful for those teams; it might have for example taken time from more useful actions.

In line with this argumentation, some studies have shown that reflexivity is not necessarily related to better performance (Wiedow & Konradt, 2011). However, adaptation to changing situations, which theoretically is an emerging factor with reflection, was related to better performance in the same study. On the other hand, reflexivity has been connected to improved performance when teams had been given feedback and stimulated to co-reflect on it (Gabelica, Van den Bossche, De Maeyer, Segers, & Gijselaers, 2014). Feedback alone was not sufficient to improve the performance of the teams, and reflexivity was needed to derive profit from the feedback. However, in the same study, the advantage gained by reflecting on information faded away with time. Authors argued that this might have been partly due to the study design or the teams misunderstanding

the purpose of the task they did. It still is noteworthy that reflection is not always leading to a better performance in long term.

The advantage of coaching reflexivity within teams has emerged also in other studies (Gurtner, Tschan, Semmer, & Nägele, 2007). Teams that got reflexivity coaching performed better than teams that did not get such. Other findings have shown reflexivity to be connected to the success of coordination (Wiedow & Konradt, 2011). Wiedows and Konradt argued that this was due to the fact that, when teams reflect about their work and strategies they naturally begin to work in a more coordinated fashion.

Moreover, reflexivity counteracted the effects of differentiation in the goal or performance orientation within teams (Nederveen Pieterse, van Knippenberg, & van Ginkel, 2011). When team reflected on their thoughts and processes, they could overcome otherwise challenging differences between the subjective goals of the members. This was especially true in face-to-face teams. Authors argued that in virtual teams where communication, and thus reflection, is not so easy, differences in the goal orientation could cause more harm. On the other hand, in the same study, the advantages of reflexivity diminished when teams were already oriented in the same manner. This could implicate that positive effects of reflexivity may occur when there is a need for a discussion and merging of different points of view. When team members are already thinking in the same way, the need for reflexivity might not be so urgent.

In the TT10 questionnaire there were four questions concerning reflexivity. These questions addressed the amount teams evaluate the quality of their interaction and quality of their work together, and the amount of adaptations done according to the results of these observations. Thus, the questionnaire is quite short in this regard and differs quite a lot from the more comprehensive style of for example Schippers, Den Hartog, and Koopman's (2007) 34-question-questionnaire of reflexivity. This shortage ensues partly from the fact that not very broad perspective on a single subscale could be taken in this kind of a multifaceted questionnaire. However, questions that were taken in the TT10 are in line with theory of reflexivity, and thus are proposed to capture the basic nature behind the concept.

Hypotheses for reflexivity were not easy to form. The subscale of reflexivity in the TT10 is a bit limited, as stated above, and moreover as the TT10 is a self-report, there are no objective measures of the success of the teams. Thus, there is for example no way to know whether the potential low scores on the shared reflection scale of the teams mirror their not so advantageous working habits or their already high performance, and thus not urgent need to reflect on it, as stated by Schippers and colleagues (2013). Also taking into account the results of Nederveen Pieterse and colleagues (2011), it is not so clear that even if teams reflected a lot, the possibly improved performance would be due to reflection. For if team members were thinking in the same manner all the time and still working well, their performance would not be affected by reflexivity because there would not be different kinds of arguments and points of view to consider. Thus, regarding these limitations, no specific hypotheses were formed for the shared reflection scale.

1.1.11. What is achieved through interaction

Even though the achievement scale - what is achieved through interaction - is not an actual subscale of the TT10, it is worth of mentioning here. This scale consists of five questions considering individual's own observations on their team's success in problem solving and completion of tasks. There is also one question considering the learning happening within a team. Thus, performance is addressed as a self-rated behavior rather than effectiveness, profitability or some other more, as could be argued, objective measure.

This definition of performance is in line with the proposition of Campbell (1990). In Campbell's model of performance, performance is separated from effectiveness and productivity. Performance is seen as an action itself, not as a consequence or a result of other actions. Performance leads to team's effectiveness and effectiveness in turn to productivity. For more elaborated discussion of the distinctions and different components of performance, reader is referred to Campbell's original text. Similar approach has been used also in other studies (e.g. Beal et al., 2003). Thus, this study continues this line of research.

All in all, this subscale is the only, but quite a short, way in which the performances of the teams are addressed in the TT10. Thus, all the hypotheses proposing that some construct is connected to better self-reported performance are compared to this subscale.

1.2. The aim of the two studies presented in this thesis

The aim of the two studies presented in this thesis is to examine the reliability, validity, and structure of the TT10 questionnaire. The TT10 is a brand-new questionnaire and thus its structure is still open to modifications. For this reason, literature-based hypotheses of the connections between the subscales of the TT10 questionnaire are formed. The questionnaire is studied in two different contexts: Study 1 with its team-based material is more in line with the original purpose of the TT10 questionnaire, which was to address different factors that may be important for problem-solving in teams. Study 2 addresses these constructs and the suitability of the questionnaire in a more restricted situation in the customer interface and could thus possibly give some interesting insights of the TT10 subscales in these situations.

1.2.1. A short definition of reliability and validity

Reliability, as used in the context of measuring instruments and questionnaires, refers to the instrument's capability to yield consistent results (Lavrakas, 2008; Tsang, Royse, & Terkawi, 2017). Reliability could be addressed in multiple ways. These include, for example, test-retest reliability (how well measures made in different times agree with each other), internal consistency (how well the questionnaire items correlate with each other and whether their results are consistent within the same constructs), and inter-rater reliability (how well different individuals agree in their evaluations of the same construct) (Tsang et al., 2017). Other ways to

address measurement are mentioned by Lavrakas (2008). However, in this thesis reliability was addressed only as the internal consistency and thus these other ways of measurement are out of the scope here.

Validity, on the other hand, refers to questionnaire's ability to measure what it is supposed to measure (Tsang et al., 2017). As is the case with reliability, also validity could be addressed in multiple ways. These include, for example, content validity, criterion-related validity, and construct validity (Lavrakas, 2008). In this thesis validity is addressed as a construct validity, which refers to the extent the results gotten from questionnaire agree with theoretical expectations. All the hypotheses of this thesis reflect this aspect of validity to some extent.

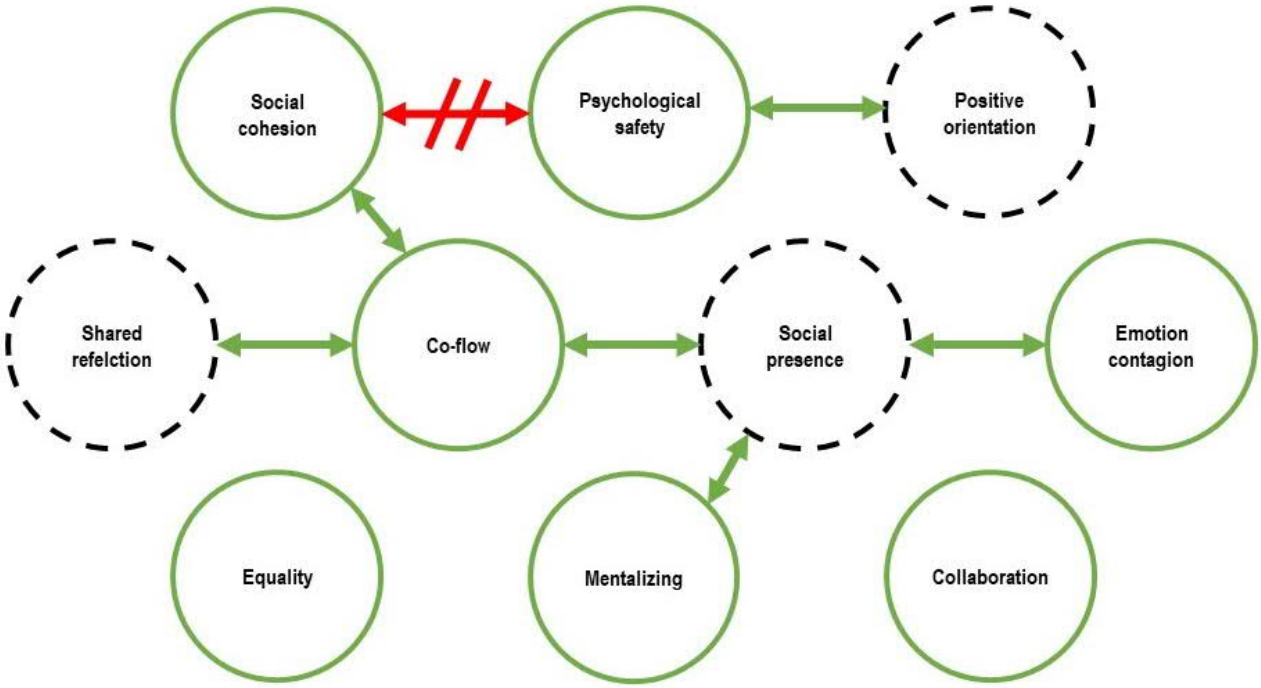

Figure 1.

The hypothesized connections between the subscales of the TT10 in study 1. The subscales presented in green, full-lined circles are hypothesized to correlate with the achievement scale. The subscales presented in the dotted-lined circles are not hypothesized to correlate with the achievement scale. All the other hypothesized correlations between the subscales are presented with green two-way arrows and the only hypothesized non-correlation is presented with a red arrow with traversed lines.

1.3. Hypotheses

1.3.1. The hypotheses of study 1

14 of the 17 hypotheses in study 1 of this thesis are depicted in Figure 1. All these hypotheses concern the relationships between different subscales of the TT10. The three additional hypotheses to those presented in Figure 1 are that mentalizing scores are hypothesized to be higher both among women than men and among teams that include more women than men. In addition, low equality and low social presence/attention scores are hypothesized to correlate with each other.

1.3.2. The hypotheses of study 2

The hypotheses of study 2 are almost the same as those of study 1. There are, however, three expectations, and thus figure 2 was created to present these hypotheses. The first of these expectations is that because in this study it is possible to study correlations between the subscales of the TT10 and the actual achievement of the pairs (measured as an amount of correctly solved puzzles, see 2.3.4. below), all the subscales that are hypothesized to correlate with the achievement scale, and are presented inside the green circles in the figure 2, are also hypothesized to correlate with the number of correctly solved puzzles. In addition to these, also the achievement scale is hypothesized to correlate with that number.

The second expectation is that in this study, equality is not hypothesized to correlate either with the achievement scale or with the actual achievement in the task of this study. This hypothesis is in line with previous studies (Halevy et al., 2012; Kim et al., 2015; Ronay et al., 2012) stating that when the success of the teams needs strong interdependence, inequality and differences in roles of the team members are profitable. In study 2 task completion is based on strong interdependence of the members of the pairs and thus this study differs strongly from study 1 in this sense.

The third expectation is that because shared reflection was not suitable for this study design, and hence was left out of the shortened version of the TT10 used in this study, there are no hypotheses concerning the connections of this subscale to other subscales.

In addition to the hypotheses presented in the figure 2, there are three more hypotheses that are in line with those of study 1. Firstly, women are hypothesized to score higher on the mentalizing scale than men. Secondly, pairs including women are hypothesized to score higher on the mentalizing scale than pairs consisting of only men. This is a bit modified version of the same hypothesis used in study 1. Thirdly, low equality and low social presence/attention scores are hypothesized to correlate with each other.

In addition to the hypotheses mentioned above, there are a few hypotheses in study 2 that are not addressed in study 1. These hypotheses concern The Reading the Mind in the Eyes (RMIE) (Baron-Cohen et al., 2001) and the Interpersonal Reactivity Index (IRI) (Davis, 1980) tests, which are used in study 2 but not in study 1, and their correlations to the TT10 questionnaire's subscales of mentalizing and emotion contagion. Firstly, it is hypothesized that the emotion contagion subscale correlates with the RMIE and the empathic concern subscale of the IRI. Secondly, it is hypothesized that the mentalizing subscale correlates with the RMIE and the perspective taking subscale of the IRI.

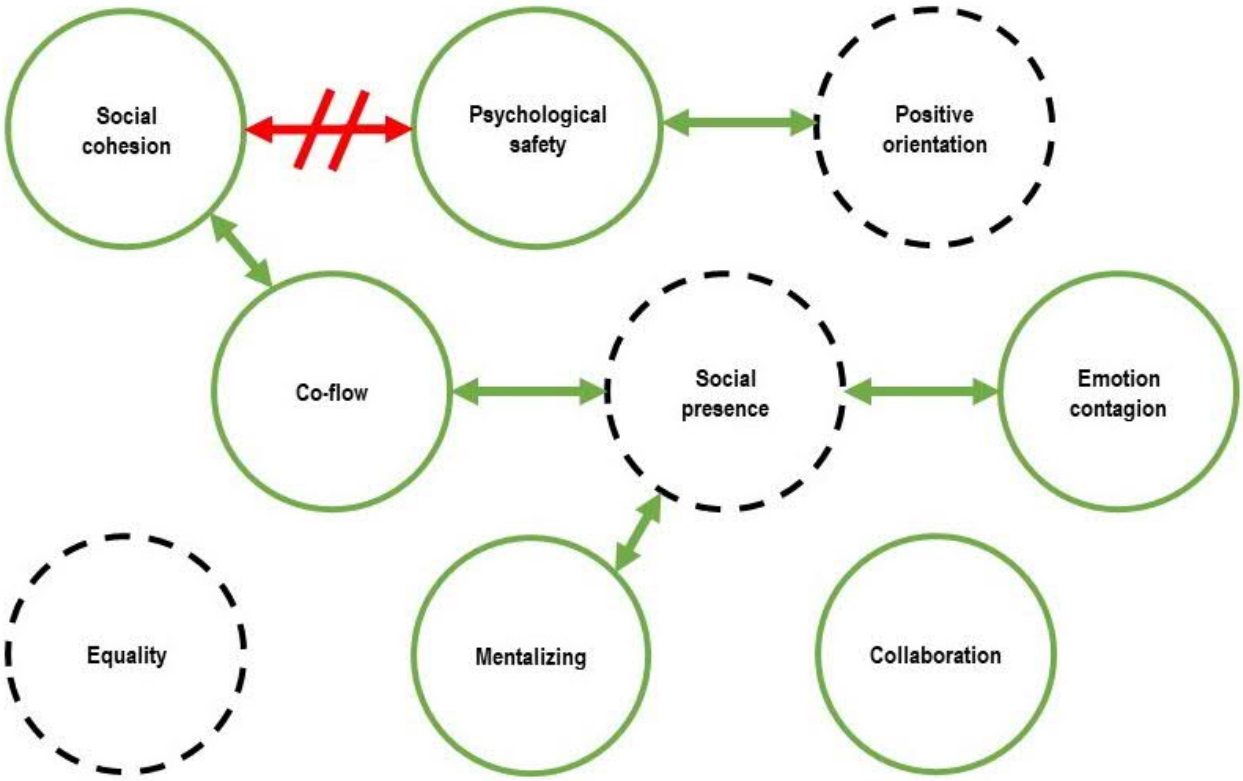

Figure 2.

The hypothesized connections between the subscales in study 2. The subscales presented in green, full-lined circles are hypothesized to correlate with the achievement scale and the number of correctly solved puzzles. The subscales presented in the dotted-lined circles are not hypothesized to correlate with the achievement scale or the number of correctly solved puzzles. All the other hypothesized correlations between the subscales are presented with green two-way arrows and the only hypothesized non-correlation is presented with a red arrow with traversed lines.

2. Methods

2.1. Participants

2.1.1. Participants of study 1

A total of 49 participants from ten different work teams participated in this study. 20 of the participants were women (40.82%), 28 were men (57.14%), and one participant (2.04%) didn't report their gender. Mean age was 35.47 years (SD = 6.04) and the age range was from 24 to 48 years. Regarding education, 3 (6.12%) of the participants had a vocational education, 6 (12.24%) had a high school level of education, 23 (46.94%) had a bachelor's degree or equivalent, and 17 (34.69%) had a master's degree or equivalent.

No statistically significant differences were between the teams regarding age ($F(9, 39) = 1.06$, $p = .41$), or educational background ($F(9, 39) = .73$, $p = .68$) of team members. Moreover, no differences were

found between teams regarding how long the team members had been in working life ($F(9, 39) = .85$, $p = .58$). 79.59% of the participants reported that their team has been in its current form under one year, 18.37% of the participants reported that their team has existed from 1 to 2 years, and one participant (2.04 %) reported that their team has been in current form over 2 years. However, this part of the questionnaire was not strictly restricted and thus these descriptive statistics should be taken cautiously, especially as one participant answered that their team has been in its current form for over 2 years but the other team members answered differently. Moreover, some members from the same team reported the size of their team in different ways. Thus, there are some obscurities regarding the actual compositions of the teams

All the participants were from Finnish organizations, but some of them worked in multicultural teams, and not all the participants had Finnish as their mother tongue. Thus, 14 participants (three of the teams) answered the questionnaire in Finnish and 35 of them answered the TT10 questionnaire in English. Neither nationality nor mother tongue of the participants were asked in order to reserve their anonymity.

When comparing the answers of the teams to the achievement scale of the TT10, it was detected that the mean scores of the teams differed from each other in a statistically significant level in one-way analysis of variance ($F(9, 39) = 2.79$, $p = .013$). Mean scores ranged on this scale from 3.80 to 4.95. This is noteworthy because without statistically significant differences between teams in this scale some of the analysis would not be reasonable to make. At the same time, it still should be remarked that even the mean of the lowest scoring team was 3.80 and thus quite high.

2.1.2. Participants of study 2

A total of 124 participants, forming 62 pairs, participated in this study. All participants were Finnish speaking adults with no diagnosed neurological impairments. Each pair comprised of an employee from the customer service team (46 (74.19%) women, mean age = 35.73 years, standard deviation (SD) = 12.14 years) and a personal customer (16 (25.81%) women, mean age = 57.49 years (3 clients did not report their age), SD = 12.01 years) of the company. Of the pairs, 12 consisted of women only, 34 consisted of both women and men, and 13 consisted men only.

The customers and employees remained anonymous to each other throughout the experiment. Only the role of the other person was shown in the chat window as a "Service Advisor" or a "Customer". The University of Helsinki Ethical Review Board in the Humanities and Social and Behavioral Sciences approved the study protocol. All participants signed a written informed consent.

2.2. Methods

2.2.1. The TT10

In study 1, the TT10 questionnaire was used in its original form. This version of the questionnaire includes 84 statements about team's and individual's functioning. Each statement is answered in a Likert scale that ranges from 1 to 5, 1 meaning not agreeing at all and 5 meaning fully agreeing to the statement. The whole questionnaire is added as an appendix 1 of this thesis. In study 1, all the questions were translated into Finnish and English, and the participants were able to answer in their preferred language.

In study 2, all the participants filled the questionnaire in Finnish. In addition, the TT10 questionnaire was shortened to 27 questions from the original 84 questions in study 2. That is because the study design was based on a test made for pairs collaborating online. As the pairs had not met before the study and they remained anonymous to each other throughout the procedure, the TT10 was not applicable in its original form in study 2.

Hence, several questions were to be left out of this shortened version of the TT10. These included for example questions about mirroring others facial expressions (from the mentalizing scale of the TT10) and feelings of being rejected for being different than others (from psychological safety scale). The whole subscale of shared reflection was left out of this shortened version, because participants did not have an opportunity to reflect on their success with their pair in this study.

Secondly, when addressing pairs, it was not possible to use the exactly same questionnaire as in study 1. The main points of the TT10 questions were the same for this study also, but vocabulary was altered to address the themes in question in a proper manner. In these altered questions the word ”team” was replaced by ”pair”. Moreover, some alterations were made to questions so that they were fit for a more limited situation in this study. In study 1 questions addressed the working habits of the teams in a general fashion. In this study, pairs were formed for just one task and for a short time frame. Thus, some questions were specified to fit to this more limited situation. All in all, the questions included in this shortened version are bolded in the appendix 1.

2.2.2. Other questionnaires used in study 2

In addition to the TT10 questionnaire, there were also other questionnaires used in study 2. Two of these are already shortly mentioned in the introduction: The Reading the Mind in the Eyes (RMIE) test (Baron-Cohen et al., 2001) and the Interpersonal Reactivity Index (IRI) (Davis, 1980). Both these tests assess the social sensitivity and theory of mind skills of the participants. They are self-reported and thus address the participant's own thoughts about their skills, unlike in the TT10 in which participants review both their own and their team's or pair's skills in these domains.

The RMIE was originally created by Baron-Cohen and colleagues in 1997. The version used here is a refined version (Baron-Cohen et al., 2001). In the RMIE participants are presented with pictures of human eyes and they have to choose the right word describing the emotion eyes are showing in pictures. This is thought to reflect the participant's theory of mind skills and ability to take the other person's perspective.

The IRI, on the other hand, was developed by Davis (1980). It consists of 28 questions that are from four separated subscales. There are seven questions in each subscale. These subscales are fantasy, perspective-taking, empathic concern and personal distress. Of these the scales of perspective-taking and empathic concern are especially relevant for this study. Perspective-taking scale addresses the ability to see other people's perspectives in real life situations, and empathic concern scale addresses the ability to feel warmth, compassion and concern when observing other individuals and their actions.

Thus, as mentioned in the introduction, the RMIE and IRI, are hypothesized to be in close connection to the mentalizing and shared reflection scales of the TT10. Thus, they are a useful addition to this study. What is especially of interest here is that the RMIE and IRI are self-reported measures of theory of mind and empathy skills. Hence, they allow for comparison of scores in self and other reviewed skills in these constructs. In the mentalizing and emotion contagion subscales of the TT10, there are questions both on participant's own skills on these domains and their evaluation of the skills of their pairs. Thus, comparing both the evaluations of participants on their own skills and the evaluations of their counterparts on the skills of their pairs in the TT10 scales of mentalizing and emotion contagion to their RMIE and IRI -scores could further validate mentalizing and emotion contagion scales.

Other questionnaires used in study 2 included abbreviated nine-item version of the 60-item Raven's Standard Progressive Matrices Test (Bilker, Hansen, Brensinger, Richard, Gur, & Gur, 2012). The mood of the participants was also asked both before and after the test with the 24-item Profile of Mood States questionnaire (POMS) (Terry, Lane, & Fogarty, 2003). Finally, participants were asked to answer three additional questions about their feelings of social closeness to their pairs. These items were answered in ten-point scale. These three questions were adapted from the paper of Tarr, Slater & Cohen (2018). However, the results of these three questionnaires were not addressed in this study and are a subject of another yet unpublished study. Thus, they are not further discussed here.

2.3. Procedure

2.3.1. The procedure of study 1

In study 1, the teams answering the TT10 online questionnaire came from two Finnish technology companies, which were partners in the University of Helsinki's HUMEX (Quantifying Human Experience for Increased Intelligence Within Work Teams and in the Customer Interface) research project funded by Business Finland.

Possibility to participate was advertised through the companies' in-house mailing lists. All answers were given anonymously in internet and participation was voluntary. At least four answers per team were required in order to analyze the data. In addition to answering the TT10 questionnaire, data was gathered on work experience and basic demographic factors. There was also an open-ended question that allowed participants to give feedback about the questionnaire.

2.3.2. The procedure of study 2

The participants of this study were recruited via a Finnish insurance company's in-house and client mailing lists. Participants were randomly assigned to pairs of one employee and one customer. The only interaction of the pairs throughout the task was via chat environment that was created for this study only. Otherwise pairs were not in contact with each other. Customers participated in the study using their personal computers. They were instructed by the researchers over phone and email. Insurance company's employees participated from their company's office where researchers were also present.

Customers answered the questionnaires (RMIE, reasoning skills test & IRI) before the experiment. They were instructed to fill questionnaires on the day before the experiment. In addition, they were asked to fill some basic information about themselves such as their age, media (social media, television) usage, years they had been in working life and in their current job, and time they spend with their family per day. On the experiment day, customers were sent mood test and written instructions to the experiment via an email approximately 30 minutes before the start of the experiment.

Employees, on the other hand, were taken to their company's meeting room and three electrocardiograph (ECG) sensors were attached to them for the purpose of heart rate measures. Heart rates of the employees were measured and transmitted to the customers during the task. However, this part of the task is not relevant for this study and is discussed elsewhere. After attaching sensors, written instructions for completing the task were given to the employees. A few minutes before the task, both employees and customers were given spoken instructions for the task, employees in person and customers via phone. After the instructions the pairs were asked to fill in the mood questionnaire. After the questionnaire was filled, they were asked to proceed to the chat environment in order to begin the task. Employees were informed to start chatting in their own manner whenever they were ready. There were no limitations to the amount of chatting during the task.

When both the customer and the employee were ready, they started the task. In the task, there were 15 minutes time during which the aim of the pairs was to build correctly as many matrice puzzles as possible. Customers saw just one matrice with incomplete symbol and another symbol piece beside the matrice. Employees on the other hand saw two similar sized matrices on their screen: one as customer was seeing it and another with a symbol the customer was expected to build according to the employee's instructions. After

the customer had built the correct symbol on the matrice, the pair proceeded to the next puzzle. After 15 minutes time was up, employees were informed to stop the task and end the discussion with customer in the manner of their choosing. The sum of correctly completed puzzles was used as an indicator of the pair's success in the task. The first puzzle was not counted in this sum for it was taken as an introductory item.

Right after the experiment both members of the pair were again asked to fill in the mood questionnaire and the abbreviated the TT10 questionnaire. In addition, they filled questionnaires for evaluating their interaction and feeling of social closeness. There was also a possibility for open feedback about the likeability of this study and its design. Due to practical reasons, the employees filled in the background questionnaires and other tests (RMIE, IRI, Raven's matrices) after the experiment.

2.4. Statistical analysis

All statistical analysis in both studies were performed with IBM SPSS Statistics (version 25, 64-bit edition) program. The same statistical methods were used in both studies expect Cronbach's alphas that were used to estimate reliability of the subscales of the TT10 in study 1 only. In study 2, Cronbach's alphas were not calculated due to the shortened version of the TT10 applied. There were two reasons for this. Firstly, and more importantly, some scales were shortened to just a couple of questions. Thus, a lot of variability within these scales was lost. Secondly, the TT10 was applied here in a way it was not necessarily designed to be applied. Hence, Cronbach's alphas would not have given any insights into the reliability of the original scales.

3. Results

3.1. The results of study 1

First Cronbach's alphas were calculated for every subscale of the TT10 to gain insights into their reliabilities. Results are collected in Table 1.

Correlation analysis was conducted to analyze hypothesized connections between different subscales of TT10 and their connections to the achievement scale. All the results of this analysis are collected to Table 2. All the correlations were statistically significant ($p < .001$ expect in two cases where $p < .01$). All the TT10 subscales correlated with the achievement-scale with a moderate ($r = .51$ for resonance scale) to large ($r = .75$ for co-flow scale) effect.

When studying mentalizing scores of the teams based on the number of women in the teams, independent samples t-test was conducted. It was found that in those teams in which there were more women than men ($n = 21$, consisting of 4 teams) mean scores of the mentalizing scale were not statistically significantly higher than those of other teams ($t(47) = 1.69$, $p = .098$). Independent samples t-test was also

conducted to address whether women score higher on the mentalizing scale than men. No statistically significant gender differences were found ($t(46) = 1.91$, $p = .063$).

The comparison of the mentalizing scores of the teams was also made based on the percentage of women in a team. There was a statistically significant difference in one-way analysis of variance between teams according to this measure ($F(7, 41) = 6.41$, $p < .001$). When further addressing differences between groups based on this statistic, it was observed that in independent samples t-test mentalizing scores were statistically significantly lower in teams in which there were 20% or fewer women (consisting of 4 teams) than in teams in which there were over 20% women (6 teams) ($t(47) = -4.82$, $p < .001$).

When studying connections between the social presence and equality scores of the teams, three lowest and three highest scoring teams on equality scale were filtered. Then independent sample t-test was conducted to address differences on mean social presence scores between these two groups. Results showed that those teams that scored lower on equality scale, scored also statistically significantly lower on social presence scale ($t(25) = -4.22$, $p < .001$) than the higher scoring teams.

Table 1.
Cronbach's alphas of every subscale of the TT10.

TT10 Subscale	Cronbach's alpha
Collaboration	.74
Co-flow	.90
Social cohesion	.88
Emotion contagion	.78
Mentalizing	.85
Social presence/attention	.65
Psychological safety	.85
Positive orientation	.83
Equality	.81
Shared reflection	.86
Achievement	.76

Table 2.

Pearson's correlations coefficients between all the TT10 subscales and the Achievement scale (Achieved). All the hypothesized correlations are bolded and the hypothesized non-correlation underlined.

The subscales of the TT10	Achieved	Collaboration	Co-flow	The subscales of the TT10							
				Sos. Cohesion	Emotion contagion	Mentalizing	Soc. Presence	Psyc. Safety	Pos. Orientation	Equality	Shar. Reflection
Achieved	1.00										
Collaboration	.70	1.00									
Co-flow	.75	.69	1.00								
Sos. Cohesion	.73	.68	.87	1.00							
Emotion contagion	.51	.57	.63	.61	1.00						
Mentalizing	.56	.56	.74	.78	.56	1.00					
Soc. Presence	.60	.57	.77	.73	.51	.81	1.00				
Psyc. Safety	.74	.67	.62	.55	.42	.52	.54	1.00			
Pos. Orientation	.74	.62	.78	.80	.59	.74	.80	.74	1.00		
Equality	.65	.63	.57	.68	.47	.58	.50	.57	.63	1.00	
Shar. Reflection	.68	.44	.64	.66	.37	.52	.47	.51	.59	.60	1.00

For all the p-values of the correlations held p<.01

3.2. The results of study 2

To test the connections between the subscales of the TT10 and the self-rated success of interaction of the pairs, correlation analysis was conducted similarly as in study 1. Results are presented in Table 3. This table also shows the correlations between all the TT10 subscales, the achievement scale, and the number of correctly solved puzzles. Exactly the same correlations were also calculated separately for clients and employees. These results are collected in Table 4.

In a similar fashion, correlations between the TT10 subscales of mentalizing and resonance/emotion contagion, the RMIE questionnaire, and the IRI subscales of Empathic concern and Perspective taking were calculated. Results are collected in Table 5. To gain a more fine-grained picture of these connections between different questionnaires, the same Pearson's correlation coefficients were also calculated for the employees and clients separately. In these calculations both mentalizing and emotion contagion subscales were divided to two parts: questions concerning the participants themselves and questions concerning their pairs. The results of these analysis are collected in Table 6 (employees) and Table 7 (clients).

To test whether women scored higher on the mentalizing scale of the TT10 than men, independent samples t-test was conducted. No statistically significant differences were found between genders ($t(121) = 1.81$, $p = .073$). No statistically significant differences were found between differently composed pairs (men only, mixed gender and women only) in one-way analysis of variance ($F(2, 120) = .97$, $p = .38$). Pairs that consisted of women only did not differ statistically significantly from teams that consisted of men only in independent samples t-test ($t(48) = .45$, $p = .66$).

When addressing the differences in social attention scores between individuals who ranked the equality of their pairs in different ways, it was noted that those whose scores on equality were lower than 3 in the TT10's Likert five-point scale ($n = 13$) reviewed the social attention of their pairs statistically significantly lower in independent samples t-test ($t(109) = -2.51$, $p = .013$) than those whose scores were higher than three in the equality scale ($n = 98$).

Table 3.

Pearson's correlation coefficients between the subscales of the TT10, the achievement scale, and the number of correctly solved puzzles. Hypothesized correlations are bolded and the hypothesized non-correlation is underlined.

				The subscales of the TT10							
The subscales of the TT10	Correctly solved puzzles	Achieved	Collaboration	Co-flow	Social cohesion	Emotion contagion	Mentalizing	Soc. Presence	Psyc. Safety	Positive orientation	Equality
Correctly solved puzzles	1.00										
Achieved	.15	1.00									
Collaboration	.11	.56*	1.00								
Co-flow	.10	.43*	.40*	1.00							
Soc. Cohesion	.15	.76*	.59*	.58*	1.00						
Emotion contagion	.05	.38*	.42*	.56*	.60*	1.00					
Mentalizing	.09	.46*	.46*	.36*	.58*	.56*	1.00				
Soc. Presence	.08	.45*	.43*	.48*	.48*	.55*	.41*	1.00			
Psyc. Safety	.07	.44*	.52*	.28*	.41*	.37*	.48*	.29*	1.00		
Pos. Orientation	.05	.61*	.64*	.65*	.75*	.68*	.62*	.60*	.55*	1.00	
Equality	.02	.32*	.32*	.37*	.38*	.44*	.28*	.35*	.27*	.46*	1.00

*. Correlation is significant at the 0.01 level (2-tailed).

Table 5.

Correlations between mentalizing and emotion contagion subscales of the TT10, results of the IRI subscales of Perspective taking and Empathic concern, and the RMIE. The hypothesized correlations are bolded.

	Questionnaire and subscale				
Questionnaire and subscale	Emotion contagion	Mentalizing	RMIE	Perspective taking (IRI)	Empathic concern (IRI)
Emotion contagion	1.00				
Mentalizing	.56*	1.00			
RMIE	-.05	.01	1.00		
Perspective taking (IRI)	.13	.15	.03	1.00	
Empathic concern (IRI)	.12	.15	.05	.50*	1.00

*. Correlation is significant at the 0.01 level (2-tailed).

Table 4.

Pearson's correlation coefficients between the subscales of the TT10, the achievement scale, and the number of correctly solved puzzles for clients and employees separately. The hypothesized correlations are bolded and the hypothesized non-correlation is underlined.

			The subscales of the TT10									
Role	The subscales of the TT10	Correctly solved puzzles	Achieved	Collaboration	Co-flow	Social cohesion	Emotion contagion	Mentalizing	Soc. Presence	Psyc. Safety	Positive orientation	Equality
Client	Correctly solved puzzles	1.00										
	Achieved	.28*	1.00									
	Collaboration	.23	.66**	1.00								
	Co-flow	.22	.59**	.33**	1.00							
	Soc. Cohesion	.30*	.82**	.60**	.70**	1.00						
	Emotion contagion	.24	.52**	.41**	.58**	.72**	1.00					
	Mentalizing	.17	.53**	.44**	.42**	.65**	.71**	1.00				
	Soc. Presence	.20	.55**	.51**	.54**	.51**	.54**	.43**	1.00			
	Psyc. Safety	.14	.39**	.44**	.29*	.36**	.37**	.42**	.24	1.00		
	Pos. Orientation	.18	.73**	.64**	.62**	.80**	.67**	.65**	.54**	.59**	1.00	
	Equality	.25	.40**	.32*	.58**	.46**	.41**	.35**	.53**	.25	.55**	1.00
Employee	Correctly solved puzzles	1.00										
	Achieved	-.01	1.00									
	Collaboration	.00	.50**	1.00								
	Co-flow	.02	.43**	.46**	1.00							
	Soc. Cohesion	-.04	.69**	.61**	.59**	1.00						
	Emotion contagion	-.13	.33*	.42**	.50**	.55**	1.00					
	Mentalizing	.00	.33**	.53**	.51**	.50**	.55**	1.00				
	Soc. Presence	-.01	.44**	.37**	.42**	.49**	.56**	.47**	1.00			
	Psyc. Safety	-.03	.44**	.67**	.48**	.45**	.51**	.48**	.41**	1.00		
	Pos. Orientation	-.07	.58**	.63**	.67**	.75**	.68**	.68**	.63**	.62**	1.00	
	Equality	-.12	.36**	.32*	.25	.38**	.47**	.31*	.27*	.39**	.41**	1.00

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Table 6.

Pearson's correlation coefficients between the TT10, the IRI, and the RMIE scores of the employees.

		Subscale and questionnaire											
Role	Subscale and questionnaire	Mentalizing (TT10)	Emotion contagion (TT10)	RMIE	Evaluation of pairs' mentalizing skills (TT10)	Evaluation of own mentalizing skills (TT10)	Evaluation of own emotion contagion skills (TT10)	Evaluation of pairs' emotion contagion skills (TT10)	Employees' perspective taking (IRI)	Employees' empathic concern (IRI)	Client's perspective taking (IRI)	Client's empathic concern (IRI)	Client's RMIE
Employee	Mentalizing (TT10)	1.00											
	Resonance (TT10)	.55**	1.00										
	RMIE	-.10	-.08	1.00									
	Evaluation of pairs' mentalizing skills (TT10)	.77**	.66**	-.04	1.00								
	Evaluation of own mentalizing skills (TT10)	.86**	.29*	-.11	.34**	1.00							
	Evaluation of own emotion contagion skills (TT10)	.59**	.90**	-.11	.66**	.34**	1.00						
	Evaluation of pairs' emotion contagion skills (TT10)	.39**	.88**	-.03	.51**	.17	.59**	1.00					
	Employees' perspective taking (IRI)	.14	.13	-.03	-.05	.25	.09	.14	1.00				
	Employees' empathic concern (IRI)	.26*	.30*	-.08	.18	.23	.24	.30*	.54**	1.00			
	Client's perspective taking (IRI)	.04	-.07	-.07	.00	.06	-.04	-.10	-.17	-.16	1.00		
	Client's empathic concern (IRI)	-.13	.01	-.06	-.11	-.10	.01	.02	-.23	-.22	.40**	1.00	
	Client's RMIE	-.01	.16	.07	.00	-.01	.02	.29*	-.12	-.01	-.04	.05	1.00

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Table 7.

Pearson's correlation coefficients between the TT10, the IRI, and the RMIE scores of the clients.

		Subscale and questionnaire											
Role	Subscale and questionnaire	Mentalizing (TT10)	Emotion contagion (TT10)	RMIE	Evaluation of pairs' mentalizing skills (TT10)	Evaluation of own mentalizing skills (TT10)	Evaluation of own emotion contagion skills (TT10)	Evaluation of pairs' emotion contagion skills (TT10)	Employees' perspective taking (IRI)	Employees' empathic concern (IRI)	asiakaspalvelijan RMIE	Client's perspective taking (IRI)	Client's empathic concern (IRI)
Client	Mentalizing (TT10)	1.00											
	Resonance (TT10)	.71**	1.00										
	RMIE	-.05	.10	1.00									
	Evaluation of pairs' mentalizing skills (TT10)	.80**	.77**	.04	1.00								
	Evaluation of own mentalizing skills (TT10)	.87**	.46**	-.12	.40**	1.00							
	Evaluation of own emotion contagion skills (TT10)	.65**	.93**	.17	.67**	.44**	1.00						
	Evaluation of pairs' emotion contagion skills (TT10)	.66**	.90**	.00	.74**	.39**	.68**	1.00					
	Employees' perspective taking (IRI)	.11	.06	-.12	.04	.14	.09	.00	1.00				
	Employees' empathic concern (IRI)	.21	.01	-.01	.08	.27*	.02	-.01	.54**	1.00			
	Employee's RMIE	.29*	.22	.07	.29*	.21	.29*	.08	-.03	-.08	1.00		
	Client's perspective taking (IRI)	.07	.21	-.04	.22	-.05	.21	.23	-.17	-.16	-.07	1.00	
	Client's empathic concern (IRI)	-.07	.00	.05	-.07	-.04	.03	-.01	-.23	-.22	-.06	.40**	1.00

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

4. Discussion

In this section, a short overall discussion is given first. The results of the two studies presented in this thesis are summarized and their implications for the applicability of the TT10 questionnaire is discussed. After that each TT10 subscale is addressed separately to gain more fine-grained view of the results. This is important as each subscale of the TT10 is based on a different theoretical background. Moreover, even though the questions of the TT10 questionnaire are based on the previous research and some of them are from previous questionnaires, there are also new questions designed for this questionnaire in almost all the subscales. Thus, the TT10 and its validity should be interpreted by taking into account all these different scales separately. Limitations to each subscale are also discussed in the context of overall discussion of that subscale. The future suggestions for research of each subscale are also given. The overall limitations of this thesis and its two studies are discussed after the discussion on the specific subscales. A short conclusion is given in the end.

4.1. General discussion

In the two studies presented in this thesis, the usefulness, reliability, and validity of the TT10 questionnaire were addressed by studying multiple literature based hypothesis about the subscales of the TT10. The TT10 is a new questionnaire that has not been much used to this day. These results give some implications for the future development of the TT10.

Firstly, all the subscales seemed to be reliable in this study. Cronbach's alphas of every subscale were high. Almost all the alphas were higher than .70 (see Table 1) that is considered to be a sign of a good internal consistency of a measurement (Lavrakas, 2008). However, Lavrakas argues that also the thresholds of .75 and .80 are used in some literature. The only subscale of which one of the alphas did not exceed the threshold value of .70 was social presence/attention. When taking into account the fact that the social presence/attention scale was not firmly based on previous research, this finding was not totally surprising. It also indicates that further evidence is needed to address the reliability of this scale. All in all, taking into account the quite low number of participants in study 1, these results were promising for the reliability of the TT10.

In addition, all the hypotheses addressing the connections of the TT10 subscales to the achievement scale and to each other gained from moderate to strong support from both studies (see tables 2 and 3) even though the hypothesized correlations were smaller in study 2 than they were in study 1. The strongest connections were between the achievement scale and collaboration, co-flow, social cohesion, and psychological safety. The Pearson's correlation coefficients between all these subscales and the achievement

scale exceeded r = .70, which is in literature held as a threshold value for strong positive correlation (Hinkle, Wiersma, & Jurs, 2003).

However, when addressing correlations between the subscales of the TT10 and pairs' actual success in their task in study 2, it was remarked that the hypothesized correlations were very weak (table 3). From these, especially low and not statistically significant correlation between the achievement scale and correctly solved puzzles is noteworthy. It was hypothesized that these two scales would correlate strongly because they should reflect the same construct even if from different perspectives. The achievement scale addresses pairs' evaluation of their success in terms of their own perceptions on their work. Correctly solved puzzles, on the other hand, is an objective measurement of their actual performance. Hypothetically, if the pairs rated their interaction as successful, they would also be expected to perform better. This was not the case in study 2. It should be still remarked that in the shortened version of the TT10 used in this study, there was just one question in the achievement scale ("Collaboration with my pair went well"). Thus, this scale was much more limited than in study 1 and these results should be interpreted cautiously.

Another noteworthy result that should be mentioned was that almost all the subscales of the TT10 correlated with each other to some extent. In the light of these studies it is not clear what caused this result. It might have been due to similarity in the TT10 questions or some other factor that is affecting the results, such as the tendency of teams and pairs to answer in a consistent manner. This held especially in study 1, in which teams did not differ from each other a lot regarding their background information. Hence, the similarity in answers might have been due the similarity in teams. However, whether this is true or not remains unclear as long as there are no objective measures of the working habits of the teams and comparisons between the different measurements are not made.

Some more obscurity to the results comes from the finding that, when comparing the connections between clients and employees separately, correlations were somewhat stronger for clients than for employees in study 2 (Table 4). Only difference was the scale of psychological safety, for which the correlation was a bit stronger for employees than clients. Statistical significance of these differences in correlations were yet not addressed here. However, taking into account the equal sample sizes for both groups and statistical significance of every correlation ($p < .01$ for all the hypothesized correlations, expect the correlation between emotion contagion and the achievement scale for employees was significant at the level of $p < .05$), these differences could be interpreted as at least worth mentioning. They could for example reflect the differences in the roles of the client and the employee in solving the task. Employees had a commanding role in solving the puzzles and hence members of a pair were not as equal as they would be in a more balanced study design or when solving a task that would require same kind of contributions from both members of the pair.

As a whole, and despite some vagueness in the results, validity and reliability of the TT10 gained support from these studies. When considering all the scales at the same time, it seems to be that the most

promising ones are the scales of psychological safety, social cohesion, collaboration, and co-flow. However, these overall results cannot give a comprehensive picture of the TT10. To gain more insights into separate subscales and their specific properties, each of the subscales is further discussed below.

4.2. Collaboration

Collaboration was defined as a team's good ability to work cooperatively (Johnson & Johnson, 2003). Collaborative teams were hypothesized to perform better than non-collaborative teams. This was based on the notions of Johnson and Johnson, that collaborative teams outperform competitive teams, especially in tasks that have specific goals or procedures. Thus, it was also presumed that correlations between collaboration and the achievement scale might be higher in study 2, in which the task was clearly defined, unlike in study 1 in which the tasks of the teams were not limited and thus the nature of their work remained somewhat unidentified.

These hypotheses were not fully supported. In study 1, collaboration was strongly positively correlated to the achievement scale ($r = .70$, $p < .01$; Table 2) but although in study 2 correlation between the same subscales of the TT10 was statistically significant ($r = .56$, $p < .01$) it was still a bit lower than in study 1. It should still be noted that collaboration was not correlated to actual success in the puzzle-solving task of study 2 ($r = .11$; Table 3). The task used in study 2 was however new and not familiar to the participants and this could have had an effect on the results. For example, it was possible that regardless how well the members of the pair collaborated, their different skill levels on this kind of a task would have made it very difficult to solve the puzzles correctly. A short time-frame could also have made it difficult to form a good collaboration within pairs.

It should be also noted that the evaluations of collaboration by the clients and the insurance company's employees were differently correlated to both the achievement scale and the number of correctly solved puzzles (Table 4). Correlations were higher for clients in both cases, even though the correlation with correctly solved puzzles was not statistically significant for either group. This difference could have stemmed from the different roles of the members in the pairs in solving puzzles. Employees had to guide clients to the correct solution and thus they were in charge of communication. Thus, clients' perspective of cooperation and their feeling of being needed to correctly solve puzzles might have reflected the basic elements of collaboration more concretely than those of the employees'.

This is not to say that the role of the clients in the collaboration was not important. Correlations between the achievement scale and collaboration were strongly positive for both clients and employees. Both members naturally affect the communication and bring their own features to it. In this study design the employee's role was however more guiding than the client's and in this sense differences in evaluating collaboration and its effects on successful completion of the task were not totally surprising. In future studies

with pairs it could be good to study pairs with a more balanced role in solving tasks to see whether the differences between the members of the pairs are diminished.

All in all, two studies presented in this paper support the validity and reliability of collaboration scale of the TT10. However, there was only one hypothesis formed based on the background literature and more elaborate research on this scale could be worthwhile. Moreover, in the article of Johnson and Johnson (2003), five possible determinants of cooperation were introduced but these were not included in the TT10. As the TT10 is a short and multifaceted questionnaire, this is justifiable, but for better theoretical understanding of collaboration and cooperation these aspects could be taken into account if the collaboration subscale were to be modified in the future.

4.3. Co-Flow

Co-flow is a term closely related to individual flow. Both terms are connected to better achievement in various tasks, to task immersiveness, and to a total concentration in the task at hand. Both individual and group flow are seen to be based on similar factors. Individual flow is based on clear set of goals, a balance between perceived skills and perceived challenges, and clear and immediate feedback (Csikszentmihalyi et al., 2014). Group flow on the other hand is reasoned to be based on three similar constructs: vision, ownership and contribution, and communication (Duncan & West, 2018).

In the studies presented in this thesis, it was firstly hypothesized that co-flow would be connected to better achievement of the teams. This hypothesis gained from moderate to strong support from both studies. In study 1 co-flow correlated with the achievement scale of the TT10 strongly positively ($r = .75$). The corresponding correlation was moderately strong in study 2 ($r = .43$). However, co-flow was not statistically significantly correlated to the amount of correctly solved puzzles in study 2.

It is also noteworthy that, as was the case with the collaboration scale, also co-flow and achievement scale and the amount of correctly solved puzzles correlated differently for clients and employees in study 2. Reasons for this are not clear but the role difference could explain some of it. However, this does not fully explain the difference, for co-flow is thought to be a feature that holds for the whole group. Thus, differences between different participants from the same team are not to be expected. More studies with the same kind of study design are needed to address this difference better.

The validity of the co-flow scale was also addressed with two other hypotheses in this study. Firstly, co-flow was hypothesized to correlate with the subscales of social attention and social cohesion, and secondly, co-flow was hypothesized, based on the literature, to correlate with the shared reflection scale. All these hypotheses gained strong support from both studies, although the hypothesis of co-flow's correlation with shared reflection was only addressed in study 1 as the scale of shared reflection was not appropriate to use in study 2. Again, correlation between subscales were stronger for clients than for employees in study 2.

On the whole co-flow subscale of the TT10 seemed to be valid and reliable, and hypotheses based on its theoretical background gained support. However, a couple of limitations should be mentioned. Shared reflection scale was used only in study 1, and taking into account the limitations in the number of participants in that study, more studies are needed to gain more insight on whether co-flow and shared reflection truly are connected with each other.

Another limitation is that the reasons behind co-flow's strong correlations with the subscales of social presence and social cohesion are not totally clear. This held especially in study 1 in which the correlation between social cohesion and co-flow was the highest correlation (r = .87) among all the correlations. It should be pointed out that the co-flow and the social cohesion subscales have same kinds of questions regarding the enjoyment of the company of the team members and whether the participant feels frustration when working with their teams. Thus, some correlation between these scales was presumable even without a strong theoretical background.

This is a particularly valuable observation, because in open-ended questions in study 1 some participants reported that the TT10 was too long and some questions were repeated too often. Thus, there is a possibility here to modify these scales in the future versions of the TT10 to better separate the scales from each other. One possibility would be to modify the co-flow scale by using three background categories of flow proposed by Duncan and West (2018).

4.4. Social Cohesion

Social cohesion is defined as a strong social and motivational bond existing between team members (Beal et al., 2003). Social cohesion has been connected to a better performance of the group. This holds especially for groups that aim to efficiency and not solely on effective solving of tasks. Thus, especially the design of study 1 of this paper was suitable to address social cohesion as in it participants were not required to perform any specific task and hence their evaluations might reflect more the processes within their teams than any specific accomplishments. Moreover, the achievement scale asked participants about their thoughts about the working in the groups and the success of interaction of the group. In doing so, it addressed the efficiency of the groups rather than measured the objective success of the groups.

Although there was only one hypothesis concerning social cohesion, this hypothesis gained strong support from both studies. Social cohesion was truly connected to a better self-rated performance of the group. The correlation between social cohesion and the achievement scale was especially high in study 2 (r = .82). Social cohesion was not however statistically significantly correlated to the amount of correctly solved puzzles. This was not so alarming taking into account the remarks of Beal and colleagues above mentioned (2003) that social cohesion is more important for the efficiency of the teams than their actual success.

Thus, social cohesion scale of the TT10 seemed to be theoretically solid in the light of limited findings of the two studies presented here. It should still be taken into account that social cohesion was addressed with just two questions in study 2. In the original form of the TT10 there are seven questions. This could have simplified the results in study 2, and in that sense it is not easy to compare the results of studies 1 and 2. Moreover, as remarked above, the questions of social cohesion and co-flow overlap with each other to some degree. Thus, more fine-grained separation of these two scales is needed to truly address whether these scales measure theoretical constructs they are supposed to measure.

4.5. Resonance/emotion contagion

Resonance/emotion contagion is the first of the two empathy related scales in the TT10. Emotion contagion refers to people's actions of mimicking the actions of the others or otherwise feeling what the others are feeling (de Waal & Preston, 2017). It is considered to represent affective empathy. All in all, the amount of literature assessing emotion contagion's effect on teamwork was limited, thus it was necessary to form the hypotheses based on the overall empathy literature.

First, it was hypothesized that emotion contagion would be correlated with a better achievement of the teams. This hypothesis gained from moderate to strong support from both studies. Connection was stronger in study 1 than in study 2. In addition, it should be again remarked that correlation between the achievement scale and emotion contagion was higher for clients than for employees. This is possibly again reflecting the role difference in solving the task.

Second, it was hypothesized that emotion contagion is connected to social presence. This is because if people do not pay attention to the others, they might not be able to recognize their emotions. This hypothesis gained strong support from both studies.

More validation for emotion contagion was gathered from study 2, in which two other empathy related questionnaires (The Reading the Mind in the Eyes (RMIE) test (Baron-Cohen et al., 2001) and the Interpersonal Reactivity Index (IRI) (Davis, 1980)) were used. However, these results did not support the validity of emotion contagion scale, because contrary to what was hypothesized, emotion contagion was not correlated neither with the RMIE questionnaire nor the empathic concern scale of the IRI. Thus, the emotion contagion scale of the TT10 did not seem to be related to the results of previous, more validated questionnaires.

However, it is to be pointed out that there is a difference in the style of these different questionnaires that could have affected the results of study 2. Emotion contagion subscale of the TT10 questionnaire is related to individual situation while the IRI and the RMIE address more universal and stable features of the participants. In line with this remark, there were statistically significant correlations both between the TT10 subscales of mentalizing and emotion contagion and between the IRI subscales of empathic concern and

perspective taking. The IRI and the TT10 were still not correlated with each other. The RMIE questionnaire was also not correlated either to the subscales of the TT10 or the subscales of the IRI. Thus, these questionnaires might measure different concepts however related they seem to be at the first glance.

In line with this remark, empathy as a construct has been criticized in some literature, or at least it is said that empathy should be better conceptualized. For example, Singer and Lamm (2009) argued that if empathy is strictly defined as sharing the feelings of the others, it is only "one part of a large spectrum of a person's possible vicarious responses toward others" (p. 82). They also stated, that in the scientific literature, there are almost as many definitions of empathy as there are researchers studying it.

Moreover, Singer and Lamm (2009) made a couple of points that are essential for this study. Firstly, they stated that emotion contagion and mimicking others' feelings are related but distinct from empathy. This is a bit different view from the suggestions that emotion contagion is a mechanism that is a part of affective mechanisms of empathy, and thus leads to it (de Waal & Preston, 2017). Hence, taking into account these contrasting views of the role of emotion contagion to the experience of empathy the results of study 2 are more understandable. If emotion contagion was not a part of empathy, there should not be strong correlations between empathy questionnaires and questions addressing emotion contagion.

Secondly, Singer and Lamm (2009) separated empathic concern from empathy. In their argumentation empathy referred to feeling others' feelings while empathic concern, with sympathy and compassion, denote that one is feeling for the other person, not feeling the other person's feelings per se. If this truly is the case, the questionnaires and hypotheses used in this study addressed two separate constructs not two sides of the same construct.

However, a multifaceted definition of empathy used in the studies presented in this thesis has gained support from other studies (e.g. Ainley, Maister, & Tsakiris, 2015) and thus it is not fully clear which of the definitions of empathy is the most suitable. Moreover, Ainley and colleagues argued that it is possible that people who have good skills in self/other distinction would themselves share others' feelings to a lesser extent. This in turn could lead to decreased experience of vicarious emotions. In their study, this showed as a non-existent correlation between emotion contagion subscale of Questionnaire of Cognitive and Affective Empathy, RMIE, all the subscales of IRI, and interoceptive awareness. Interoceptive awareness was defined as an extent to which one's internal bodily cues consciously influence behavior, feelings and cognition. Even though interoceptive awareness was not directly addressed in studies presented in this thesis, it could be seen closely related to the theoretical concepts used in them and could therefore have an effect on their findings.

Finally, because emotion contagion subscale was shortened to just two questions in study 2, its findings are very prone to be compromised. That is especially true taking into account the two separated viewpoints of emotion contagion covered by the two separate questions: the first question asked how the participant's feelings were shared by their pair, and second asked how participant shared their pair's feelings

themselves. Thus, only one question referred to the participant's own emotion contagion skills and one to their review of their pair's skills. This is a limited way to address a seemingly complex construct and could have a great effect on the results.

This holds especially because RMIE and IRI are questionnaires addressing participant's own skills. Thus, there was just one question in this shortened version of the emotion contagion scale comparative to RMIE and IRI results. Moreover, this one question did not correlate statistically significantly with RMIE and IRI (see Tables 6 and 7). These results thus remain limited. In future studies it is proposed to use more comprehensive form of this scale to get more complete picture of emotion contagion and its connections to other questionnaires.

4.6. Mentalizing

Mentalizing is the second empathy related subscale in the TT10. Mentalizing was defined as an ability to perceive others as separate individuals who have their own thoughts and emotions (Goldman, 2006). Mentalizing also includes an ability to separate one's own and other's thoughts, feelings and beliefs from reality (Frith & Frith, 2003).

Hypotheses concerning mentalizing did not gain consistent support. Mentalizing was statistically significantly correlated to the achievement scale as hypothesized, but these correlations ranged from moderate to strong in their volume between the two studies of this thesis. They were also a bit lower than other hypothesized correlations between the achievement scale and the other TT10 subscales.

Moreover, there was no support for the hypotheses concerning gender differences in mentalizing scores, even though there seemed to be some differences between differently composed teams in study 1. These differences were still not statistically significant. However, a statistically significant difference between teams was found when dividing teams into two groups based on the percentage of women in them: those, which had 20% or more women and those, which had less than 20% women. However, this was a bit artificial way to study this topic as this division was not based on the previous research. It could be argued that these results still gave support to the hypothesis that mentalizing scores are higher in teams in which there are more women than men. But while there is no evidence that this division should be done exactly by comparing teams that have 20% or more women, the meaning of this result remains somewhat unclear. Especially because other results in these studies did not support gender-based differences in mentalizing.

Even though these hypotheses did not gain strong support, it does not mean that the mentalizing scale is not working. Overall, there have been discussions whether women possess more empathy than men (Singer & Lamm, 2009). One mechanism affecting these results could be that for women it is more socially desirable to show empathy related behavior. However, taking into account the continuously evolving nature of society and gender roles, it is possible that some differences will diminish over time. Thus, even though it was

hypothesized that mentalizing scores would be higher among women and in teams in which there are more women than men, the fact that these hypotheses did not gain support could be due to several factors that were not controlled in this study, that this could not be considered a deficiency of mentalizing scale.

In addition to the previous remarks, mentalizing scale was not correlated either with the RMIE questionnaire or the perspective taking subscale of the IRI in study 2. However, these results could also be due to difficulties in defining and measuring empathy overall, as was the case with emotion contagion. The differences between the questions in the questionnaires used could even strengthen this effect. Moreover, mentalizing, as emotion contagion, was addressed with two types of questions in the shortened version of the TT10 in study 2. There were two questions about participant's own mentalizing skills and two questions about participant's evaluation of their pair's skills. Thus, it remains questionable how much these limitations in the number of questions effect the results, especially because RMIE and IRI are questionnaires regarding only the features of the respondents themselves, and the mentalizing scale addressed two separate scales in a very short way.

All in all, the mentalizing scale was connected to better achievement to some extent, but due the limitations in other results, its validity remains questionable. Taking into account that other questionnaires addressing mentalizing or related constructs are more or less based on the individual's view of themselves, in further versions of the TT10 it would be useful to make mentalizing scale more corresponding to them. Or at least it would be valuable to separate individual mentalizing from mentalizing skills of the group, as the scale which includes questions of these both perspectives could remain vague in its nature. The results of such a scale are not so easily connected to the theoretical background.

4.7. Social presence/attention

Social presence was addressed in this thesis more in a commonplace manner than has been the case in scientific literature in general. Social presence/attention was addressed in the TT10 with questions about how a participant pays attention to other team members and how they pay attention to the participant. A more common way to address the topic in literature has been by studying the direction of gaze. However, it was not possible in a questionnaire-based study.

Two hypotheses could still be made based on this subscale, and both gained strong support from the studies of this thesis. The first hypothesis was that low equality scores would be correlated to lower social presence scores. This was in line with literature that states that those having power tend to overlook lower ranked individuals and use stereotypes when thinking of them (e.g. Overbeck & Park, 2006; Rodriguez-Bailon et al., 2000). The results in study 1 showed that the social presence scores were lower in the teams whose equality scores were low also. Similarly, in study 2 those individuals ranking the equality of their pair low ranked their pair's social presence scores lower than those whose equality scores were higher.

However, some limitations should be mentioned. These limitations concern mostly the equality scale but are discussed here because of their implications to these results. First, in study 2 there were significantly fewer low scoring individuals (n = 13) than higher scoring (n = 98) on the equality scale. This affects the generalization of the results of study 2.

Secondly, equality was addressed with just one question in study 2 and thus its connections to the other scales should be considered very cautiously. Thirdly, pairs in study 2 were automatically unequal in solving the puzzles due to the study design. This naturally could have affected the results. For example, all participants with the lowest equality scores were employees. That might tell about the role of the employees in the sense that clients needed their instruction when solving the puzzles. That might have led in the more passive role of the clients and, thus, employees' feelings of a nonresponsive relationship. Thus, the very likely explanation for the correlation between the low equality scores and low social presence scores in study 2 was related to study design, not to the hypothesized and literature-based explanation that higher ranked individuals paid less attention to the lower ranked individuals as hypothesized (e.g. Overbeck & Park, 2006; Rodriguez-Bailon et al., 2000).

The fourth limitation is that there were no teams with extremely low scores on the equality scale in study 1. Even though the results still supported the hypothesis, a bigger number of unequal teams participating could have made the evidence more reliable. Finally, it should be still mentioned that the TT10's equality scale takes into account just the participant's self-reviewed perspective on their team's equality. Thus, no objectively measured hierarchical, or vertical (Bunderson & Van der Vegt, 2018), structures were addressed in these studies. Actual hierarchical structure should be taken in account if this hypothesis were to be addressed perfectly. That is because without the knowledge about the actual hierarchical structures within teams, the reason for the correlation between social presence and equality scales remains somewhat unclear, as was the case in study 2.

The second hypothesis about the social presence scale was that it would be connected to the mentalizing scale. This hypothesis gained strong support from study 1 but moderate support from study 2. This difference cannot be based solely on the fact that in study 2 pairs were only interacting online. Mentalizing has been shown to be a meaningful entity in both online and face-to-face interactions (Engel et al., 2014). In addition, social attention was required in such a manner that is suitable to both situations. Thus, the whole meaning of this difference between studies remains somewhat unclear. One possibility is that it was related to differences in the study designs. For example, it is possible that social presence did not play a big role in the task of study 2 in which a lot of attention was given to solving the puzzles. The shortened version of the TT10 questionnaire could also have its effect on the results. All in all, the connection between mentalizing and social attention should be further addressed in future studies to investigate whether there really is a connection between these scales or not.

4.8. Psychological safety

Psychological safety was defined as people's perception regarding whether they are allowed to speak openly and express themselves freely in teams or groups (Edmondson, 1999; Edmondson & Lei, 2014). A climate of psychological safety allows teams to learn from their mistakes by discussion and thus improve their performance (Carmeli & Gittell, 2009; Edmondson, 1999). Thus, it was hypothesized that psychological safety is correlated to a better achievement.

This hypothesis gained strong support from both studies, even though in study 2 the correlation was lower than in study 1. This might be due the study design. In study 2, the role of psychological safety was perhaps not as big as it was in study 1, because pairs only solved puzzles during a short time frame. Thus, there was not necessarily many occasions where the members of the pair could have felt a need to propose new ideas. Also, the clear division of roles to employees and clients could have diminished the role of psychological safety in study 2. That is because if it is evident that the other member of the pair is an employee in charge of the situation, it could automatically be easier for the client to seek assistance and support from the employee. Finally, it has been proposed that psychological safety is more important in teams in which tasks are not so clearly defined and in which the need for innovation is thus stronger (Kostopoulos & Bozionelosa, 2011). In study 2 this was not clearly the case as the task was quite simple. This could have diminished the importance of psychological safety.

On the other hand, the complexity of the task of the team was not addressed in study 1 either. However, teams in it represented modern companies and thus it could be assumed that their tasks are not as straightforward as was the case in study 2. Regarding the results of study 1, it should also be pointed out that psychological safety was unexpectedly high in every team that answered the questionnaire. Only one team's mean value was under 4. In that sense, it's difficult to separate teams from each other and in these studies, it was not clear how low psychological safety scores are related to achievement. One of the possible explanations for this bias in results is that only the most successful teams that have great communication strategies answered the questionnaire. It's possible that the questionnaire seemed more intriguing to them than to teams not so fluent in everyday communication.

Other possibility is that the construct of psychological safety was not fully captured with the TT10-questionnaire. However, that seems unlikely as most of the questions were derived from a quite widely used questionnaire by Edmondson (1999). It should be remarked that most of the teams answered the questionnaire in English, even though teams were from Finnish companies and not all the participants hold English as their mother tongue. Therefore, this leaves a possibility of misunderstanding some of the questions. However, a need for studies in populations whose mother tongue is not English has been expressed (Edmondson & Lei, 2014). These studies address this need, even though the results are somewhat limited in their scope.

Another hypothesis concerning psychological safety was that it is not connected to social cohesion. This was based on the remarks of Edmondson (1999) that in teams with a strong cohesion, tendency to speak about problems with the team might diminish. This hypothesis did not gain support from these studies as psychological safety and social cohesion correlated moderately both in study 1 and in study 2. However almost all the scales of the TT10 seemed to correlate with each other and thus it is not clear how much of this is due the real theoretical connections, and how much is related to similar types of questions or participant's tendencies to answer them in a similar fashion.

All in all, the subscale of psychological safety seems to be one of the most promising ones of the TT10's subscales, because it is well grounded on previous questionnaires. New questions can also be seen as a complement to the original questionnaire by Edmondson (1999), bringing more tangible features to it. However, more studies are needed with a bigger variation of participating teams. As long as lower scoring teams on this scale are missing, its validity in the Finnish population cannot be truly addressed.

4.9. Positive orientation

Positive orientation was addressed with questions about mutual respect, encouragement, and a willingness to help each other among the team members. However, the scientific literature on positive orientation suggested that it actually is a part of psychological safety (Carmeli et al., 2009; Carmeli & Gittell, 2009), or at least it leads to stronger psychological safety in teams. Thus, the only hypothesis formed for this subscale was that it is strongly correlated to psychological safety.

This hypothesis gained strong support from both studies. It is yet noteworthy that correlation between subscales was stronger in study 1, and in study 2 positive orientation's correlation between the achievement scale was higher than that for psychological safety. This might have been due to the nature of the questions used. In the positive orientation scale, questions are straightforward and maybe a better fit for a short task like the one used in study 2. It is possible that there is no need for risk-taking or a great worry about being a laughingstock in suggesting something new when people are solving relatively simple puzzles. Participants also knew that they would be in contact with their pair for only a short period of time. Thus, it is possible that questions regarding encouragement and trust fit to this kind of short situation better, and therefore if these elements of interaction emerge, that would lead to a better evaluation of interaction.

Another possibility is that positive orientation, as addressed in the TT10, towards others comes first and lays the foundation for psychological safety to emerge in long term. If this truly was the case, it would not show in a short study like study 2. However, it would be in line with the results of both Carmeli and colleagues (2009) and Carmeli and Gittell (2009) that proposed that psychological safety is based on high-quality relationships that included mutual respect between team members among other descriptors and is hence a noteworthy possibility. Whether this possibility holds, is out of this study's scope. Longitudinal studies should

be arranged to address this connection. Based on the two studies presented in this paper, it could only be said that even though the positive orientation subscale of the TT10 seems to be subordinate to the psychological safety subscale, it might have its own interesting propositions for scientific literature. Whether the TT10 should be just a cross-sectional questionnaire used to give teams information about their current state, or if it was to be used as a scientific instrument used in various study-settings remains to be decided. Depending on this decision, positive orientation subscale could be deleted and some of its questions added to psychological safety subscale, or it could be further developed to address possible longitudinal and theoretical connections.

4.10. Equality

In these two studies, equality was addressed in terms of the hierarchical structure of the teams and differences in power in the team's decision making, as well as equal opportunities to make propositions and being heard by the other team members. Bunderson and Van der Vegt (2018) suggested that the differences between the team members can be of two types: vertical and horizontal differences. They also proposed that all studies addressing equality should take these two aspects into account. However, equality was studied in this paper and in the TT10 in terms of vertical differences only. Horizontal differences (differences in background information) were not addressed, partly due to the participant related limitations. This held especially in study 1, in which the number of participants was very low and in which there were no statistically significant differences in the background information of the teams. And because every participant evaluated only the members of their own teams, and there were no significant differences between the teams regarding their compositions, all the evaluations reflected more or less the features of similar teams. In the future studies, a bigger number of participants and more variation in the set of teams is needed to address the horizontal differences and their effect on the equality scale better.

Hypotheses in study 1 and 2 differed for the equality subscale. In study 1 the hypothesis was that equality is advantageous for the success of the teams. This hypothesis gained strong support. Thus, this result was in line with previous studies (Engel et al., 2014; Renger & Simon, 2011; Woolley et al., 2010) that showed that when teams are not strictly interdependent in solving their tasks, equality leads to better success. However, it should be remarked that the actual tasks of the teams were not addressed in this study. Thus, this result is based on assumption that team members all had a possibility to have a unique influence on their team's success and that they did their parts at least partly individually, so that not all their work was depended on others. In future studies, the actual tasks of the teams should be inquired to the extent it is possible while taking into account anonymity and confidentiality.

Unlike in study 1, in study 2 equality was hypothesized to correlate neither with achievement scale of the TT10 nor with the number of correctly solved puzzles. These hypotheses gained support. Equality was the

subscale of the TT10 that correlated the least with both achievement scale and the number of correctly solved puzzles.

When addressing clients and employees separately, one noteworthy difference was detected. For clients, correlation with equality and the number of correctly solved puzzles was lowly positive and for employees lowly negative. Despite of the fact that both these correlations were in the range of “Little if any correlation” based on the book of (Hinkle et al., 2003), this difference could still tell something about the role difference in this study. It is possible that for clients, perceived equality in communication is more important than for employees who are in charge of the completion of the task. It might be that clients require more feeling of being listened to and paid attention to in order to complete the task more successfully. This proposition is a promising one, especially when taking into account the fact that in the shortened version of the equality scale used in study 2, there was just one question addressing equality: ‘When I expressed something, my pair reacted to it.’ At the same time this same fact is however a big limitation in study 2. This one question hardly captures the whole complex concept of equality. Thus, in future studies a longer version of this scale is suggested to be applied. In studies addressing equality in working environments in general, it is also suggested to pay attention to propositions of Bunderson and Van der Vegt (2018), who stated that equality should always be studied considering both vertical and horizontal differences between individuals. It would also be useful to address the actual hierarchical structures of the teams. The TT10 is a self-rated questionnaire, and both similarities and differences between actual and perceived hierarchies could be interesting to study.

4.11. Shared reflection

No hypotheses were formed for shared reflection scale in either study 1 or study 2. This was because based on the scientific literature, there were no clear ways to address shared reflection in these studies. Shared reflection has been shown both to be (e.g. Gurtner et al., 2007; Schippers et al., 2013) and not to be (e.g. Wiedow & Konradt, 2011) profitable for teams. It has been argued that shared reflection is not useful in cases in which teams are already thinking and acting in the same way (Nederveen Pieterse et al., 2011) or in which teams are already performing well (Schippers et al., 2013). In these both cases reflection can be more time and energy consuming than useful, as there would not be much to reflect upon.

In studies 1 and 2 the performance levels, value differences, and actual working habits of the teams were not addressed at all. Thus, it would have remained unclear, what the shared reflection scores of the teams meant. Well performing teams could have got both high and low scores on this scale, but without the knowledge of their performance and background, these results would have been vague. Moreover, the study design of study 2 did not allow shared reflection to happen within pairs. Thus, the validity of shared reflection

scale remains unclear in these studies and should be addressed in future studies with study designs taking better into account the teams' actual levels of performance.

4.12. What is achieved through interaction

The achievement scale is not an actual subscale of the TT10 questionnaire. It still played a crucial role in the studies presented in this thesis and is thus shortly mentioned here. It was mostly correlated to the subscales of the TT10 in hypothesized ways and hence seemed to be a valid scale itself. Moreover, its reliability was acceptable. However, the achievement scale was not correlated to the number of correctly solved puzzles in study 2, as mentioned above. Thus, it is not clear whether this scale is actually addressing the objectively measured success of the teams or not. It could, for example, reflect more the efficiency of the teams than their actual success in solving tasks. The shortened version of the TT10 might have also led to this result. In future studies using the TT10 it would be wise, if possible, to include objective measures of the success of the teams to further study this finding. In the light of these studies the modification of this scale itself is not recommended as it still was quite strongly connected to the subscales it was hypothesized to be.

4.13. Limitations of the two studies presented in this thesis

4.13.1. Limitations of study 1

The first apparent limitation of study 1 was that the number of participants was very low. Only forty-nine participants from ten teams filled the TT10 questionnaire. Thus, results of this study should be taken cautiously. For example, no factor analysis could be made due to this limited number of participants. Correlations that ranged from moderate to high between different subscales were hence left partly unexplained. Some of these were of course in line with previous research. But when taken in account that almost all the correlations between different subscales exceeded .50, the lowest one being .42, further studies should address the possible factorial structure of the TT10. This not done, it could be argued that the construct validity of the TT10 is limitedly addressed.

On the other hand, these strong correlations between different scales might be due the other clear limitation of study 1: there were no significant differences between the background of the teams. Every team also rated their achievement very high, even though there were statistically significant differences between teams on the achievement scale. However, no low performing teams were participating in this study because in a scale from one to five, 3.80 is to be considered quite high and clearly above median (2.5) of the scale. This bias in answers might be due to the voluntary nature of answering the questionnaire. It is possible that only those teams that were already performing well were able to gather over four participants that were required to get the team in this study to answer the questionnaire. Another explanation would be that

achievement scale itself provokes answers of a certain kind. This, however, is not so logical explanation as the selection bias.

It is also a possible limitation that some of the participants answered the TT10 questionnaire in English and some in Finnish. Thus, there is a possibility of understanding some questions in different ways. Moreover, not all the participants had a possibility to answer the questionnaire in their mother tongue and this could have some effects on their answers.

4.13.2. Limitations of study 2

The first obvious limitation in study 2 is that the TT10 is a questionnaire designed to address team features and in this study, teams consisted of pairs. Even though, in the literature referenced in this thesis, there has not been strict limitations to the compositions of the teams in regarding the number of members, pairs are not necessarily the best entities to apply the TT10 with. Firstly, the TT10 required some modifications to address pairs. Thus, it is not clear whether some information was lost in doing so. Secondly, when addressing pairs instead of teams of multiple members, naturally emerging consequence is that addressed interpersonal relationships are simpler in their nature. Or at least, there is no possibility for team members to relate to other members in different ways because there is only one relationship in a pair.

The second limitation in study 2 was that participants had such a limited and short time in connection to each other. They talked to each other via internet chat just for 20 minutes. It is very questionable how deep interpersonal relationships could occur in such a short time frame. It is also not clear how well a questionnaire like the TT10 addresses interpersonal relationships in this kind of a situation.

Third limitation is that the TT10 questionnaire was shortened remarkably for this study. In the original version of the TT10 there were 84 questions (when including also the achievement scale) and multiple questions concerning each subscale. In the shortened version used in study 2, there were only 27 questions. For example, the subscale of co-flow was addressed just with 2 questions in study 2 compared to 11 questions in the original form. It is questionable whether these questions that remained in the study truly capture the true nature of theoretical construct they are supposed to do. Still, most of the hypotheses gained from moderate to strong support. Thus, even though it could be argued that this version of the questionnaire still worked well, the results could have possibly been more definitive with more information or a longer questionnaire.

4.14. Conclusions

All in all, two studies presented in this thesis gave moderate, and partly strong, support for the validity and reliability of the TT10 questionnaire. In the light of these studies, the most promising ones of the subscales of the TT10 questionnaire were collaboration, co-flow and social cohesion (even though these subscales overlapped to some extent), and psychological safety. All the 10 subscales still had their limitations, but this is not that alarming taking into account that the TT10 is a new questionnaire that was used in these studies for

the first time. All the new questionnaires require fine graining. And in this kind of a questionnaire including a lot of different theoretical constructs, the further modification is especially important. To develop a reasonably short questionnaire that at the same time addresses such a multifaceted theoretical background satisfactorily is not a simple task and requires most likely plenty of further studies.

Nevertheless, these studies are altogether promising for the future of the TT10. An especially noteworthy result was that within the TT10, all the hypothesized scales correlated strongly and positively with the achievement scale. This is a valuable result, as the questionnaire was originally designed to address themes leading to a better interaction and success in teams. This goal seems to be reached even though individual subscales still need modification and further validation from the future studies.

5. References

Ainley, V., Maister, L., & Tsakiris, M. (2015). Heartfelt empathy? No association between interoceptive awareness, questionnaire measures of empathy, reading the mind in the eyes task or the director task. Front. Psychol. 6: 554. doi: 10.3389/fpsyg.2015.00554

Baron-Cohen, S., Jolliffe, T., Mortimore, C., & Robertson, M. (1997). Another advanced test of theory of mind: Evidence from very high-functioning adults with autism or Asperger Syndrome. Journal of Child Psychology and Psychiatry, 38, 813–822.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The “Reading the Mind in the Eyes” Test Revised Version: A Study with Normal Adults, and Adults with Asperger Syndrome or High-functioning Autism. J. Child Psychol. Psychiat. 42, 2, 241–251. https://doi.org/10.1017/S0021963001006643

Bateman, A. & Fonagy, P. (2013). Mentalization-Based Treatment. Psychoanalytic Inquiry, 33:6, 595-613. DOI: 10.1080/07351690.2013.835170

Beal, D. J., Cohen, R. R., Burke, M. J., & McLendon, C. L. (2003). Cohesion and Performance in Groups: A Meta-Analytic Clarification of Construct Relations. Journal of Applied Psychology, Vol. 88, No. 6, 989–1004.

Bilker, W., Hansen, J., Brensinger, C., Richard, J., Gur, R., & Gur, R., (2012). Development of Abbreviated Nine-item Forms of the Raven’s Standard Progressive Matrices Test. Assessment 19, 3, 354–369. https://doi.org/10.1177/1073191112446655

Biocca, F., Harms, C., & Gregg, J. (2001). The Networked Minds Measure of Social Presence: Pilot Test of the Factor Structure and Concurrent Validity. 4th Annual International Workshop of Presence, 21(January), 9–11.

Bunderson, J. S. & Van der Vegt, G. S. (2018). Diversity and Inequality in Management Teams: A Review and Integration of Research on Vertical and Horizontal Member Differences. Annual Review of Organizational Psychology and Organizational Behavior 5: 47–73. https://doi.org/10.1146/annurev-orgpsych-032117-104500

Campbell, J. P. (1990). Modeling the performance prediction problem in industrial and organizational psychology. In M. D. Dunnette & L. M. Hough (Eds.), Handbook of industrial and organizational psychology (2nd ed., Vol. 1, pp. 687–732). Palo Alto, CA: Consulting Psychologists Press.

Carmeli, A., Brueller, D., & Dutton, J. E. (2009). Learning Behaviours in the Workplace: The Role of High-quality Interpersonal Relationships and Psychological Safety. Systems Research and Behavioral Science 26, 81-98. DOI:10.1002/sres.932

Carmeli, A. & Gittell, J. H. (2009). High-quality relationships, psychological safety, and learning from failures in work organizations. Journal of Organizational Behavior 30, 709–729. DOI: 10.1002/job.565

Chikersal P., Tomprou M., Kim Y. J., Woolley A. W. & Dabbish L. (2017). Deep Structures of Collaboration: Physiological Correlates of Collective Intelligence and Group Satisfaction. Proc 2017 ACM Conf Comput Support Coop Work Soc Comput - CSCW '17: 873-888. http://dx.doi.org/10.1145/2998181.2998250

Cross, R., & Parker, A. (2004). Charged Up: Creating Energy in Organizations. Journal of Organizational Excellence: 3–14

Csikszentmihályi M (1990). Flow: The psychology of optimal experience. New York, NY: Harper Perennial.

Csikszentmihalyi M. (2014). Flow and the Foundations of Positive Psychology - The Collected Works of Mihaly Csikszentmihalyi. Springer Dordrecht

Csikszentmihalyi, M., Abuhamdeh, S., & Nakamura, J., (2014). Flow. In Flow and the Foundations of Positive Psychology - The Collected Works of Mihaly Csikszentmihalyi (pp. 227-238). Springer Dordrecht. Originally in 50

Elliot, A. J., & Dweck, C. S., (2005) Handbook of Competence and Motivation. New York: The Guilford Press, pp 598-608

Davis, M. (1980). A Multidimensional Approach to Individual Di erences in Empathy. JSAS Catalog of Selected Documents in Psychology 10, p. 85. https://doi.org/10.1037/0022-3514.44.1.113

Deaner, R. O. & Platt, M. L. (2003). Reflexive Social Attention in Monkeys and Humans. Current Biology, Vol. 13, 1609–1613. DOI 10.1016/j.cub.2003.08.025

de Kort, Y. A. W., Ijsselsteijn, W. A., & Poels, K. (2007). Digital Games as Social Presence Technology: Development of the Social Presence in Gaming Questionnaire (SPGQ). In Presence. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.483.4786&rep=rep1&type=pdf

Deutsch, M. (1949). An experimental study of the effects of cooperation and competition upon group process. Human Relations, 2: 129–152.

de Waal, F. B. M., & Preston, S. D. (2017). Mammalian empathy: behavioural manifestations and neural basis. Nature Reviews Neuroscience, 18(8), 498–509.

Duncan J. & West R. E. (2018). Conceptualizing group flow: A framework. Academic Journals Vol. 13(1), 1-11, DOI: 10.5897/ERR2017.3313

Edmondson A. (1999). Psychological Safety and Learning Behavior in Work Teams. Administrative Science Quarterly, 44: 350-383.

Edmondson, A. C., & Lei, Z. (2014). Psychological Safety: The History, Renaissance, and Future of an Interpersonal Construct. Annu. Rev. Organ. Psychol. Organ. Behav. 1: 23–43.

Engel D., Woolley A. W., Jing L. X., Chabris C. F., & Malone T. W. (2014). Reading the Mind in the Eyes or Reading between the Lines? Theory of Mind Predicts Collective Intelligence Equally Well Online and Face-To-Face. PLoS ONE 9(12): e115212. doi:10.1371/journal.pone.0115212

Frith, U. & Frith, C. D. (2003). Development and neurophysiology of mentalizing. Phil. Trans. R. Soc. Lond. B. Vol 358, 459-473. http://doi.org/10.1098/rstb.2002.1218

Gabelica, C., Van den Bossche, P., De Maeyer, S., Segers, M., & Gijselaers, W. (2014). The effect of team feedback and guided reflexivity on team performance change. Learning and Instruction 34, 86-96.

Gallese, V. & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. Trends in Cognitive Sciences, Vol. 2, No. 12, 493-501.

Goldman, A. I. (2006). Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading. Oxford: Oxford University Press.

Goleman, D. (1998). Working With Emotinoal Intelligence. Bantam Books, New York.

Gurtner, A., Tschan, F., Semmer, N. K., & Nägele, C. (2007). Getting groups to develop good strategies: Effects of reflexivity interventions on team process, team performance, and shared mental models. Organizational Behavior and Human Decision Processes 102, 127–142.

Halevy, N., Chou, E. Y., Galinsky, A. D., & Murnighan, J. K. (2012). When Hierarchy Wins: Evidence From the National Basketball Association. Social Psychological and Personality Science 3(4): 398-406.

Hessels, R. S., Holleman, G. A., Kingstone, A., Hooge, I. T. C., & Kemner, C. (2019). Gaze allocation in face-to-face communication is a□ected primarily by task structure and social context, not stimulus-driven factors. Cognition 184, 28-43. https://doi.org/10.1016/j.cognition.2018.12.005

Hinkle, D. E., Wiersma, W., & Jurs, S. G. (2003). Applied Statistics for the Behavioral Sciences. Houghton Mifflin Company. Boston.

Johnson, D. W., & Johnson, R. T. (2003). Training for cooperative group work. In International handbook of organizational teamwork and cooperative working (pp. 167–183). Wiley Online Library.

Kim, Y. J., Engel, D., Woolley, A. W., Lin, J., McArthur, N., & Malone, T. W. (2015). Work Together, Play Smart: Collective Intelligence in League of Legends Teams. Paper presented at the 2015 Collective Intelligence Conference, Santa Clara, CA.

Klein, J. T., Shepherd, S. V., & Platt, M. L. (2009). Social Attention and the Brain. Current Biology 19, R958–R962 DOI 10.1016/j.cub.2009.08.010

Knapp, R. (2010). Collective (Team) Learning Process Models: A Conceptual Review. Human Resource Development Review 9(3), 285–299.

Kostopoulos, K. C. & Bozionelosa N. (2011). Team Exploratory and Exploitative Learning: Psychological Safety, Task Conflict, and Team Performance. Group & Organization Management 36(3): 385–415.

Lavrakas, P. J. (2008). Encyclopedia of survey research methods. Thousand Oaks, Calif.: SAGE Publications.

LePine, J. A., Piccolo, R. F., Jackson, C. L., Mathieu, J. E., & Saul, J. R. (2008). A Meta-analysis of Teamwork Processes: Tests of a Multidimensional Model And Relationships With Team Effectiveness Criteria. Personnel Psychology, 61, 273–307.

Luca, J. & Cowan, E. (2001). Does emotional intelligence a□ect successful teamwork? In Meeting at the crossroads. Proceedings of the 18th Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education. Melbourne, Australia, 9-12 December 2001

MacDonald, R., Byrne, C., & Carlton, L. (2006). Creativity and flow in musical composition: an empirical investigation. Music and Psychology Research vol 34(3): 292-306.

Mareschal, I., Calder, A. J., & Clifford, C.W.G. (2013). Humans Have an Expectation That Gaze Is Directed Toward Them. Current Biology 23, 717–721. http://dx.doi.org/10.1016/j.cub.2013.03.030

Nederveen Pieterse, A., van Knippenberg, D., & van Ginkel, W. P. (2011). Diversity in goal orientation, team reflexivity, and team performance. Organizational Behavior and Human Decision Processes 114, 153–164.

Overbeck, J. R. & Park, B. (2006). Powerful perceivers, powerless objects: Flexibility of powerholders' social attention. Organizational Behavior and Human Decision Processes 99, 227–243

Preston, S. D. & de Waal, F. B. M. (2002). Empathy: Its ultimate and proximate bases. Behavioral And Brain Sciences 25, 1-72.

Renger, D. & Simon, B. (2011). Social recognition as an equal: The role of equality‐based respect in group life. European Journal of Social Psychology, 41, 501–507.

Rodriguez-Bailon, R., Moya, M., & Yzerbyt, V. (2000). Why do superiors attend to negative stereotypic information about their subordinates? Effects of power legitimacy on social perception. European Journal of Social Psychology 30, 651-671.

Ronay, R., Greenaway, K., Anicich, E. M., & Galinsky, A. D. (2012). The Path to Glory Is Paved With Hierarchy: When Hierarchical Differentiation Increases Group Effectiveness. Psychological Science 23(6): 669–677.

Sawyer, K. (2007). Group Genius – The creative power of collaboration. Revised and updated, 2017 edition. Basic Books, New York.

Schippers, M. C., Den Hartog, D. N., & Koopman, P. L. (2007). Reflexivity in Teams: A Measure and Correlates. Applied Psychology: An International Review 56 (2), 189–211. doi: 10.1111/j.1464-0597.2006.00250.x

Schippers, M. C., Homan, A. C., & Knippenberg, D. V. (2013). To reflect or not to reflect: Prior team performance as a boundary condition of the effects of reflexivity on learning and final team performance. Journal of Organizational Behavior, J. Organiz. Behav. 34, 6–23.

Sharp, C. & Fonagy, P. (2008). The Parent's Capacity to Treat the Child as a Psychological Agent: Constructs, Measures and Implications for Developmental Psychopathology. Social Development, 17: 737-754. doi:10.1111/j.1467-9507.2007.00457.x

Singer, T. (2006). The neuronal basis and ontogeny of empathy and mind reading:Review of literature and implications for future research. Neuroscience and Biobehavioral Reviews 30, 855–863.

Singer, T. & Lamm, C. (2009). The Social Neuroscience of Empathy. The Year in Cognitive Neuroscience 2009: Ann. N.Y. Acad. Sci. 1156: 81–96. doi: 10.1111/j.1749-6632.2009.04418.x

Singh, B., Winkel, D. E., & Selvarajan, T. T. (2013). Managing diversity at work: Does psychological safety hold the key to racial differences in employee performance? Journal of Occupational and Organizational Psychology, 86, 242–263

Tarr, B., Slater, M., & Cohen, E. (2018). Synchrony and social connection in immersive Virtual Reality. Scientifc Reports 8, 3693. https://doi.org/10.1038/s41598-018-21765-4

Terry, P. C., Lane, A. M., & Fogarty, G. J. (2003). Construct validity of the Profle of Mood States - Adolescents for use with adults. Psychology of Sport and Exercise 4, 2, 125–139. https://doi.org/10.1016/S1469-0292(01)

Tsang, S., Royse, C. F., & Terkawi, A. S. (2017). Guidelines for developing, translating, and validating a questionnaire in perioperative and pain medicine. Saudi journal of anaesthesia, 11(Suppl 1), S80–S89. doi:10.4103/sja.SJA_203_17

Wageman R., Hackman J. R. & Lehman E. (2005). Team Diagnostic Survey - Development of an Instrument. The Journal of Applied Behavioral Science Vol. 41 No. 4, 373-398

West, B. J., Patera, J. L., & Carsten, M. K. (2009). Team level positivity: investigating positive psychological capacities and team level outcomes. Journal of Organizational Behavior 30, 249–267

West, M. A. (1996). Reflexivity and work group effectiveness: A conceptual integration. In M. A. West (Ed.), Handbook of work group psychology. Chichester: John Wiley & Sons Ltd., 555–579.

Wiedow, A. & Konradt, U. (2011). Two-Dimensional Structure of Team Process Improvement: Team Reflection and Team Adaptation. Small Group Research 42(1), 32–54.

Woolley A. W., Chabris C. F., Pentland A., Hashmi N. & Malone T. W. (2010). Evidence for a collective intelligence factor in the performance of human groups. Science, vol. 330(6004): 686-688. doi:10.1126/science.1193147.

APPENDIX 1

All of the questions of the TT10 questionnaire

- All of the questions were answered in Likert scale from 1 to 5, in which 1 means refers to not at all agreeing and 5 to totally agreeing on the claim presented.
- (r) in the end of a question refers to reversed question when calculating the scores (i.e. participant's answer of 5 is actually 1, 4 is 2 and so on).
- Bolded questions/claims were the ones used in the shortened version of TT10 in study 2 of this thesis, with the exception that the questions were modified to address pairs rather than teams.
- Some of the questions were taken from previous questionnaires. These questions are marked and references to the original questionnaire are given.

Collaboration

- I feel happy when my teammates succeed.
- I receive help from my team if I need it.
- My teammates and I need each other.
- I try to outdo my teammates.(r)
- I feel envy towards my teammates.(r)
- I do not receive help from my team even when I need it.(r)
- We work together with my teammates.
- I feel like I am competing against my teammates.(r)
- Others try to outdo me.(r)
- We flexibly share our work load.

Co-flow

- We have a shared goal.
- My mind wanders during our interaction.(r)
- I am completely immersed in our task.
- During our interaction I do not focus on other things besides my team members and the task at hand.
- I get excited during our collaboration.
- I get bored during our collaboration.(r)

- During our collaboration, I forget about everything else.
- I feel joy during our collaboration.
- I find our collaboration frustrating.
- I do not enjoy our collaboration.
- Our collaboration makes me feel good.
- I enjoy our collaboration.

Social cohesion

Two of the questions in this scale (marked with *) were adapted from the General Satisfaction Index by Wageman et al. (2005) and two questions (marked with **) were adapted from the Psychological Involvement subscale of a social presence questionnaire developed by de Kort et al. (2007).

- Generally speaking, I am very satisfied with this team.*
- Working with my team members is frustrating. (r)*
- I enjoy my teammates' company.**
- My team members are like friends to me.
- I would gladly spend free time with my team members.
- I feel connected to the others.**
- I do not especially enjoy my team members' company. (r)

Resonance/emotion contagion

Two of the questions (marked with an asterisk) were adapted from the Empathy scale of The Networked Minds measure (Biocca et al. 2001).

- My feelings are transferred to my teammates. For example, when I am happy the others are happy.*
- My teammates' feelings do not affect me.(r)*
- My feelings do not affect my teammates' feelings.(r)*
- My teammates' feelings are transferred to me. For example, when the others are happy I am happy.*
- At times, I notice that we mirror each others postures (for example by leaning forward or positioning our hands the same way).
- My team members' gestures and facial expressions are contagious (for example yawning, nodding, and smiling).

Mentalizing

- I understand others' feelings.
- I think about what others are feeling.
- I think about what others are thinking.
- Others understand my thoughts.
- I put myself in my teammate's position.
- I understand others' thoughts.
- Others understand my feelings.
- Others do not understand my point of view.(r)
- I empathize with my team members.

Social presence/attention

- I focus on what my teammates have to say.
- I listen carefully to my teammates.
- My attention is focused on other things than my teammates or our task during our collaboration. (r)
- Others focus on what I have to say.

Psychological safety

The questions marked with an asterisk were adapted from the psychological safety questionnaire by Edmondson (1999).

- It is completely safe to take a risk during our conversation.*
- I was able to propose things, although I was not sure of their success.
- If I make a mistake, I feel that it is held against me.*
- If I do not know something I can bring it up without a fear of becoming a laughing-stock.
- In this team, it is easy to discuss difficult issues and problems.*
- I can express myself if I think my teammates' suggestions or thoughts are poor without them taking it as an insult.
- At times I feel rejected for being different.(r)*
- It is easy for us to compliment each others' suggestions or thoughts.
- I can present unfinished ideas to my team.
- It is difficult to ask other members of this team for help.(r)*

Positive orientation

- I do not trust my team members.(r)
- I am encouraging towards others.
- Others are encouraging towards me.
- I trust my team members.
- I would help my team members even if it would produce a disadvantage to myself.
- I appreciate my team members.
- I feel appreciated by my team members.

Equality

- Everyone sends messages or talks an equal amount.
- Certain team members make the decisions.
- Everyone has a similar opportunity to make decisions.
- Our messages or speaking turns are usually short.
- We are equal.
- The opinions of certain people are usually more important than those of others.(r)
- I have to interact in a way that does not suit my natural interaction style and preferences.(r)
- One person talks or sends more messages than others.(r)
- When someone expresses something, the others react immediately.
- I can join the conversation with a frequency that suits me best.

Shared reflection

- We evaluate the quality of our interaction together.
- We try to improve the quality of our interaction together if we notice that something does not work.
- We try to improve our ways of working together if we notice that something does not work.
- We evaluate our ways of working together with my team.

What is achieved through interaction

- I learn things during our interaction.
- Collaboration with my team goes well.
- Collaboration leads to the advancement or completion of tasks.
- Our interaction supports problem-solving.
- Interaction makes the advancement or completion of tasks more difficult.