

MSc in People Analytics

Capstone Project:

**Development and validation of the Teamwork Big Five
Questionnaire for Hybrid/Remote teams**

Wordcount: 14135

Author: Gabor Nagy

Date: 11 January 2023

Abstract

In this paper, I present the development of a teamwork effectiveness questionnaire for team development purposes in virtual teams: the Teamwork Big Five Questionnaire for Hybrid/Remote Teams. I analyze the different priorities between surveys used in academic research and team development; I point out the deficiencies of current instruments from the viewpoint of practitioners, and I argue for using a formative instead of a reflective construct. I present the key considerations, decisions, and steps of the development process, the resulting model, and the instrument. The paper also describes the validation process, the validity data collected, and their analysis.

Introduction

Background

This paper is about the development of a questionnaire for practitioners: team development consultants, Human Resources (HR) and business leaders, People Analytics (PA) teams, team leaders and team members. Its main goal is to contribute to the practice of measuring and developing the effectiveness of teamwork in the hybrid/remote teams of organizations. This paper does not answer a research question in a particular organization; it aims to provide a tool that can help answer practical questions from team leaders in many organizations about how to develop their teams.

Teamwork matters for a reason. Teamwork has been recommended by some researchers since the 1980s as a better work organizational structure than a hierarchical organization for all non-routine work (Sundstrom et al., 1990). Indeed, teamwork has been used in some large organizations already in the 90s at the top level (top management teams) and the bottom level (quality circles, operator teams) where the complex challenges exceed the problem-solving abilities of individuals, as Beckmann (1995) has indicated. Fast forward to today's VUCA (volatility, uncertainty, complexity, and ambiguity) world, the need for increased organizational agility and innovation has led to the widespread adoption of teams in many modern organizations.

Hybrid and teleworking are on the rise. The rise of the team as an organizational form has been intertwined with another trend, the development of IT and collaboration technology, which enabled the transformation of intellectual work into telework, and has also made teamwork becoming partly or fully virtual. The move towards teleworking was accelerated by the corona pandemic and has led to the situation today where remote or hybrid teamwork is a fundamental form of work in many organizations. This trend can be illustrated by the rapid growth of the global market size for team collaboration software (15 billion US dollars in 2020, 17 billion in 2021, and forecasted to exceed 40 billion in 2028) according to Fortune Business Insights (2022); or with a Gartner (2021) survey which reported a 44% increase in the use of collaboration software tools in the period of 2019 – 2021.

Hybrid work arrangements and remote work make teamwork more difficult, but the challenges already start with co-located teams. Belbin (1981) described a host of negative dynamics that can hinder team performance, as he observed in a pioneering experimental study over eight years. Lencioni's book, the "The Five Dysfunctions of a Team" (2002), which describes another set of pitfalls that teams face, made it onto many notable bestseller lists (The New York Times Best Seller List, Business Week, Wall Street Journal and USA Today), which indicates that many people resonate with the problems addressed in the book. Leading a group of people and turning them into a high-

performing team has never been an easy task, and now add to that the increasing challenges that team virtuality brings - challenges around communication, planning, coordination, progress monitoring and supporting each other (Martins, Gilson & Maynard, 2004) - and it's no surprise that teams don't always excel.

As teamwork becomes more prevalent, so does corporate interest in improving the effectiveness of teamwork, and there is an entire branch of the training industry devoted to team development, as well as a considerable academic interest in relevant research. Salas et al. (2007) found 130 models related to teamwork effectiveness or some of its components in their literature review and described the field as "the golden age for team research". Salas et al. (2008) suggest that the factors that influence teamwork effectiveness are known, but there is a lack of their measurement instruments. This work aims to address one aspect of the measurement deficit by developing a practice-oriented team diagnostic tool for hybrid/remote teams.

Key definitions

In this paper, a team is defined as two or more people working interdependently toward a common goal (adopted and simplified from Salas et al., 2007).

The name "Teamwork Big Five Questionnaire for Hybrid/Remote Teams" (TBFQHR) refers to the new instrument presented in this paper.

Hybrid/remote teams refer to all teams whose members do not work in the same location. Although these teams are referred to as virtual teams in the academic literature, and according to the Martins, Gilson, and Maynard (2004) research review, many researchers consider virtuality to be a characteristic of all modern teams to some degree, this paper follows the terminology used in human resource management. Teams that work predominantly in the same location and whose members can meet physically on a daily basis are referred to in this paper as "co-located teams" and all other teams as "hybrid/remote teams".

Salas et al. (2005) emphasize the importance of distinguishing between Teamwork Effectiveness and Team Performance, which is a measure of the quantity and quality of team output. Team Performance does not take into account how the team achieved the performance; it is all about the output numbers; it can be affected by conditions beyond the team's control, such as economic conditions or extreme customer demands. Defining and measuring Team Performance is different for a team doing software development, marketing, or operating construction machines. Teamwork effectiveness, on the other hand, refers to the interactions between team members, and many aspects of it are similar across a variety of teams. This paper is focusing on measuring Teamwork Effectiveness.

Objectives of this paper

- Analyzing the comprehensive models of teamwork effectiveness for generic and hybrid/remote teams;
- Analyzing the most relevant existing survey questionnaire instruments;
- Introducing the development process of TBFQHR;
- Presenting and discussing the validity study results.

Significance of the new instrument

Measuring and improving Team Effectiveness with questionnaires means business. Lurey and Raisinghani (2001) have made a number of arguments that team members' perceptions of Team Effectiveness are valid predictors of Team Performance. They argue that because team members perform the work processes themselves, they are the closest to the work, and their perceptions provide a reasonable basis for evaluating Team Performance. They also point out the difficulty of objectively measuring Team Performance since teams often perform non-routine work under constantly changing conditions. In addition, Salas et al. (2005) describe a range of causal relationships to explain what managers intuitively know: Teamwork Effectiveness is an enabler of Team Performance. These arguments all highlight the importance of Teamwork Effectiveness and the surveys that can measure them.

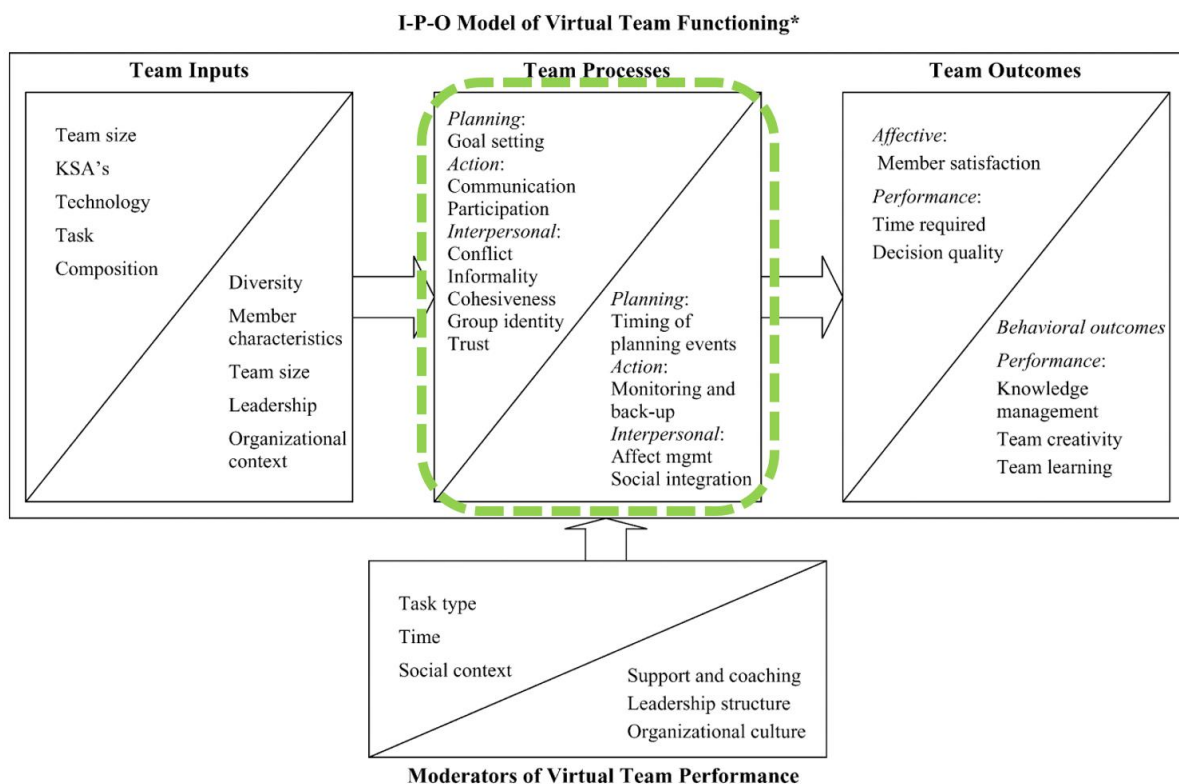
Salas et al. pointed out the lack of instruments to measure teamwork effectiveness as early as 2008, and the current literature review presented in this paper shows that there are still deficiencies on the measurement front. Prior to this work, there was no practice-oriented comprehensive survey on team effectiveness for hybrid/remote teams in the public domain. The purpose of serving team development in organizations as opposed to academic research led to important distinctions and innovative features in the new "Teamwork Big Five Questionnaire for Hybrid/Remote Teams" (TBFQHR) presented.

Literature Review (LR)

The LR presents relevant previous work from the perspective of TBFQHR. For each paper, there is a summary of key findings and critiques. At the end of the LR, there is a summary of lessons learned and an insight into the use of reflective versus formative modeling approaches.

Team Effectiveness models

According to Martins, Gilson, and Maynard (2004), the Input-Process-Outcome model (I-P-O model) developed by Hackman and Morris (1975) has become the predominant framework for considering team functioning holistically at work and this framework is also appropriate for hybrid/remote teams. The following is an illustration of this model by Martins, Gilson, and Maynard (2004):



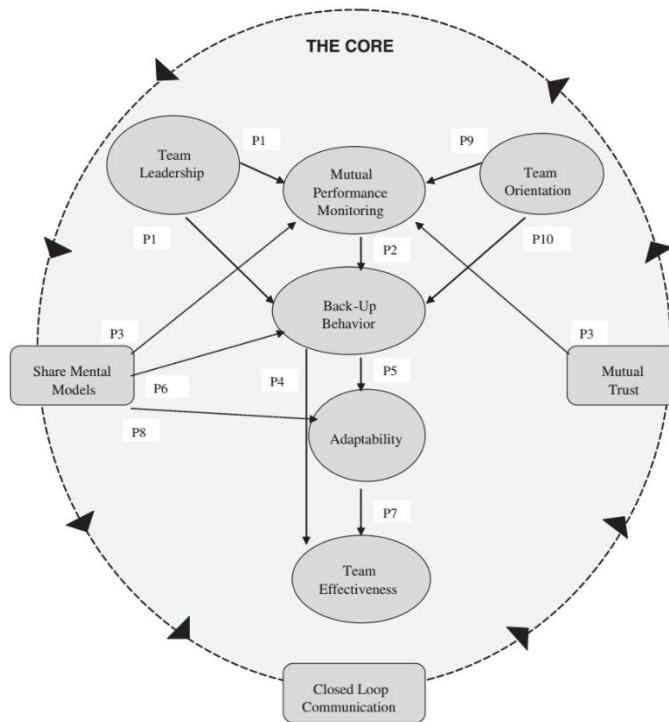
*Within each category of variables, those that have been examined in research on VTs are listed above, and those in need of future research are listed below, the diagonal.

Figure 1. The I-P-O model of team functioning by Martins, Gilson, and Maynard (2004).

The I-P-O model is also relevant to TBFQHR because it can provide a bird's eye view of how teams function. TBFQHR focuses on the middle part "Team Processes" (highlighted in green), because the Team Inputs, Team Outcomes and the Moderators are specific to the team, while team processes are very similar across teams. An additional note: "Team learning" is considered as a team process (conducting learning activities such as reflection, discussions, feedback) in the TBFQHR, while the I-P-O model views "Team learning" as an outcome (the increased knowledge of the team).

The next model, the Teamwork Big Five by Salas et al. (2005), provided an essential input to TBFQHR. Professor Eduardo Salas spent decades on researching teams, he is a co-author of numerous papers and books, and his work is very often referenced by other researchers. Salas et al. (2005) point out that researchers have developed many models of team effectiveness, but there is a lot of overlap

between them; at the same time the team development industry is lacking a comprehensive framework around which to build a professional consensus. In order to create such an integrative framework, Salas et al (2005) conducted a literature review and thematic analysis of previous research on team effectiveness and integrated all important findings into a framework, which they called the "Teamwork Big Five". They crystallized five universal core components of teamwork: Team Leadership, Mutual Performance Monitoring, Backup Behavior, Adaptability, and Team Orientation, enabled by three supporting mechanisms: Shared Mental Models, Closed Loop Communication, and Mutual Trust, making up 8 themes in total. The figure below from their paper illustrates the main hypothetical causal relationships between the components and Team Effectiveness.



Even though the "Teamwork Big Five" model is based on thematic analysis (sorting content into categories by experts), not based on factor analysis or other statistical methods, yet it is outstanding in terms of its integrative ambition. Another achievement of the model is that it highlights the dynamic, adaptive interactions in teams by emphasizing the five core components of teamwork.

From the perspective of team development practice and TBFQHR, there are also points of critique. The definition of Closed Loop Communication is very narrow and only covers a few aspects of effective communication. In addition, the model dates back to 2005 and may not take into account the latest research,

Figure 2. The Teamwork Big Five model, by Salas et al. (2005).

especially the recent findings from the remarkable shift to hybrid/remote working. Salas et al. (2005) have provided detailed behavioral markers that could be further developed to operationalize the model, but there was no questionnaire instrument based on this model prior to the TBFQHR.

The Teamwork Big Five model of Salas et al. (2005) was used in a relevant qualitative study by Hosseini et al. (2018). The researchers interviewed professionals who worked in hybrid/remote teams and asked them open-ended questions about team effectiveness. They then coded the transcripts and counted the frequency of the themes mentioned. They began coding the transcripts using the components of the Teamwork Big Five framework, but created additional themes when other issues were mentioned frequently. The following chart shows the counted frequencies of mentions by code families, codes, and subcodes:

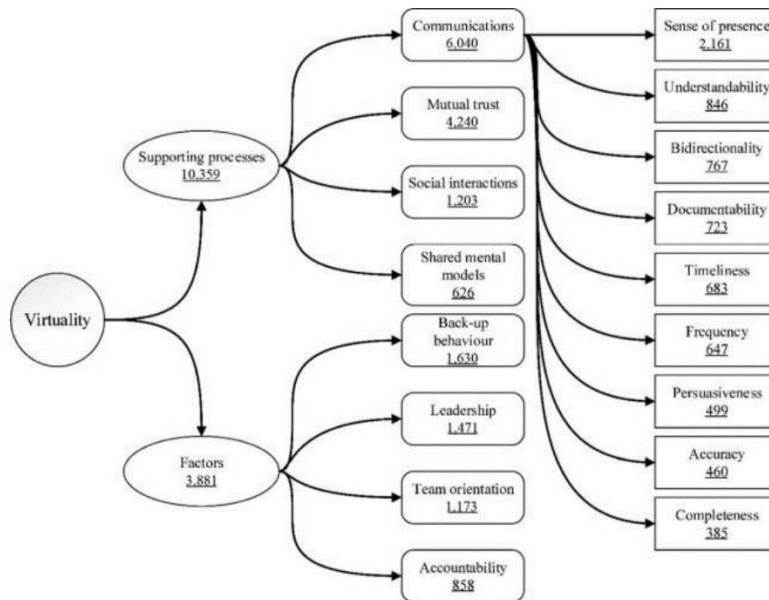


Figure 3. Key topics of team effectiveness in virtual teams by Hosseini et al. (2018).

Frequencies of associated words - the underscored numbers - indicate the relative importance of the topics. It is interesting to compare these results with the original Teamwork Big Five model: first of all, there is much more to communication than Closed Loop Communication; we can also see that the Mutual Performance Monitoring and Adaptivity components have fallen away; and new topics, Accountability and Social interactions, have emerged. A similar approach was used in constructing the themes of the TBFQHR: sorting content first with the Teamwork Big Five framework and then modifying the themes as needed, resulting in a somewhat modified version of the Teamwork Big Five model.

Previously available team effectiveness surveys

There are very few comprehensive survey instruments on team processes. One of them is the Team Effectiveness Questionnaire (TEQ) (Zaffit, Perez, and Adams, 2006), which was developed to measure the team effectiveness of engineering students working in teams at the University of Nebraska-Lincoln. The construct is a reflexive model that includes 7 subscales: 1. Productive Conflict Resolution, 2. Mature Communication, 3. Role Clarity, 4. Accountable Interdependence, 5. Goal Clarification, 6. Common Purpose, 7. Psychological Safety. The factors consist of 4-6 items and have a Cronbach's alpha between 0.513 and 0.745 (n=987). Unfortunately, in the scope of this study information about the development methodology and the items have not been found. It is noteworthy that the correlations between the factors are high, ranging from 0.408 to 0.746.

The most notable existing team survey is the Team Climate Inventory (TCI) created by Anderson and West (1998), which has been used in many studies and was translated into various languages. The survey originates from an earlier theoretical work of the authors, where they identified four dimensions that contribute to the innovativeness of teams: 1. Vision, 2. Participative safety, 3. Task orientation, 4. Support for innovation. The authors then developed a survey questionnaire and confirmed the four-factor structure with factor analysis (even though there are some unexplained item cross-loadings, and Kivimäki et al. (1997) argue for a five-factor solution). The authors also conducted a longitudinal study on the predictive validity of their model, comparing the TCI scores with independent external evaluations, and found the following: Support for innovation was the only valid predictor of overall innovativeness of the teams; Task orientation predicted administrative efficiency; and Participative safety predicted the number of creative ideas. When one reads the TCI items, this makes sense because the content of the items directly relates to these areas. This confirms that team members' perceptions assessed by survey questions correspond well to real-world phenomena, in line with the arguments of Lurey and Raisinghani (2001) and Salas et al. (2005), mentioned before in the Introduction of this paper. This provides further evidence of the usefulness of team questionnaires. It is also noteworthy that all dimensions of the TCI are significantly positively correlated with each other (r between 0.35 - 0.62). The TCI authors mention the possibility of an overall halo effect, which requires further research in other settings. From a practitioner's perspective, there is an uncommon user-friendly approach in the TCI in the way the first 13 items of the questionnaire were worded. Exploratory questions were used, such as "How clear are your team's goals to you?" with specifically worded response options, not just the usual Likert-type statements with options to express levels of agreement or disagreement. This way of formulating items is more laborious, but it is much closer to people's actual conversational habits. For some reason, however, this user-friendly format was not consistently implemented; instead, the usual Likert statements were used in most parts of the 61-item survey. From the viewpoint of TBFQHR, the main criticism of the TCI is that some items are repeated with little variations, for example: Item 38. "People in the team co-operate in order to help develop and apply new ideas"; Item 39. "Members of the team provide and share resources to help in the application of new ideas"; Item 40. "Team members provide practical support for new ideas and their application". Such repetition is common in academic research and increases the Cronbach's alpha values of scales, but is unusual in business communications at work. Participants may rightly feel that such repetitive questions waste their time.

Another remarkable team effectiveness survey is The Five Behaviors of a Cohesive Team (FBCT), based on Lencioni's bestseller book, the "The Five Dysfunctions of a Team" (2002). The survey is part of a team development package along with a personality assessment and facilitated activities (Wiley, 2023). The survey is based on a team effectiveness model consisting of five scales: Trust, Conflict, Commitment, Accountability, and Results. Each scale consists of four Likert-type items. One cannot find a single peer-reviewed scientific paper published on the FBCT instrument. The only validity report (Wiley, 2014) was published by the vendor; the report provides item means, item standard deviations, Cronbach's alpha values of the scales and intercorrelation values between the scales. All dimensions of the FBCT are significantly positively correlated with each other (r between 0.57 - 0.80). What makes this survey unique is its focus on business application and its popularity. The exact numbers are not published, but based on the vendor's partner list and an Internet search, it can be estimated that there are at least a hundred partners worldwide who paid just for the accreditation thousands of dollars plus the recurring fees of \$450 per year, and the total number of people who participated in this survey could be over one million (Wiley, 2023). This is in stark

contrast to the other surveys I have discussed here, which are practically unknown beyond the circle of academic research (the published studies mention maximum few hundred participants). There is a lot to learn from this tool about what makes a survey attractive to business users. The survey is short, with a total of 20 items; all five scales have simple, jargon-free names; the authors didn't employ factor analysis in their questionnaire development methodology; their efforts went into crafting the items and the report: the report includes background information on each questionnaire item and shows the response distribution for each item with an explanation of what that specific result means to the team.

In the scope of this LR, only one comprehensive team effectiveness survey could be found that was specifically developed for virtual teams, the Virtual Team Survey (VTS) by Lurey and Raisinghani (2001). Their survey takes a holistic approach and measures all factors that influence team functioning according to the I-P-O model: Team forming (items 1 -9), Team member selection procedures (items 10-16), Team member relations (17-25), Team processes (items 26-41), Learning and development (items 42-49), Executive leadership style (items 50-55), Communication patterns (items 56-61), Team performance and Team member satisfaction (items 62-70). There are also items about the Reward system, Job characteristics, Organizational culture, Communication tools and technologies and Motivation scattered throughout the questionnaire. The questionnaire uses the format of Likert-type items with five grades. The Team performance and the Team member satisfaction scales are used as criteria for evaluating the validity of the questionnaire. Results: all scales positively correlate with the criteria; Team member relations and Team processes have the highest correlations with criteria (Team performance and the Team member satisfaction scales). The authors interpret these results as evidence for the validity of the VTS. It is noteworthy that after researching so many factors that influence the team functioning in virtual teams, it is the classic group dynamics that stand out the most. From the perspective of TBFQHR, their conclusion is noteworthy:

“...Much of the data resulting from the research suggests that many of the issues that affect virtual teams are similar in nature to those that affect co-located teams. This study has demonstrated that virtual teams are first and foremost teams (Lurey & Raisinghani, 2001).”

They emphasized that building a high-performing virtual team is very similar to building a high-performing co-located team. They also collected qualitative data on the challenges of virtual teamwork and identified communication as a major issue. The criticism of the VTS from the practitioner-oriented TBFQHR perspective is that it contains too many assumptions that may not apply to some teams and some survey participants, for example: Item 1 (I was invited to participate in the formation of this team) and Item 2 (Team members were asked for their suggestions when the team was originally formed) are not relevant for participants who joined the team later; or Item 7 (New team members are quickly brought up to speed when they join the team) is not relevant for teams that don't have new team members.

Lin, Standing, and Liu (2008) set out to integrate previous studies to create a comprehensive model of effective virtual teams. Their study included literature review, meta-analysis of 50 earlier studies, a field experiment, and also quantitative research with a questionnaire. They used Structural Equation Modeling, and confirmed the following relationships between the factors they examined:

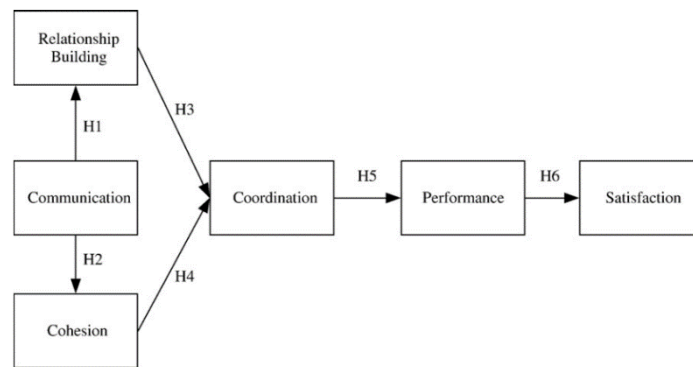


Figure 4. The relationship between the key factors of virtual team effectiveness. Lin, Standing, & Liu (2008).

The strength of their study is the complex research design and sophisticated data analysis which could identify causal relationships between the factors. The main weakness from the viewpoint of TBFQHR is the content validity of their questionnaire design: not all known factors have been addressed by the questionnaire items; there is not a single questionnaire item that explicitly addresses the challenges of virtual work; and they included the item “I am as interested in building a good relationship as in completing the case study” as an item measuring “Communication”, while they have a scale “Relationship Building”. This item is a good example of the ongoing debate about the role of content validity and the evaluation of validity of psychometric instruments in general (Spoto et al., 2023). From the practice-oriented approach of TBFQHR, this item does not represent “Communication” neither “Relationship Building” in a clear and unequivocal way, and it should have been excluded from the outset.

In another relevant recent quantitative study, Alsharo, Gregg, and Ramirez (2017) examined the relationship between knowledge sharing, trust, and teamwork effectiveness in virtual teams. They used Lurey and Raisinghani's (2001) VTS to measure teamwork effectiveness and used scales from various other previously published studies to measure knowledge sharing, trust, and collaboration. This shows that the VTS is still an authoritative questionnaire for measuring teamwork effectiveness in virtual teams.

Challenges of virtual teams

The purpose of this part of the literature review was to summarize the challenges identified by researchers and practitioners related to teamwork in hybrid/remote teams. As mentioned before, Lurey and Raisinghani (2001) concluded that virtual teams are very similar to co-located teams, with the only difference being in the area of communication, but it is worth taking a current look at the issue. Today, hybrid/remote teams are much more common, and many researchers have studied why they are successful or underperforming. Morrison-Smith and Ruiz (2020) conducted a systematic literature review and summarized the key findings from 255 studies; other studies used qualitative research, meta-analysis, experiments, surveys (Lin, Standing, & Liu 2008; Hosseini et al. 2018; Kahlow, Klecka, & Ruppel 2020; Reyes et al. 2021). Due to the TBFQHR's practitioner-oriented approach, this study also included some non-academic sources (Mortinsen & Haas, 2021; Nelson, 2022; Gartner, 2022) and extracted the practical points particularly relevant to teamwork effectiveness in hybrid/remote teams. The recurring topics across the sources have been included in the item content of TBFQHR at the item development stage.

Lessons from the Literature Review

Team surveys in academic research and team development

The LR revealed there are key differences between team effectiveness survey instruments for academic research and team development.

- 1. Participant experience.** In both fields, great emphasis is placed on the ethical use of surveys, such as confidentiality, privacy and data protection. However, in the reviewed academic research papers, the questionnaire items are considered a means to an end, with the end being the quantitative analysis of some construct to answer the research questions. There is no evidence in the papers that the direct impact of completing the surveys on individuals is considered (I am not talking here about clarity, readability and the mandatory ethical principle of doing no harm): There is little discussion about the survey items in the scientific papers, and many papers don't even include the survey items they used. In the practice of team development, however, it is quite different. Any competent HR professional would carefully check whether each item in the survey is appropriate and meaningful for their specific environment before implementing a survey. The practice-oriented FBCT instrument, contrary to the academic ones, uses the items as the key content to educate, generate insights and provide guidance for actions. The commercial success of the FBCT instrument shows the need for recognizing the importance of participant experience, to facilitate valuable reflection through a survey with educational effect. Completing a questionnaire should be viewed as a learning opportunity, in which the questionnaire items play the major role.
- 2. Reporting the survey results to generate actionable insights.** Supporting the particular teams that complete the survey is not the main purpose of academic work, and the scholarly papers don't report on how the survey results were fed back to the participating teams. For example, the reviewed studies had a goal of finding out if the four factors of team climate predict team innovativeness (TCI), or if knowledge sharing contributes to trust (Alsharo, Gregg, & Ramirez, 2017). Regardless of how valuable the results are to researchers, the generic answers to these questions provide limited help for the specific challenges faced by participating teams. As Dykes (2020) points out, the purpose of data analytics in organizations is to drive change; effective teamwork analytics must not only help organizations know how their teams are performing by providing a few overall scores, but also must help every team leader to see where there is room for improvement, and what they should do to improve. The reviewed academic surveys are not optimized for this purpose.

By combining the above observations on the deficiencies of existing instruments, the development direction of TBFQHR becomes clear: the new survey needs questionnaire items that are informative, not repetitive, allow for meaningful reflection, and educate about teamwork effectiveness. The survey items should provide not only a measure of effective teamwork, but also a description, a prescription of effective teams. They should function like an ingredient list of a cooking recipe: They **should contain all the necessary important ingredients of teamwork effectiveness** and be concise at the same time. This is fully in line with the definition of content validity, which requires that all salient theoretical issues related to the construct must be represented (Haynes, Richard, & Kubany, 1995). While companies have a reasonable expectation to minimize the time their employees spend on surveys, a team survey takes less than 10 minutes, which makes maximizing the value they get from such an exercise the priority. Each participating team should receive a report that reinforces what the participants learned during completing the survey, highlights the team's strengths and

areas for improvement, and provides the team with a road map for further development. If the report presents only scale-level results, important details and insights are lost: namely, participants' insights they gained by completing the questionnaire. Coltman et al. (2008) state it like this:

“While aggregating indicators to create a construct achieves the objective of model parsimony, it may come at a significant cost in terms of the loss of the rich, diverse and unique information the individual indicators provide (Coltman et al., 2008)”.

An additional drawback is that when only scale-level results are reported, two additional steps are required: 1. the scales need to be defined, explained, and understood; 2. the scale-level results need to be interpreted, they need to be translated back into practice to explain what it means at the behavioral level. If one works with the items, these additional steps can be saved. All of these points highlight the importance of focusing on the questionnaire items.

Reflective and Formative Models

Why are item-level scores often under-utilized by researchers? As Coltman et al. (2008) point out, academic research in psychology is dominated by the reflective modeling approach. Reflective models assume that latent entities have a causal effect on relevant phenomena, e.g., that general mental ability is the cause of someone scoring high or low on a broad range of mental tasks. Much research is being done to identify the latent entities that are "reflected" in real-world phenomena, and this approach has helped psychology become more empirical. Today, empirical evidence of construct validity through factor analysis is a prerequisite for measurement instruments in the academic world. In this approach, the phenomena represented in the questionnaire items are considered to be samples only. If a questionnaire item does not fit the reflective model, it is considered a "bad item" that is not relevant, gets deleted, and replaced with an item that fits the model better. However, reflective modeling is not the only approach to making sense of the world around us. For example, rather than thinking of "computer literacy" as a latent entity that makes someone use an Internet browser and word processing software efficiently, we can conversely think of "computer literacy" as a constructed concept that is not more than a practical label made up of its components such as being able to use an Internet browser and word processing software efficiently. This approach is called formative modeling, and although it is not commonly used in psychology, it is often applied in economy, marketing, or sociology.

There are proponents of both the reflective and formative models, which approach is better suited to modeling teamwork effectiveness? Coltman et al. (2008) advocate taking a middle ground and incorporating elements from both approaches. The intended practice-oriented focus of the TBFQHR, the goal of providing a recipe-like description and prescription for teamwork effectiveness rather than a contribution to theory, the emphasis on the questionnaire items and less assumption about the construct, all speak for using formative modeling. At the same time, the statistical toolkit of reflective models should be used to analyze the construct and revisit this question based on the results.

Development Methodology

According to Oosterveld, Vorst, and Smits (2019), the four main steps of questionnaire development are as follows: Concept Analysis, Item Creation, Scale Construction, and Evaluation, and the work has

an iterative nature. They also classified the test development strategies, and the main approaches are: Rational/Intuitive (guided by theory and expert knowledge); Inductive/Factor Analytic (starting with a larger item set based on minimal assumptions and extracting the construct using factor analysis); and External/Criteria-driven (starting with a larger item set based on minimal assumptions and selecting the items that best predict the clear criteria previously established). The External/Criteria-driven approach is immediately ruled out in the case of teamwork effectiveness because establishing objective criteria for teams that operate in diverse areas is very difficult (Lurey & Raisinghani, 2001). The Inductive/Factor Analytic approach requires willing respondents who are available to answer longer questionnaires with some items that are not obviously relevant, and in this study, finding participating organizations was expected to be a challenge; another difficulty with this method is that the researcher cannot tell the participating organizations what the report will look like because the factors extracted and measured are not known at the outset; therefore, this approach would meet challenges in this study. The Rational/Intuitive method was selected in this study: there is already knowledge about the factors of team effectiveness (as summarized in the Literature Review); there was access to subject matter experts (SMEs) needed for this method; and the questionnaire designed with this approach can have the greatest learning effect as the experts' knowledge get turned into questionnaire items. The Rational/Intuitive method is also most aligned with the formative modeling approach. This study will also use statistical methods to assess the relationship between items, but factor analysis will not be the principal development method.

Concept Analysis, Item Creation, Theme Construction

The first step in the Rational/Intuitive development process of TBFQHR was the LR introduced before. The LR work generated two main outputs: the list of success factors for co-located and for hybrid/remote teams. The lists were compiled using descriptions of behaviors or specific phenomena that each team member must have experienced instead of higher level concepts like “accountability” or “collaboration”. The behavioral markers from Salas (2005) and the interview scripts from Hosseini (2008) have been particularly useful. Then these two lists were merged into one linguistically coherent, unified list of questionnaire items. The next step was to sort these items into topic groups based on their content. The structure of Salas' Teamwork Big Five and Supporting Factors served as the initial framework, and more Supporting Factors were added as needed. At the end of this stage, TBFQHR had 42 items grouped into 11 themes as displayed in the Appendix Figure A1 (this was not the final TBFQHR).

The above description may sound like a completely rational and linear process, but in reality it was also a creative and intuitive work. Although the work relied heavily on the literature review, the author's personal experiences certainly influenced which aspects were selected and how they were expressed when writing the items. Such personal bias and possible omission of important elements are unavoidable and usually addressed through more researchers working on a topic and SME feedback. This study used the SME feedback strategy.

SME feedback

Ten SMEs were recruited to support the development with their feedback and input; they all worked as team development consultants and also had managerial experience. The following table shows their background:

SME	Senior Consultant	HR	OD	Team leadership experience	Ph. D.
SME1	1	1	1	1	
SME2	1		1	1	
SME3	1		1	1	1
SME4	1		1	1	1
SME5	1		1	1	1
SME6	1	1	1	1	
SME7	1		1	1	
SME8	1	1	1	1	
SME9	1	1		1	
SME10	1		1	1	
Total	10	4	9	10	3

Figure 5. Professional background of the SMEs.

The first version of TBFQHR was sent to them to evaluate item clarity, item adequacy and the tenability of clustering the items into the various themes. All SMEs sent back suggestions for improvement. Some points were criticized quite strongly. The SMEs made many suggestions for items that they felt were missing. Some pointed out additional materials that should be included in the work. There were extended conversations with 3 SMEs to discuss the items and construct in detail. Since the intended purpose of the TBFQHR is to provide a team development tool for practitioners such as these SMEs, their input was seriously considered. This resulted in a thorough revision of the items, and even the themes were changed. As a result, the original themes of Salas's (2005) Big Five were modified, particularly the theme "Leadership," to avoid emphasizing the leader's role as an individual, which is not conducive to group dynamics in team development. For example, instead of using the item "Aligning all team members with the team's goals" and classifying it under the theme "Leadership," the item "Having a shared understanding of how to achieve the team's goals" was used in a new theme called "Shared Goals" instead of Leadership. The concept of "Closed-loop communication" by Salas (2005) was criticized by SMEs on the grounds that it might work in the military, but not in business, and the Communication theme was also significantly updated. The names of some other themes were also changed to use simpler business language. In the end, the TBFQHR had 60 items grouped into 10 themes that were used in this study, this means a 43% increase in the number of items due to the SME feedback. The updated TBFQHR was sent again to the SMEs for a second round of feedback, this time resulting in only minor linguistic adjustments to improve clarity. With all things considered, feedback from the SMEs has contributed significantly to the development of the TBFQHR.

The TBFQHR instrument

These are the final items and themes as a result of the above development steps:

#	Items	Themes
1	Having clearly defined team goals	Shared Goals
2	Having motivating team goals	
3	Having agreement and alignment of all team members with the team's goals	
4	Having a shared understanding of how to achieve the team's goals	
5	Knowing each other's skills, experience, strengths, and limitations	Coordination
6	Having clearly defined roles and responsibilities in the team	

#	Items	Themes
7	Having a reasonable distribution of tasks	
8	Having a clear work plan (with action steps, milestones, sprints,...)	
9	Proactively informing each other about accomplishments, difficulties, and delays we encounter	
10	Being able to quickly bring together ad hoc problem-solving meetings with the relevant people	
11	Placing team goals above our individual interests	Team Spirit
12	Adapting our working styles to each other	
13	Creating a friendly and collaborative team atmosphere	
14	Including all team members in discussions about the team's functioning	
15	Giving each other suggestions and guidance	
16	Helping each other out when someone needs support	
17	Giving each other encouragement and recognition	
18	Resolving disagreements and conflicts constructively	
19	Keeping the team together as a cohesive "One Team" (no cliques, no fault lines, no internal fighting)	
20	Conducting effective team-building activities to keep team spirit high	Attention to Results
21	Having all team members complete their part of the work on time and to a high standard	
22	Reviewing the team's results, achievements, and performance gaps regularly	
23	Actively participating in discussions to improve team performance	
24	Actively seeking feedback from each other and outside of the team to continuously improve	
25	Challenging each other in a positive way by providing constructive, valuable feedback to each other	Adaptability
26	Adapting quickly to changing needs or environments to ensure the long-term success of the team	
27	Addressing problems that arise in a timely manner	
28	Developing new approaches, experimenting with new methods, and continuously improving the way we work	
29	Keeping up to date with the latest knowledge and trends	
30	Regularly reflecting on what has happened in the team and learning from the experiences	Psychological safety
31	Listening to each other	
32	Creating an inclusive team atmosphere in which every member feels accepted and valued, regardless of age, gender, cultural, regional, linguistic, or other personal differences	
33	Creating an atmosphere of psychological safety in which team members can speak their minds, admit mistakes, talk about their emotions, and share divergent ideas	
34	Having in-depth discussions and debates	
35	Having sufficient informal communication in the team	

#	Items	Themes
36	Developing and consistently using communication standards (which channel to use in different situations, how quickly to respond to different messages, etc.)	Effective Communication
37	Using a common vocabulary that all team members understand in the same way	
38	Making good use of asynchronous communication (email and messaging) to allow for fewer interruptions and more focused work	
39	Making sure that messages get received and understood accurately	
40	Avoiding wasted time caused by unnecessary or badly prepared meetings	
41	Conducting effective meetings regardless of where team members are located	
42	Using the video conferencing software effectively	
43	I think the team has the manpower and the skills to be successful	Organizational Support
44	I think the team has the necessary organizational resources (budget, time, autonomy, tools, and access to information) to be successful	
45	I feel well-informed about the broader context of our work (what is going on in the organization, among the internal/external customers,...)	
46	I have access to the information I need to do my job	
47	I am satisfied with the IT (information technology) tools we use in the team	
48	I have no problem getting used to all the different IT systems we need to use in the team	
49	I am satisfied with the degree of autonomy I have over where and when I do my work	
50	I am satisfied with the organization's policies related to hybrid/remote working	Coping with Remote Conditions
51	My home office (or other teleworking) workspace is properly equipped and well-connected to the Internet	
52	I can stay focused on work without too many distractions in my home office (or other teleworking) environment	
53	I can cope with the current workload without the danger of burnout	
54	I can manage the balance between work and private life in the current (hybrid/remote) work model	
55	Currently, I have no problem with the potential stress associated with hybrid/remote work (e.g., less informal interaction, more strenuous communication, screen fatigue)	Motivation
56	I am satisfied with the learning opportunities provided by the team and the organisation	
57	My work in the team is well aligned with both my strengths and interests	
58	I feel that I am making an important contribution to the team with my work	
59	I feel that I get fair recognition for a job well done	
60	I really like my work in this team	

Figure 6. The TBFQHR model.

It is important to note that the model is formative, meaning that the items of primary importance, and the themes are defined by the items. Measurement, reporting, and interpretation (development of actionable findings) should be based primarily on the items, not the themes. The themes are used only to provide an overview of the content, like sections in a bookstore. Such an overview is practical for forming a mental representation of the content, because a list of 60 items is overwhelming to the mind.

The TBFQHR instrument is developed following the psychometric paradigm of Classical Test Theory versus Item Response Theory. There are many reasons for this: most notable are the formative modeling and the emphasis on the educational value of reflecting on the items rather than obtaining a single score. TBFQHR uses a five-point Likert scale in ascending order. There are critiques to using Likert scales: there are disputes whether the collected data must be analyzed as interval or ordinal data; there are concerns about using a middle point; and while certainly there are biases related to Likert scales (primacy effect and acquiescence bias are the most notable), Likert scales have also many advantages: the most proven method of surveying attitudes; writing items with a wide range of content is possible; people in the TBFQHR target group are familiar with completing such surveys; and according to well-researched studies, the disadvantages are not significant (Willits et al. 2016; Chyung et al. 2017; Chyung et al. 2018).

TBFQHR is intended to be user-friendly, and there was an effort to make it more conversational than the traditional Statement + Degree of agreement (Strongly disagree; Disagree; Neither agree nor disagree; Agree; Strongly agree response anchors). The inspiration came from the large set of Likert-type scale response anchors collected by Vagias (2006), and the innovative solution took this form: Generic question + Item + Response anchors; the general question being: "How well is the team doing in the following areas?" + Item (for example Item1: "Having clearly defined team goals") + Response anchors: Big improvement needed; Some improvement needed; Neither good or bad; Good ; Excellent. The advantage of this format is good text flow and great alignment with the purpose of the survey: the TBFQHR wants to explore how well the team is doing in the areas defined by the items, and the questionnaire now asks participants that very question. The disadvantage is that these response anchors have been newly created, they need to be tested and validated. Only the first 42 items could be addressed with this combo of question and response anchors. The miscellaneous questions regarding to working conditions, stress and individual motivation semantically had to be asked differently, and TBFQHR used the traditional Statement + Degree of agreement (Strongly disagree; Disagree; Neither agree nor disagree; Agree; Strongly agree response anchors) format for the remaining 18 items. Regardless of which response anchors are used, the responses always coded as 1, 2, 3, 4, 5, ascending from left to right. The items are all positive statements, to provide the description and prescription of effective teams, there are no negatively keyed items.

Ethical perspective

Using questionnaire-based surveys in the workplace can have negative effects on participants that need to be considered. Beyond the obvious GDPR and data protection requirements, the use of TBFQHR may raise other practical issues that need to be considered, which are summarised in the table below:

Risk	Measures to minimize the risk
An actor with potentially bad intention may try to use the responses against individual participants	The questionnaire items don't ask any confidential or sensitive data; TBFQHR will be administered either anonymously, without collecting name or other personal identification data, or the personal identification data will be deleted as soon as participants completed the survey. Only anonymized responses will be evaluated. The privacy protection must be ensured by the online questionnaire platform; Only aggregated reports will be generated, and only teams with 5 or more participants will receive reports.
Some respondents may have negative feelings about their team (may feel isolated, frustration with colleagues or leadership, or may have received unfair treatment), and the questionnaire may recall these negative experiences, and may generate additional negative thoughts and emotions.	Users of TBFQHR will be informed about these risks. HR professionals and team leaders need to pay special attention to low scores on items related to inclusion, stress and mental health, and follow up with field observation and one-to-one interviews.
If respondents don't get any feedback after the survey, and if they perceive that nothing has been done to follow up the survey, they may get disappointed and their employee engagement may get reduced.	Users of TBFQHR will be informed about this risk, and they will be trained in conducting a follow up workshop with their teams.

Participation should always be voluntary; organizations, teams, and individuals must be able to refuse to participate without facing consequences. Participation should be motivated by benefits (learning about the key factors of effective teamwork; receiving a roadmap as a team to further accelerate team development), not coercion.

Validity Research Methodology

This study will employ a validity research strategy of conducting statistical analysis of responses collected from employees working in hybrid/remote teams. The analysis will look at the response patterns to each TBFQHR items, among the items, and among the items and two additional criteria questions. The following analysis will be performed:

Item descriptive statistics

It is important to list the items with the highest/lowest means and standard deviations, to check whether the result is reasonable, or due to an error such as the wording of the item. Extreme values may point to problems.

Item correlation with the total score

If all TBFQHR items indeed measure components of teamwork effectiveness, then a significantly positive correlation is expected between all items and the total score, and as a minimum criterion, no negative correlations expected.

Item correlations with criteria

Except for the 60 questionnaire items that make up the TBFQHR model, responses to two additional overall evaluation questions will be also gathered to serve as criteria, one is about the overall experience of team members, and the other one is about team performance:

Criterion1: "Please rate your overall experience of working in this team from 1 to 10: (1=Bad experience, 10=Great experience)"

Criterion2: "Please rate the overall work performance of the team from 1 to 10 in terms of the quantity, quality, and value of output: (1=Poor, 10=Excellent)"

As Diamantopoulos, Sarstedt, and Fuchs (2012) pointed out, multi-item scales are more reliable than single-item scales in general. At the same time, there are also arguments in favor of using single-item scales: they are short and practical, and in some cases, they measure nearly as accurately as multi-item scales. In this study, single-item scales were used because the TBFQHR, with its 60 items, is not a short survey and an additional multi-item scale for criteria was not attractive; the 1-10 rating is commonly used in management and marketing; the 1-10 rating format is somewhat different from the 5-point Likert scales of the TBFQHR, and reducing the covariance caused by the biases stemming from using the same format in the TBFQHR and the criteria was a priority.

The higher and more positive correlations between the TBFQHR items and the criteria are signs of validity; no negative correlations expected. It is also expected that items with "feel good" content will correlate higher with Criterion1, and items with more performance-oriented content will correlate higher with Criterion2.

Internal Consistency and Factor Analysis

These statistics will be performed to better understand the inter-item correlations and potentially discover an underlying empirical factor structure. There are no explicit expectations because of the formative modeling approach does not require these evaluation methods.

Inter-rater Reliability

While it is normal for members of the same team to have somewhat different perceptions on their teams, we would expect to see common trends in their responses, therefore, examining the consistency of responses within teams is also a possible measure of validity. Such consistency is referred to as inter-rater reliability, and in the case of Likert-item responses from non-randomly selected multiple raters, the recommended method of evaluation by Koo & Li (2016) is the Intraclass

Correlation Coefficient (ICC) using two-way mixed modeling and following the 95% confidence interval convention in reporting results. The expected ICC average minimal value is above 0.5 (values less than 0.5 indicate poor, between 0.5 and 0.75 moderate, between 0.75 and 0.9 good, and greater than 0.90 excellent inter-rater reliability) (Koo & Li, 2016).

Data collection

The data collection was organized through the HR departments or the management of the participating companies. 16 companies from different industries participated in the survey. Geographically, the companies are located in Germany, Hungary, Italy, Pakistan and the United Kingdom. The following table shows the basic information of the participating companies. The company names are hidden for confidentiality reasons, and the organizations are represented by letters of the alphabet:

	Code	Industry	Teams	Participants
1	A	Consulting	1	5
2	B	Engineering	1	7
3	C	Manufacturing	1	5
4	D	Engineering	1	9
5	E	Services	1	10
6	F	Services	1	5
7	G	IT	1	4
8	H	Chemicals	2	13
9	I	Property	1	6
10	J	Consulting	1	5
11	K	Consulting	1	7
12	L	Education	1	8
13	M	Services	1	5
14	N	IT	6	88
15	O	Infrastructure	3	27
16	P	Consumer products	23	100
Total:			46	304

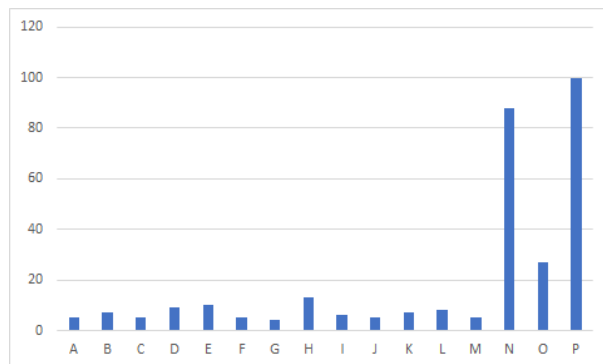


Figure 7. Participating organizations.

The chart shows visually the distribution of participants per organization. Gender distribution of participants was Female 57%, Male 38%, Unknown 5%.

Team sizes: Participants came from 46 different teams, with a maximum size of 31 and a minimum size of 1, with the median size being 5. Team sizes refer to the number of people who actually participated in the survey, not the number of people working in those teams. The response rate in organization N was 87% and in organization P it was 64%; in the other smaller organizations with 1-3 participating teams, it was 100% or close to it. The chart on the right shows the distribution of participants per team.

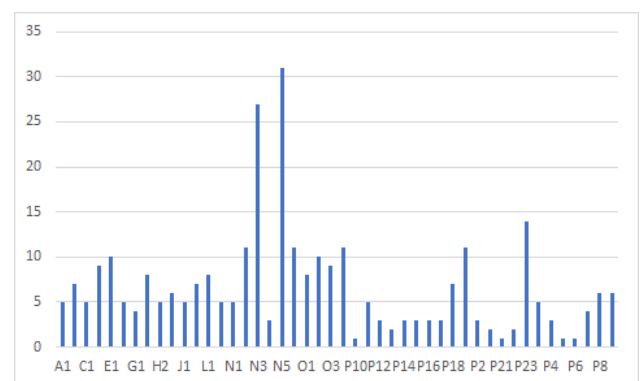


Figure 8. Team sizes chart.

Data analysis results and discussion

In this section, the results of the analysis of the collected data are presented and discussed from the perspective of the validity of TBFQHR. The relevant results of the analysis are included in the text and not in the Appendix, whenever this was possible for better readability. For the same reason, the detailed discussions directly follow the results and a summary of the main points can be read in the next section "Conclusions and Recommendations".

The analysis was carried out using Microsoft Excel 2000 and IBM SPSS v20.

Data validation

304 participants completed the questionnaire and all did it online, which ensured that there were no missing data or unclear answers. Unfortunately, the online platform does not provide information on questionnaire completion time, and response sets with potentially unrealistically fast completion times such as 1 or 2 minutes could not be checked and eliminated. The standard deviations of the responses to the 60 items for each participant were checked and three people answered with a standard deviation of zero: two answered all questions with the highest score (5) and one with the middle option (3). It is extremely unlikely that anyone answering this survey accurately would respond this way. But what if their answer expressed their overall, true opinion? Considering that they are from different teams (N2, N5, N6) and it is 3 out of 304 cases, which is less than 1%, and that I do not know their processing times, their data was not deleted. By keeping these data, the results of the following analysis will not be improved, on the contrary, if anything, all the indicators examined will be slightly weaker. For the same reason, there was no further check for other suspicious response patterns.

The sample size is adequate for the intended analysis, even exploratory factor analysis can be considered (de Winter, Dodou, & Wieringa, 2009).

Item descriptive statistics

It is important to list the items with the highest/lowest means and standard deviations, to check whether the result is reasonable, or due to an error in wording the item. Extreme values may point to problems. The full table is shown in the Appendix Figure A2.

Top ten items with the highest means (n=304):

	Items	Means
Item51	My home office (or other teleworking) workspace is properly equipped and well-connected to the Internet	4.2928
Item52	I can stay focused on work without too many distractions in my home office (or other teleworking) environment	4.2697
Item16	Helping each other out when someone needs support	4.2303
Item60	I really like my work in this team	4.2039
Item58	I feel that I am making an important contribution to the team with my work	4.1349
Item48	I have no problem getting used to all the different IT systems we need to use in the team	4.1316
Item43	I think the team has the manpower and the skills to be successful	4.0921
Item13	Creating a friendly and collaborative team atmosphere	4.0757
Item54	I can manage the balance between work and private life in the current (hybrid/remote) work model	4.0757
Item55	Currently, I have no problem with the potential stress associated with hybrid/remote work (e.g., less informal interaction, more strenuous communication, screen fatigue)	4.0493

Bottom ten items with the lowest means (n=304):

	Items	Means
Item29	Keeping up to date with the latest knowledge and trends	3.5197
Item8	Having a clear work plan (with action steps, milestones, sprints,...)	3.4770
Item22	Reviewing the team's results, achievements, and performance gaps regularly	3.4671
Item24	Actively seeking feedback from each other and outside of the team to continuously improve	3.4572
Item7	Having a reasonable distribution of tasks	3.4539
Item28	Developing new approaches, experimenting with new methods, and continuously improving the way we work	3.4507
Item4	Having a shared understanding of how to achieve the team's goals	3.4309
Item3	Having agreement and alignment of all team members with the team's goals	3.4276
Item20	Conducting effective team-building activities to keep team spirit high	3.3947
Item2	Having motivating team goals	3.3586

The value range of the items is between 1 and 5, and there are no items with suspiciously low mean values. The items with the highest means have higher mean values than expected, but upon closer examination of the items, it is quite conceivable that the participating teams actually achieved high levels in these areas. Most of these items are related to relatively “easier” issues: physical infrastructure, working conditions, motivation, mental health, and mutual help between team members; and in our participating companies these issues could be indeed well addressed. The high values, however, could be also signs of acquiescence bias or social desirability. Items with the innovative Likert response anchor performed well; none of them has ended up in the highest top ten group, suggesting that the new format is not particularly susceptible to acquiescence or social desirability bias.

The standard deviation values show how differently the participants responded to the items. Low values may indicate a problem here: If all participants answered exactly the same, the item has no added value because it doesn't differentiate between teams with high or low teamwork effectiveness.

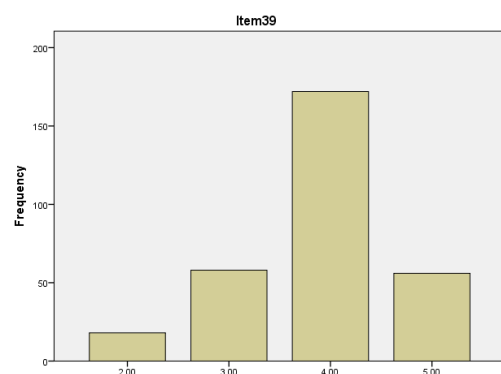
Top five items with the highest standard deviations (n=304):

	Items	Standard Deviations
Item33	Creating an atmosphere of psychological safety in which team members can speak their minds, admit mistakes, talk about their emotions, and share divergent ideas	1.13566
Item2	Having motivating team goals	1.13440
Item6	Having clearly defined roles and responsibilities in the team	1.12527
Item20	Conducting effective team-building activities to keep team spirit high	1.12375
Item19	Keeping the team together as a cohesive "One Team" (no cliques, no fault lines, no internal fighting)	1.12228

Bottom ten items with the lowest standard deviations (n=304):

	Items	Standard Deviations
Item48	I have no problem getting used to all the different IT systems we need to use in the team	.85738
Item51	My home office (or other teleworking) workspace is properly equipped and well-connected to the Internet	.85393
Item16	Helping each other out when someone needs support	.85187
Item46	I have access to the information I need to do my job	.84575
Item55	Currently, I have no problem with the potential stress associated with hybrid/remote work (e.g., less informal interaction, more strenuous communication, screen fatigue)	.82105
Item37	Using a common vocabulary that all team members understand in the same way	.82050
Item60	I really like my work in this team	.81923
Item58	I feel that I am making an important contribution to the team with my work	.79907
Item57	My work in the team is well aligned with both my strengths and interests	.78213
Item39	Making sure that messages get received and understood accurately	.77342

The difference between the strongest and weakest item is 0.37. There is no established hard limit for what can be considered an acceptable minimum standard deviation for a single five-point Likert item, but we can look at the histogram of responses of the weakest item, Item39, on the right. We can see that even this item generates some variance. The next question is whether the variance generated is meaningful, and to answer that question we need to look at the correlations.



Items with the innovative Likert response anchor performed well, earning the top positions; and only two of them have ended up in the bottom ten group, suggesting that the new format has performed at least as well as the traditional response anchors in generating variance.

Correlations with the total score

Careful observation of correlations is important in establishing construct validity of formative models. If all TBFQHR items indeed measure components of teamwork effectiveness, then a significantly positive correlation is expected between all items and the total score, and as a minimum criterion, no negative correlations expected. The results show that indeed all items are positively correlating with the corrected total score (corrected means that the item investigated was left out from calculating the total score to remove the item's own variance), as shown in Appendix Figure A3. In addition, even the inter-item correlation table doesn't have a single negative number, as shown in Appendix Figure A4.

Top ten items with the highest correlation values with the (corrected) total score:

Item33	Creating an atmosphere of psychological safety in which team members can speak their minds, admit mistakes, talk about their emotions, and share divergent ideas	0.752806
Item19	Keeping the team together as a cohesive "One Team" (no cliques, no fault lines, no internal fighting)	0.730337
Item2	Having motivating team goals	0.728291
Item41	Conducting effective meetings regardless of where team members are located	0.722519
Item26	Adapting quickly to changing needs or environments to ensure the long-term success of the team	0.719263
Item32	Creating an inclusive team atmosphere in which every member feels accepted and valued, regardless of age, gender, cultural, regional, linguistic, or other personal differences	0.718659
Item23	Actively participating in discussions to improve team performance	0.718281
Item25	Challenging each other in a positive way by providing constructive, valuable feedback to each other	0.715071
Item3	Having agreement and alignment of all team members with the team's goals	0.707041
Item4	Having a shared understanding of how to achieve the team's goals	0.705718

Bottom ten items with the lowest correlation values with the (corrected) total score:

Item38	Making good use of asynchronous communication (email and messaging) to allow for fewer interruptions and more focused work	.513
Item43	I think the team has the manpower and the skills to be successful	.510
Item49	I am satisfied with the degree of autonomy I have over where and when I do my work	.491
Item47	I am satisfied with the IT (information technology) tools we use in the team	.460
Item54	I can manage the balance between work and private life in the current (hybrid/remote) work model	.459
Item50	I am satisfied with the organization's policies related to hybrid/remote working	.429
Item53	I can cope with the current workload without the danger of burnout	.415
Item55	Currently, I have no problem with the potential stress associated with hybrid/remote work (e.g., less informal interaction, more strenuous communication, screen fatigue)	.405
Item52	I can stay focused on work without too many distractions in my home office (or other teleworking) environment	.315
Item51	My home office (or other teleworking) workspace is properly equipped and well-connected to the Internet	.296

It is worth to note that 9 items with the lowest correlations are from the second part of the survey, related to the working conditions and motivation; and Item38 is related to virtual teamwork. It shows that the items that specifically address the challenges of virtual teams tend to contribute slightly less variance to the total score than the traditional team effectiveness components.

All items exceed the criterion of not negatively correlating with the total score.

Item correlations with criteria

Except for the 60 questionnaire items that make up the TBFQHR model, responses to two additional overall evaluation questions were also gathered to serve as criteria.

Criterion1: "Please rate your overall experience of working in this team from 1 to 10: (1=Bad experience, 10=Great experience)". Item correlations with the Team Member Experience score are between 0.63-0.18, please refer to Appendix Figure A5.

Top ten items with the highest correlation values with the Team Member Experience score:

Nr.	Item	Correlation w Team member experience
60	I really like my work in this team	0.63
19	Keeping the team together as a cohesive "One Team" (no cliques, no fault lines, no internal fighting)	0.59
57	My work in the team is well aligned with both my strengths and interests	0.58
13	Creating a friendly and collaborative team atmosphere	0.57
33	Creating an atmosphere of psychological safety in which team members can speak their minds, admit mistakes, talk about their emotions, and share divergent ideas	0.56
2	Having motivating team goals	0.54
31	Listening to each other	0.54
32	Creating an inclusive team atmosphere in which every member feels accepted and valued, regardless of age, gender, cultural, regional, linguistic, or other personal differences	0.54
14	Including all team members in discussions about the team's functioning	0.53
15	Giving each other suggestions and guidance	0.53

Bottom ten items with the lowest correlation values with the Team Member Experience score:

	Items	Correlation w Team member experience
28	Developing new approaches, experimenting with new methods, and continuously improving the way we work	0.34
53	I can cope with the current workload without the danger of burnout	0.34
48	I have no problem getting used to all the different IT systems we need to use in the team	0.32
54	I can manage the balance between work and private life in the current (hybrid/remote) work model	0.32
55	Currently, I have no problem with the potential stress associated with hybrid/remote work (e.g., less informal interaction, more strenuous communication, screen fatigue)	0.31
38	Making good use of asynchronous communication (email and messaging) to allow for fewer interruptions and more focused work	0.28
47	I am satisfied with the IT (information technology) tools we use in the team	0.27
50	I am satisfied with the organization's policies related to hybrid/remote working	0.26
51	My home office (or other teleworking) workspace is properly equipped and well-connected to the Internet	0.19
52	I can stay focused on work without too many distractions in my home office (or other teleworking) environment	0.18

All items positively correlate with the Team Member Experience score. Significance: Item52 $p=0.002$, Item51 $p=0.001$, and the other items are even more significant. Items with the innovative Likert response anchor performed well, they tend to be more in the top group than in the bottom.

Criterion2: “Please rate the overall work performance of the team from 1 to 10 in terms of the quantity, quality, and value of output: (1=Poor, 10=Excellent)”. Item correlations with this Team Performance score are between 0.58-0.07, please refer to Appendix Figure A6.

Top ten items with the highest correlation values with the Team Performance score:

	Items	Team Performance
Item21	Having all team members complete their part of the work on time and to a high standard	0.58
Item26	Adapting quickly to changing needs or environments to ensure the long-term success of the team	0.55
Item32	Creating an inclusive team atmosphere in which every member feels accepted and valued, regardless of age, gender, cultural, regional, linguistic, or other personal differences	0.51
Item34	Having in-depth discussions and debates	0.51
Item2	Having motivating team goals	0.5
Item19	Keeping the team together as a cohesive "One Team" (no cliques, no fault lines, no internal fighting)	0.5
Item33	Creating an atmosphere of psychological safety in which team members can speak their minds, admit mistakes, talk about their emotions, and share divergent ideas	0.5
Item3	Having agreement and alignment of all team members with the team's goals	0.49
Item29	Keeping up to date with the latest knowledge and trends	0.49
Item4	Having a shared understanding of how to achieve the team's goals	0.48

Bottom ten items with the lowest correlation values with the Team Performance score:

	Items	Team Performance
Item53	I can cope with the current workload without the danger of burnout	0.33
Item37	Using a common vocabulary that all team members understand in the same way	0.31
Item58	I feel that I am making an important contribution to the team with my work	0.31
Item47	I am satisfied with the IT (information technology) tools we use in the team	0.28
Item49	I am satisfied with the degree of autonomy I have over where and when I do my work	0.26
Item55	Currently, I have no problem with the potential stress associated with hybrid/remote work (e.g., less informal interaction, more strenuous communication, screen fatigue)	0.26
Item48	I have no problem getting used to all the different IT systems we need to use in the team	0.24
Item50	I am satisfied with the organization's policies related to hybrid/remote working	0.2
Item52	I can stay focused on work without too many distractions in my home office (or other teleworking) environment	0.13
Item51	My home office (or other teleworking) workspace is properly equipped and well-connected to the Internet	0.07

There are no negative correlations. Except for Item51 and Item52, all items positively correlate with the Team Performance score. Significance: Item50 $p=0.001$, and the other items are even more significant. In these teams, the objective work conditions correlate less with the perceptions on Team Performance.

Items with the innovative Likert response anchor performed well, they tend to be more in the top group than in the bottom.

Internal consistency

In reflective models, the Cronbach's Alpha value, which is a measure of internal consistency of scales, is often used as an estimate of test reliability. The expected Cronbach's Alpha values depend on the type of scale and the number of items, but usually a value above 0.7 is considered acceptable and above 0.8 excellent. Although the Cronbach's Alpha measure is not required for evaluating formative models that do not require a high degree of internal consistency, it can still be informative to calculate Cronbach's Alpha values for the entire questionnaire and also for the themes in the TBFQHR model. This shows how coherent (focused) or diverse (broad) content is addressed in each theme. If we consider the 60 items as a single scale of teamwork effectiveness, the Cronbach's Alpha value of TBFQHR would be 0.974 (n=304). This may sound like an excellent figure, and it does indeed show that whatever the items measure together, they measure it in a reliable way; at the same time, this superscale contains 60 items, which is an extremely high number for a scale. The Cronbach's Alpha values of the ten themes are:

Theme	Number of items	Cronbach's Alpha
Shared Goals	4	0.91
Coordination	6	0.84
Team Spirit	10	0.92
Attention to Results	5	0.87
Adaptability	5	0.87
Psychological safety	5	0.89
Effective Communication	7	0.89
Organizational Support	8	0.82
Coping with Remote Conditions	5	0.74
Motivation	5	0.85

According to these alpha values, all themes contain fairly coherent topics. If they had been designed using the reflective methodology, the themes would pass as reliable scales measuring latent constructs. Does this mean that the intuitive item sorting method used in forming the TBFQHR themes resulted in an empirically valid structure of themes? Well, not really. As it was mentioned before, all of the items in the TBFQHR are positively correlated with each other, and there are many possible ways to group them into themes to achieve acceptable Cronbach's Alpha values. Performing a factor analysis is needed to empirically identify an underlying structure. Here it can be concluded from the Cronbach's Alpha values that sorting the 60 items into the ten themes did not produce themes with empirically disparate content.

Factor analysis

Even though TBFQHR was developed with formative modeling, using factor analysis to better understand the empirical factor structure of the TBFQHR model may be helpful in identifying problems or further development. Factor analysis may also add to our knowledge about teamwork effectiveness: If teamwork effectiveness is strongly affected by some distinctive factors, then they should show up as latent constructs reflected in the variance of the TBFQHR data. The overall Kaiser-Meyer-Olkin (KMO) factor adequacy test score of the data is 0.951, indicating that there is enough commonality in the data (0.80 to 0.89 meritorious, 0.90 to 1.00 excellent). Bartlett's Sphericity Test yielded a p. value of 0, which is less than 0.05, indicating that there is sufficient correlation between the variables for factor analysis. Multicollinearity of the data is an issue that needs to be considered. Because the scales of the other reflective instruments reviewed are also significantly correlated with each other, the oblimin rotation was used to extract the components, which allows the factors to correlate with each other. The resulting solution contains 10 components with eigenvalues greater than 1. Item loadings on the factors are shown in the component matrix in the Appendix Figure A7. The first factor explains 40% of the variance, and every item is significantly correlated with this first factor. Because so many items load on this factor, the nature of this factor cannot be determined in the scope of this study. This might be similar to the overall halo-effect mentioned by Anderson and West (1998), when reporting the factor analysis results of their TCI data, and it may be an interesting research topic for the future. The second factor explains 5% of the total variance, with the items relating to satisfaction with working conditions loading on this factor. The third factor explains 4% of the total variance, with mainly the items related to communication loading on this factor; the fourth factor explains 3% of the total variance, with the items related to motivation loading mostly on this factor. The other 6 extracted factors together explain 13% of the variance, 2-3% each, and the highest loading items also load on other factors.

Inter-rater Reliability

While it is normal for members of the same team to have somewhat different perceptions on their teams, we would expect to see common trends in their responses. If team members' ratings of their teams are not consistent at all, it is likely not only a problem with their observational skills, but also a problem with the rating method (in this case, the TBFQHR questionnaire). Therefore, examining the consistency of responses within teams is also a possible measure of validity. Such consistency is referred to as inter-rater reliability, and in the case of Likert-item responses from non-randomly selected multiple raters, the recommended method of evaluation by Koo & Li (2016) is the Intraclass Correlation Coefficient (ICC) using two-way mixed modeling and following the 95% confidence interval convention in reporting results. I need to exclude the teams with only one member (inter-rater reliability cannot be applied by definition); and I know that the two large teams in organization N (N3 and N5) have a complex sub-team structure, so I will also exclude them from this statistic; and I also exclude the 3 participants with responses of 0 standard deviation because in these small datasets their data would have relatively more weight, which may cause a larger error. That leaves us 40 teams with a total of 240 participants. I calculated the ICC value for each 40 teams and the mean ICC is **0.71**. Values less than 0.5 indicate poor, between 0.5 and 0.75 moderate, between 0.75 and 0.9 good, and greater than 0.90 excellent inter-rater reliability (Koo & Li, 2016). I included the results table as Appendix Figure A8.

Conclusion and Recommendations

Key findings

The TBFQHR instrument was developed with a practice-oriented, content-centred approach for measuring and developing team effectiveness in hybrid/remote teams. It used the formative modeling approach and 42 of its 60 items feature an innovative version of Likert scale response anchors. The development included extensive qualitative SME feedback.

The validity of the questionnaire was examined by a quantitative research with an adequate sample size (n=304). The item descriptive statistics were in the expected ranges; all items significantly positively correlated with the total score; except for two items, all items significantly positively correlated with the two criteria scores (Team Member Experience and Team Performance scores), and there was no negative correlation with the criteria scores. Inter-rater Reliability was calculated in 40 teams and produced an acceptable result.

The 60 items of TBFQHR are grouped into 10 themes based on content similarity. The 10 themes showed acceptable Cronbach's Alpha values. All themes positively correlate with each other (Appendix Figure A9). This result is consistent with the other instruments reviewed (TEQ, TCI, FBCT, VTS) which also all showed positive correlations between their scales.

Limitations and further work

Even though many organizations and teams participated in the quantitative research, two organizations account for 62% of the sample, which may limit the generalizability of the results somewhat.

The reliability of a questionnaire is best evaluated by a test-retest study, this was not included in this study and should be conducted.

The strongest validity evidence of formative questionnaires is correlation with external data independent from the questionnaire. The criteria data included in this study does not meet this requirement satisfactorily, and a study with external data independent from participants' perceptions would be needed.

The factor analysis included in this study was not conclusive, especially Factor 1 which explains 40% of the variance would need further research to determine whether it is a kind of bias or a dominant factor of team effectiveness.

Concluding thoughts

This paper takes the reader on a journey to view teamwork effectiveness at multiple levels of analysis: Qualitative interviews such as Hosseini et al. (2018) that are closest to the rich tapestry of life; some themes emerge from the coding of these interviews; then researchers empirically validate individual facets of these themes and identify the individual factors that influence team performance; then researchers such as Salas et al. (2005) or Lurey and Raisinghani (2001) attempt to combine these factors and create comprehensive models with hypothesized causal relationships among the factors; and studies such as Lin, Standing, and Liu (2008) attempt to test these relationships. TBFQHR fits into this picture by attempting to take the learning gained from academic research back to where it all began: to the leaders and members of teams.

Appendix

Figure A1. The first draft of TBFQHR prior to the SME feedback (with 42 items in 11 themes):

#	Themes	Items
1	Team Leadership	The goals of the team are clearly defined.
2		The roles and responsibilities are clear in the team.
3		Team members know exactly what work and performance are expected of them.
4		The distribution of tasks in the team is reasonable and fair.
5		Recognition and praise are distributed fairly to team members.
6	Attention to Results	Team members pay attention to the overall progress of the team's work.
7		Team members look out for each other's work.
8		I get the necessary critical feedback on my work from my teammates.
9	Mutual Support	I get a lot of encouragement and recognition from my teammates.
10		Team members give each other suggestions and ideas.
11		Team members help overburdened teammates to complete tasks.
12	Adaptability	The team adapts quickly to changes (e.g. changes in the needs of the team's customers).
13		The team regularly reflects on its performance to learn from the experience.
14	Team Orientation	Team members actively participate in discussions about the functioning of the team.
15		Team members place team goals above individual goals.
16		Team members adjust their work based on input and feedback from other team members.
17		Conflicts within the team are handled constructively through open, direct communication.
18		The team is "one team" - no cliques, no fault lines, no competing sub-groups.
19	Shared Team Know-How	Team members understand the interdependencies between their roles (how their work is tied to one another).
20		Team members know how to coordinate their work with each other.
21		Team members know each other's skills, experience, strengths/weaknesses, and working styles.
22		The team has a supportive work culture, which greatly influences the behavior of team members.
23	Mutual Trust	Team members have strong interpersonal relationships and know about each other's personal situations.
24		Team members feel safe to openly criticize ideas and plans.
25		Team members feel safe to bring up problems and tough issues.
26		Team members are willing to admit mistakes and accept feedback.
27	Effective Communication	There is sufficient communication in the team.
28		It is easy to reach out and talk to (or schedule talks with) other team members.
29		Team members are well aware of the communication channels and customs within the team.
30		It is standard practice in the team to confirm that messages have been received and understood accurately.
31		The team uses a common vocabulary that all team members understand in the same way.
32	Organizational Support	I feel well informed about the "big picture" (what is going on in the organization, among the internal/external customers,...)
33		I feel that the organization recognizes the importance of the team's work.
34		I have sufficient skill development opportunities at work.
35		I have access to the information I need to do my job.
36		I have the authorization/empowerment I need to do my job.
37	IT Tools for Collaboration	The team has appropriate IT (information technology) tools for collaboration.
38		Team members use well the IT (information technology) tools for collaboration.
39	Individual Circumstances	My current physical remote or teleworking environment (home office) enables me to work effectively.
40		I can stay focused on work without too many distractions in the current (hybrid or remote) working arrangement.
41		I can manage work-related stress in the current (hybrid or remote) working arrangement.

#	Themes	Items
42		I can manage the balance between work and private life in the current (hybrid or remote) working arrangement.

Figure A2. Item means and standard deviations (n=304)

Item Descriptive Statistics			
	Mean	Std. Deviation	N
Item1	3.5296	1.10763	304
Item2	3.3586	1.13440	304
Item3	3.4276	1.09365	304
Item4	3.4309	1.09687	304
Item5	3.6447	1.07420	304
Item6	3.7171	1.12527	304
Item7	3.4539	1.06451	304
Item8	3.4770	1.10480	304
Item9	3.6349	1.06624	304
Item10	3.8454	1.05263	304
Item11	3.6349	.96210	304
Item12	3.5888	.97426	304
Item13	4.0757	.98044	304
Item14	3.7664	1.09366	304
Item15	3.9638	.93449	304
Item16	4.2303	.85187	304
Item17	3.8717	1.01798	304
Item18	3.6842	.97101	304
Item19	3.6447	1.12228	304
Item20	3.3947	1.12375	304
Item21	3.6875	1.02657	304
Item22	3.4671	1.01746	304
Item23	3.6612	.99521	304
Item24	3.4572	.99411	304
Item25	3.5592	.94211	304
Item26	3.6711	.97680	304
Item27	3.6645	.98130	304
Item28	3.4507	1.03926	304
Item29	3.5197	1.08071	304
Item30	3.5625	1.02254	304
Item31	3.8618	.99867	304
Item32	3.9211	1.08410	304
Item33	3.7763	1.13566	304
Item34	3.5362	1.04297	304
Item35	3.7928	1.00156	304
Item36	3.6974	.91946	304
Item37	4.0066	.82050	304
Item38	3.8421	.86062	304
Item39	3.8750	.77342	304
Item40	3.5724	1.05993	304
Item41	3.9079	.96715	304
Item42	3.9803	.98148	304
Item43	4.0921	.92886	304
Item44	3.6645	1.05895	304
Item45	3.5987	.98006	304
Item46	3.9704	.84575	304
Item47	3.6217	1.06473	304
Item48	4.1316	.85738	304
Item49	3.9836	.96287	304
Item50	3.7599	1.11318	304
Item51	4.2928	.85393	304
Item52	4.2697	.88232	304
Item53	3.7928	1.02436	304
Item54	4.0757	.89969	304
Item55	4.0493	.82105	304
Item56	3.6711	1.00017	304
Item57	3.9539	.78213	304
Item58	4.1349	.79907	304
Item59	3.6974	1.01174	304
Item60	4.2039	.81923	304

Figure A3. Item – corrected total score correlations (n=304)

Item-Total Score Correlations			
	Item	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Item1	Having clearly defined team goals	.695	.974
Item2	Having motivating team goals	.728	.974
Item3	Having agreement and alignment of all team members with the team's goals	.707	.974
Item4	Having a shared understanding of how to achieve the team's goals	.706	.974
Item5	Knowing each other's skills, experience, strengths, and limitations	.607	.974
Item6	Having clearly defined roles and responsibilities in the team	.631	.974
Item7	Having a reasonable distribution of tasks	.631	.974
Item8	Having a clear work plan (with action steps, milestones, sprints,...)	.645	.974
Item9	Proactively informing each other about accomplishments, difficulties, and delays we encounter	.645	.974
Item10	Being able to quickly bring together ad hoc problem-solving meetings with the relevant people	.596	.974
Item11	Placing team goals above our individual interests	.673	.974
Item12	Adapting our working styles to each other	.619	.974
Item13	Creating a friendly and collaborative team atmosphere	.656	.974
Item14	Including all team members in discussions about the team's functioning	.674	.974
Item15	Giving each other suggestions and guidance	.681	.974
Item16	Helping each other out when someone needs support	.607	.974
Item17	Giving each other encouragement and recognition	.646	.974
Item18	Resolving disagreements and conflicts constructively	.686	.974
Item19	Keeping the team together as a cohesive "One Team" (no cliques, no fault lines, no internal fighting)	.730	.974
Item20	Conducting effective team-building activities to keep team spirit high	.656	.974
Item21	Having all team members complete their part of the work on time and to a high standard	.700	.974
Item22	Reviewing the team's results, achievements, and performance gaps regularly	.665	.974
Item23	Actively participating in discussions to improve team performance	.718	.974
Item24	Actively seeking feedback from each other and outside of the team to continuously improve	.641	.974
Item25	Challenging each other in a positive way by providing constructive, valuable feedback to each other	.715	.974
Item26	Adapting quickly to changing needs or environments to ensure the long-term success of the team	.719	.974
Item27	Addressing problems that arise in a timely manner	.664	.974
Item28	Developing new approaches, experimenting with new methods, and continuously improving the way we work	.628	.974
Item29	Keeping up to date with the latest knowledge and trends	.621	.974
Item30	Regularly reflecting on what has happened in the team and learning from the experiences	.692	.974
Item31	Listening to each other	.695	.974
Item32	Creating an inclusive team atmosphere in which every member feels accepted and valued, regardless of age, gender, or ethnicity	.719	.974
Item33	Creating an atmosphere of psychological safety in which team members can speak their minds, admit mistakes, and ask for help	.753	.974
Item34	Having in-depth discussions and debates	.693	.974
Item35	Having sufficient informal communication in the team	.618	.974
Item36	Developing and consistently using communication standards (which channel to use in different situations, how quickly to respond, etc.)	.618	.974
Item37	Using a common vocabulary that all team members understand in the same way	.602	.974
Item38	Making good use of asynchronous communication (email and messaging) to allow for fewer interruptions and more focused work	.513	.974
Item39	Making sure that messages get received and understood accurately	.655	.974
Item40	Avoiding wasted time caused by unnecessary or badly prepared meetings	.662	.974
Item41	Conducting effective meetings regardless of where team members are located	.723	.974
Item42	Using the video conferencing software effectively	.633	.974
Item43	I think the team has the manpower and the skills to be successful	.510	.974
Item44	I think the team has the necessary organizational resources (budget, time, autonomy, tools, and access to information)	.516	.974
Item45	I feel well-informed about the broader context of our work (what is going on in the organization, among the internal/external stakeholders)	.624	.974
Item46	I have access to the information I need to do my job	.559	.974
Item47	I am satisfied with the IT (information technology) tools we use in the team	.460	.974
Item48	I have no problem getting used to all the different IT systems we need to use in the team	.514	.974
Item49	I am satisfied with the degree of autonomy I have over where and when I do my work	.491	.974
Item50	I am satisfied with the organization's policies related to hybrid/remote working	.429	.974
Item51	My home office (or other teleworking) workspace is properly equipped and well-connected to the Internet	.296	.975
Item52	I can stay focused on work without too many distractions in my home office (or other teleworking) environment	.315	.975
Item53	I can cope with the current workload without the danger of burnout	.415	.974
Item54	I can manage the balance between work and private life in the current (hybrid/remote) work model	.459	.974
Item55	Currently, I have no problem with the potential stress associated with hybrid/remote work (e.g., less informal interaction)	.405	.974
Item56	I am satisfied with the learning opportunities provided by the team and the organisation	.610	.974
Item57	My work in the team is well aligned with both my strengths and interests	.662	.974
Item58	I feel that I am making an important contribution to the team with my work	.530	.974
Item59	I feel that I get fair recognition for a job well done	.542	.974
Item60	I really like my work in this team	.625	.974

Figure A5. Item correlations with the Team Member Experience score: (n=304)

Nr.	Items	Team member experience
60	I really like my work in this team	0.63
19	Keeping the team together as a cohesive "One Team" (no cliques, no fault lines, no internal fighting)	0.59
57	My work in the team is well aligned with both my strengths and interests	0.58
13	Creating a friendly and collaborative team atmosphere	0.57
33	Creating an atmosphere of psychological safety in which team members can speak their minds, admit mistakes, talk about their emotions, and share divergent ideas	0.56
2	Having motivating team goals	0.54
31	Listening to each other	0.54
32	Creating an inclusive team atmosphere in which every member feels accepted and valued, regardless of age, gender, cultural, regional, linguistic, or other personal differences	0.54
14	Including all team members in discussions about the team's functioning	0.53
15	Giving each other suggestions and guidance	0.53
11	Placing team goals above our individual interests	0.52
26	Adapting quickly to changing needs or environments to ensure the long-term success of the team	0.52
1	Having clearly defined team goals	0.51
4	Having a shared understanding of how to achieve the team's goals	0.51
43	I think the team has the manpower and the skills to be successful	0.51
3	Having agreement and alignment of all team members with the team's goals	0.49
25	Challenging each other in a positive way by providing constructive, valuable feedback to each other	0.49
34	Having in-depth discussions and debates	0.49
17	Giving each other encouragement and recognition	0.48
23	Actively participating in discussions to improve team performance	0.48
56	I am satisfied with the learning opportunities provided by the team and the organisation	0.48
18	Resolving disagreements and conflicts constructively	0.47
20	Conducting effective team-building activities to keep team spirit high	0.47
45	I feel well-informed about the broader context of our work (what is going on in the organization, among the internal/external customers,...)	0.47
12	Adapting our working styles to each other	0.46
16	Helping each other out when someone needs support	0.46
21	Having all team members complete their part of the work on time and to a high standard	0.46
35	Having sufficient informal communication in the team	0.46
41	Conducting effective meetings regardless of where team members are located	0.46
7	Having a reasonable distribution of tasks	0.44
58	I feel that I am making an important contribution to the team with my work	0.44
24	Actively seeking feedback from each other and outside of the team to continuously improve	0.43
59	I feel that I get fair recognition for a job well done	0.43
5	Knowing each other's skills, experience, strengths, and limitations	0.42
6	Having clearly defined roles and responsibilities in the team	0.42
30	Regularly reflecting on what has happened in the team and learning from the experiences	0.42
40	Avoiding wasted time caused by unnecessary or badly prepared meetings	0.42
29	Keeping up to date with the latest knowledge and trends	0.41
42	Using the video conferencing software effectively	0.41
8	Having a clear work plan (with action steps, milestones, sprints,...)	0.4
9	Proactively informing each other about accomplishments, difficulties, and delays we encounter	0.4
22	Reviewing the team's results, achievements, and performance gaps regularly	0.4
36	Developing and consistently using communication standards (which channel to use in different situations, how quickly to respond to different messages, etc.)	0.39
37	Using a common vocabulary that all team members understand in the same way	0.39
44	I think the team has the necessary organizational resources (budget, time, autonomy, tools, and access to information) to be successful	0.38
46	I have access to the information I need to do my job	0.38
39	Making sure that messages get received and understood accurately	0.37
27	Addressing problems that arise in a timely manner	0.36
49	I am satisfied with the degree of autonomy I have over where and when I do my work	0.36
10	Being able to quickly bring together ad hoc problem-solving meetings with the relevant people	0.35
28	Developing new approaches, experimenting with new methods, and continuously improving the way we work	0.34
53	I can cope with the current workload without the danger of burnout	0.34
48	I have no problem getting used to all the different IT systems we need to use in the team	0.32
54	I can manage the balance between work and private life in the current (hybrid/remote) work model	0.32
55	Currently, I have no problem with the potential stress associated with hybrid/remote work (e.g., less informal interaction, more strenuous communication, screen fatigue)	0.31
38	Making good use of asynchronous communication (email and messaging) to allow for fewer interruptions and more focused work	0.28
47	I am satisfied with the IT (information technology) tools we use in the team	0.27
50	I am satisfied with the organization's policies related to hybrid/remote working	0.26
51	My home office (or other teleworking) workspace is properly equipped and well-connected to the Internet	0.19
52	I can stay focused on work without too many distractions in my home office (or other teleworking) environment	0.18

Figure A6. Item correlations with the Team Performance scale: (n=304)

	Items	Team performance
Item21	Having all team members complete their part of the work on time and to a high standard	0.58
Item26	Adapting quickly to changing needs or environments to ensure the long-term success of the team	0.55
Item32	Creating an inclusive team atmosphere in which every member feels accepted and valued, regardless of age, gender, cultural, regional, linguistic, or other personal differences	0.51
Item34	Having in-depth discussions and debates	0.51
Item2	Having motivating team goals	0.5
Item19	Keeping the team together as a cohesive "One Team" (no cliques, no fault lines, no internal fighting)	0.5
Item33	Creating an atmosphere of psychological safety in which team members can speak their minds, admit mistakes, talk about their emotions, and share divergent ideas	0.5
Item3	Having agreement and alignment of all team members with the team's goals	0.49
Item29	Keeping up to date with the latest knowledge and trends	0.49
Item4	Having a shared understanding of how to achieve the team's goals	0.48
Item11	Placing team goals above our individual interests	0.48
Item24	Actively seeking feedback from each other and outside of the team to continuously improve	0.48
Item12	Adapting our working styles to each other	0.47
Item18	Resolving disagreements and conflicts constructively	0.47
Item27	Addressing problems that arise in a timely manner	0.47
Item40	Avoiding wasted time caused by unnecessary or badly prepared meetings	0.47
Item1	Having clearly defined team goals	0.46
Item6	Having clearly defined roles and responsibilities in the team	0.46
Item8	Having a clear work plan (with action steps, milestones, sprints,--)	0.46
Item15	Giving each other suggestions and guidance	0.46
Item25	Challenging each other in a positive way by providing constructive, valuable feedback to each other	0.46
Item30	Regularly reflecting on what has happened in the team and learning from the experiences	0.46
Item41	Conducting effective meetings regardless of where team members are located	0.46
Item14	Including all team members in discussions about the team's functioning	0.45
Item23	Actively participating in discussions to improve team performance	0.45
Item28	Developing new approaches, experimenting with new methods, and continuously improving the way we work	0.45
Item39	Making sure that messages get received and understood accurately	0.45
Item7	Having a reasonable distribution of tasks	0.44
Item57	My work in the team is well aligned with both my strengths and interests	0.44
Item9	Proactively informing each other about accomplishments, difficulties, and delays we encounter	0.43
Item22	Reviewing the team's results, achievements, and performance gaps regularly	0.43
Item20	Conducting effective team-building activities to keep team spirit high	0.42
Item42	Using the video conferencing software effectively	0.42
Item43	I think the team has the manpower and the skills to be successful	0.42
Item10	Being able to quickly bring together ad hoc problem-solving meetings with the relevant people	0.4
Item13	Creating a friendly and collaborative team atmosphere	0.4
Item16	Helping each other out when someone needs support	0.4
Item31	Listening to each other	0.4
Item36	Developing and consistently using communication standards (which channel to use in different situations, how quickly to respond to different messages, etc.)	0.4
Item38	Making good use of asynchronous communication (email and messaging) to allow for fewer interruptions and more focused work	0.4
Item44	I think the team has the necessary organizational resources (budget, time, autonomy, tools, and access to information) to be successful	0.4
Item17	Giving each other encouragement and recognition	0.39
Item35	Having sufficient informal communication in the team	0.39
Item45	I feel well-informed about the broader context of our work (what is going on in the organization, among the internal/external customers,...)	0.39
Item56	I am satisfied with the learning opportunities provided by the team and the organisation	0.39
Item46	I have access to the information I need to do my job	0.37
Item5	Knowing each other's skills, experience, strengths, and limitations	0.35
Item54	I can manage the balance between work and private life in the current (hybrid/remote) work model	0.35
Item59	I feel that I get fair recognition for a job well done	0.35
Item53	I can cope with the current workload without the danger of burnout	0.33
Item37	Using a common vocabulary that all team members understand in the same way	0.31
Item58	I feel that I am making an important contribution to the team with my work	0.31
Item47	I am satisfied with the IT (information technology) tools we use in the team	0.28
Item49	I am satisfied with the degree of autonomy I have over where and when I do my work	0.26
Item55	Currently, I have no problem with the potential stress associated with hybrid/remote work (e.g., less informal interaction, more strenuous communication, screen fatigue)	0.26
Item48	I have no problem getting used to all the different IT systems we need to use in the team	0.24
Item50	I am satisfied with the organization's policies related to hybrid/remote working	0.2
Item52	I can stay focused on work without too many distractions in my home office (or other teleworking) environment	0.13
Item51	My home office (or other teleworking) workspace is properly equipped and well-connected to the Internet	0.07

Figure A7. Component matrix (n=304)

	Component									
	1	2	3	4	5	6	7	8	9	10
Item1	.707	.059	-.288	-.273	.214	-.040	-.089	-.029	-.018	-.191
Item2	.740	.059	-.251	-.189	.127	-.036	-.023	-.077	-.039	-.051
Item3	.719	-.004	-.344	-.188	.100	-.167	.174	-.162	.017	.014
Item4	.720	-.031	-.332	-.176	.100	-.202	.117	-.165	-.097	-.094
Item5	.623	-.037	-.311	-.039	.190	-.246	.072	.178	.153	-.040
Item6	.648	-.087	-.270	-.219	.123	-.153	.049	.135	.166	-.127
Item7	.644	.035	-.218	-.219	.045	.101	-.034	-.134	-.151	-.015
Item8	.661	.040	-.206	-.259	-.030	.223	-.009	.021	-.215	-.001
Item9	.667	-.232	-.092	-.092	.089	.166	.103	.181	-.037	.042
Item10	.613	-.013	-.076	-.232	-.011	.137	.223	.070	.055	.004
Item11	.695	-.225	-.155	-.105	-.053	.016	.111	.159	-.014	.013
Item12	.641	-.260	-.138	-.035	-.015	-.060	.200	.082	.222	.154
Item13	.677	-.263	-.055	.226	.216	-.111	-.088	-.058	-.085	-.024
Item14	.698	-.272	-.220	.115	-.019	.087	-.090	.135	-.136	.012
Item15	.703	-.317	-.095	.157	.195	.029	-.036	.170	.061	-.031
Item16	.630	-.328	.006	.251	.276	.065	.044	.113	.059	.030
Item17	.670	-.333	-.003	.342	.139	-.041	-.002	.082	.029	.141
Item18	.709	-.327	-.097	.226	-.023	-.085	.045	.043	-.030	.161
Item19	.750	-.233	-.090	.176	.044	-.021	-.124	-.105	-.110	-.012
Item20	.670	.050	-.184	.081	-.183	.214	-.060	.034	-.196	.049
Item21	.719	-.106	.093	-.143	.022	.160	.045	-.288	.000	-.050
Item22	.683	-.016	-.072	-.224	-.200	.339	.024	.003	-.078	.102
Item23	.737	-.082	-.108	-.092	-.132	.188	-.066	-.070	-.180	.035
Item24	.665	-.157	-.099	-.068	-.153	.066	-.045	-.202	-.056	.171
Item25	.735	-.168	.023	-.032	.013	.125	-.045	-.255	.084	.032
Item26	.738	-.065	.162	-.010	-.096	-.002	-.074	-.172	.358	-.027
Item27	.684	.004	.081	-.171	-.261	.056	.140	-.098	.255	.002
Item28	.650	-.033	.195	-.040	-.294	.078	.072	-.057	.304	.192
Item29	.642	-.036	.230	-.039	-.265	.104	-.065	.018	.186	.036
Item30	.715	-.070	.164	-.041	-.200	.046	-.058	-.136	.060	.072
Item31	.717	-.223	.175	.232	.098	-.132	-.026	.123	-.161	-.037
Item32	.740	-.147	.255	.164	.043	-.136	-.060	-.084	.013	-.035
Item33	.775	-.186	.200	.172	.022	-.114	-.117	-.082	-.042	-.026
Item34	.719	-.187	.233	.086	-.158	.007	-.115	-.082	-.109	.042
Item35	.640	-.069	.222	-.047	.033	-.020	-.046	.212	.093	.022
Item36	.639	.006	.247	-.215	-.123	-.165	-.069	.307	.040	.007
Item37	.617	.170	.294	-.064	.124	-.223	.053	.102	-.244	.060
Item38	.536	.042	.421	-.294	-.032	-.266	.109	.084	-.261	.085
Item39	.671	.107	.382	-.110	-.036	-.112	.070	.062	-.190	.029
Item40	.685	.013	.367	-.196	-.108	-.095	-.043	-.071	-.063	-.047
Item41	.737	.095	.321	-.117	-.002	-.012	.089	-.057	-.069	-.201
Item42	.644	.118	.214	.068	.119	.008	.241	-.153	.129	-.247
Item43	.517	.181	-.097	.006	.221	.038	-.421	-.136	.197	-.309
Item44	.525	.392	-.126	-.057	-.196	.020	-.314	.150	.097	-.090
Item45	.637	.169	-.029	.054	-.110	-.035	-.247	.345	.187	.016
Item46	.565	.304	-.096	-.099	-.097	.064	-.206	.450	.023	-.003
Item47	.467	.419	-.070	.103	-.244	.076	.250	.146	-.205	-.178
Item48	.525	.160	.228	.071	.080	.029	.117	.065	.084	-.461
Item49	.492	.428	-.178	.288	-.160	-.170	.292	.028	.056	.091
Item50	.435	.355	-.143	.379	-.088	-.127	.454	-.008	.091	.069
Item51	.296	.312	.180	.056	.417	.492	.138	.194	-.002	.144
Item52	.321	.207	.256	.059	.567	.455	.089	-.020	.072	.027
Item53	.419	.436	-.092	-.260	.233	-.245	-.242	-.076	.043	.243
Item54	.457	.509	.038	-.016	.158	-.088	.050	-.176	.091	.347
Item55	.409	.363	.093	.022	.466	-.121	-.191	-.077	-.029	.262
Item56	.617	.344	-.190	.171	-.197	.035	-.127	-.129	-.085	-.002
Item57	.666	.275	-.080	.329	-.046	-.060	-.139	-.147	-.063	-.116
Item58	.537	.250	.059	.408	-.068	.053	.018	-.057	-.121	-.121
Item59	.549	.225	-.189	.379	-.238	.099	-.120	-.017	-.009	.081

Extraction Method: Principal Component Analysis.

a. 10 components extracted.

Figure A8. Inter-rater Reliability

Intraclass Correlation Coefficient (ICC) in each team with two or more participants. The calculation used two-way mixed modeling and reported at 95% confidence interval.

Team	Nr of team members	ICC
A1	5	0.83
B1	7	0.72
C1	5	0.58
D1	9	0.84
E1	10	0.75
F1	5	0.78
G1	4	0.83
H1	8	0.7
H2	5	0.62
I1	6	0.64
J1	5	0.74
K1	7	0.66
L1	8	0.79
M1	5	0.73
N1	5	0.69
N2	10	0.77
N4	3	0.52
N6	10	0.86
O1	8	0.78
O2	10	0.87
O3	9	0.8
P01	11	0.8
P02	3	0.7
P03	5	0.76
P04	3	0.78
P07	4	0.62
P08	6	0.7
P09	6	0.61
P11	5	0.77
P12	3	0.65
P13	2	0.35
P14	3	0.68
P15	3	0.63
P16	3	0.76
P17	3	0.66
P18	7	0.7
P19	11	0.89
P20	2	0.61
P22	2	0.47
P23	14	0.86
Mean		0.7125

Figure A9. Correlations between themes

Correlations											
		Sharedg	Coordin	Teamspir	Attentres	Adapt	Psysafety	Effcomm	Orgsupp	Coping	Motiv
Sharedg	Pearson Correlation	1	.796**	.683**	.680**	.589**	.585**	.603**	.584**	.465**	.572**
	Sig. (2-tailed)		.000	.000	.000	.000	.000	.000	.000	.000	.000
	N	304	304	304	304	304	304	304	304	304	304
Coordin	Pearson Correlation	.796**	1	.785**	.763**	.666**	.650**	.658**	.641**	.447**	.557**
	Sig. (2-tailed)	.000		.000	.000	.000	.000	.000	.000	.000	.000
	N	304	304	304	304	304	304	304	304	304	304
Teamspir	Pearson Correlation	.683**	.785**	1	.768**	.697**	.801**	.636**	.592**	.349**	.627**
	Sig. (2-tailed)	.000	.000		.000	.000	.000	.000	.000	.000	.000
	N	304	304	304	304	304	304	304	304	304	304
Attentres	Pearson Correlation	.680**	.763**	.768**	1	.759**	.726**	.683**	.587**	.406**	.595**
	Sig. (2-tailed)	.000	.000	.000		.000	.000	.000	.000	.000	.000
	N	304	304	304	304	304	304	304	304	304	304
Adapt	Pearson Correlation	.589**	.666**	.697**	.759**	1	.731**	.742**	.622**	.369**	.582**
	Sig. (2-tailed)	.000	.000	.000	.000		.000	.000	.000	.000	.000
	N	304	304	304	304	304	304	304	304	304	304
Psysafety	Pearson Correlation	.585**	.650**	.801**	.726**	.731**	1	.772**	.576**	.398**	.612**
	Sig. (2-tailed)	.000	.000	.000	.000	.000		.000	.000	.000	.000
	N	304	304	304	304	304	304	304	304	304	304
Effcomm	Pearson Correlation	.603**	.658**	.636**	.683**	.742**	.772**	1	.635**	.506**	.557**
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000		.000	.000	.000
	N	304	304	304	304	304	304	304	304	304	304
Orgsupp	Pearson Correlation	.584**	.641**	.592**	.587**	.622**	.576**	.635**	1	.532**	.746**
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.000		.000
	N	304	304	304	304	304	304	304	304	304	304
Coping	Pearson Correlation	.465**	.447**	.349**	.406**	.369**	.398**	.506**	.532**	1	.459**
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.000	.000	
	N	304	304	304	304	304	304	304	304	304	304
Motiv	Pearson Correlation	.572**	.557**	.627**	.595**	.582**	.612**	.557**	.746**	.459**	1
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.000	.000	
	N	304	304	304	304	304	304	304	304	304	304

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

References

- Alsharo, M., Gregg, D. & Ramirez, R. (2017) Virtual team effectiveness: The role of knowledge sharing and trust, *Information & management*, vol. 54, no. 4, pp. 479-490.
- Anderson, N.R. & West, M.A. (1998) Measuring climate for work group innovation: development and validation of the team climate inventory, *Journal of organizational behavior*, vol. 19, no. 3, pp. 235-258.
- Beckmann, M.J. (1995) Teamwork versus delegation — comparison of Japanese and German management styles, *Japan and the world economy*, vol. 7, no. 2, pp. 139-144.
- Belbin, R.M. (1981) *Management teams: why they succeed or fail*, Butterworth-Heinemann, Oxford.
- Chyung, S.Y., Kennedy, M. & Campbell, I. (2018) Evidence-Based Survey Design: The Use of Ascending or Descending Order of Likert-Type Response Options, *Performance improvement* (International Society for Performance Improvement), vol. 57, no. 9, pp. 9-16.
- Chyung, S.Y., Roberts, K., Swanson, I. & Hankinson, A. (2017) Evidence-Based Survey Design: The Use of a Midpoint on the Likert Scale, *Performance improvement* (International Society for Performance Improvement), vol. 56, no. 10, pp. 15-23.
- Coltman, T., Devinney, T.M., Midgley, D.F. & Venaik, S. (2008) Formative versus reflective measurement models: Two applications of formative measurement, *Journal of business research*, vol. 61, no. 12, pp. 1250-1262.
- de Winter, J. C. F., Dodou, D. & Wieringa, P.A. (2009) Exploratory Factor Analysis With Small Sample Sizes, *Multivariate behavioral research*, vol. 44, no. 2, pp. 147-181.
- Delise, L.A., Allen-Gorman, C., Brooks, A.M., Rentsch, J.R. & Steele-Johnson, D. (2010) The effects of team training on team outcomes: A meta-analysis, *Performance improvement quarterly*, vol. 22, no. 4, pp. 53-80.

- Diamantopoulos, A., Sarstedt, M., Fuchs, C. (2012) Guidelines for choosing between multi-item and single-item scales for construct measurement: a predictive validity perspective. *J. of the Acad. Mark. Sci.* 40, 434–449.
- Dykes, B. (2020) *Effective data storytelling: how to drive change with data, narrative and visuals*, Wiley, Hoboken, New Jersey.
- Fortune Business Insights (2021) Team collaboration software market report.
<https://www.fortunebusinessinsights.com/industry-reports/team-collaboration-software-market-101327> (Accessed: 12 December 2022)
- Gartner (2021) <https://www.gartner.com/en/newsroom/press-releases/2021-08-23-gartner-survey-reveals-44-percent-rise-in-workers-use-of-collaboration-tools-since-2019> (Accessed: 12 December 2022)
- Gartner (2022) *Evolve culture and leadership for the hybrid workplace* [eBook],
<https://emtemp.gcom.cloud/ngw/globalassets/en/human-resources/documents/trends/evolve-culture-and-leadership-for-the-hybrid-workplace-ebook.pdf>, (Accessed: 12 December 2022)
- Grossman, R. and Miller, J.P. (2021) Optimizing Team Effectiveness: Key Takeaways From the Science With a Focus on the Virtual Context, *American journal of health promotion*, vol. 35, no. 5, pp. 732-737.
- Haynes, S.N., Richard, D.C.S. & Kubany, E.S. (1995) Content Validity in Psychological Assessment: A Functional Approach to Concepts and Methods, *Psychological assessment*, vol. 7, no. 3, pp. 238-247.
- Hosseini, M.R., Bosch-Sijtsema, P., Arashpour, M., Chileshe, N. & Merschbrock, C. (2018) A qualitative investigation of perceived impacts of virtuality on effectiveness of hybrid construction project teams, *Construction innovation*, vol. 18, no. 1, pp. 109-131.
- Kahlow, J., Klecka, H. & Ruppel, E. (2020) What the differences in conflict between online and face-to-face work groups mean for hybrid groups: A state-of-the-art review, *Review of Communication Research*, vol. 8, pp. 51-77.
- Kivimäki, M., Kuk, G., Elovainio, M., Thomson, L., Kalliomäki-Levanto, T. & Heikkilä, A. (1997) The Team Climate Inventory (TCI)-four or five factors? Testing the structure of TCI in samples of low

and high complexity jobs, *Journal of occupational and organizational psychology*, vol. 70, no. 4, pp. 375-389.

Koo, T. K., & Li, M. Y. (2016) A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of chiropractic medicine*, 15(2), 155–163.

Lin, C., Standing, C. & Liu, Y. (2008) A model to develop effective virtual teams, *Decision Support Systems*, vol. 45, no. 4, pp. 1031-1045.

Lurey, J.S. & Raisinghani, M.S. (2001) An empirical study of best practices in virtual teams, *Information & management*, vol. 38, no. 8, pp. 523-544.

Martins, L.L., Gilson, L.L. & Maynard, M.T. (2004) Virtual Teams: What Do We Know and Where Do We Go From Here?, *Journal of management*, vol. 30, no. 6, pp. 805-835.

Morrison-Smith, S. & Ruiz, J. (2020) Challenges and barriers in virtual teams: a literature review, *SN Applied Sciences*, vol. 2, no. 6.

Mortinsen, M. & Haas, M. (2021) Making the hybrid workplace fair, *Harvard Business Review*, <https://hbr.org/2021/02/making-the-hybrid-workplace-fair> (Accessed: 15 October 2022)

Nelson, Bob (2022) Managing Virtual Workers in the Era of Hybrid Work, <https://www.trainingmagnetnetwork.com/events/3168> (Accessed: 15 October 2022)

Oosterveld, P., Vorst, H.C.M. & Smits, N. (2019) Methods for questionnaire design: a taxonomy linking procedures to test goals, *Quality of life research*, vol. 28, no. 9, pp. 2501-2512.

Reyes, D.L., Luna, M. & Salas, E. (2021) Challenges for team leaders transitioning from face-to-face to virtual teams, *Organizational dynamics*, vol. 50, no. 2, pp. 100785.

Salas, E., Stagl, K. C., Burke, C. S., and Goodwin, G. F. (2007) Fostering team effectiveness in organizations: Toward an integrative theoretical framework of team performance, *Modeling complex systems: Motivation, cognition and social processes: Nebraska Symposium on Motivation* (Vol. 51, pp. 185–243). Lincoln: University of Nebraska Press.

Salas, E., Sims, D.E. and Burke, C.S., (2005) Is there a “big five” in teamwork?, *Small group research*, 36(5), pp.555-599.

Salas, E., Cooke, N. J., & Rosen, M. A. (2008) On Teams, Teamwork, and Team Performance: Discoveries and Developments, *Human Factors*, 50(3), 540–547.

Schroeder, K.A., Sorensen, P.F. & Yaeger, T.F. (2014) Accelerating Global Hybrid Team Effectiveness, *Research in Organizational Change and Development*, Emerald Group Publishing Limited, pp. 335-364.

Spoto, A., Nucci, M., Prunetti, E. & Vicovaro, M. (2023) Improving content validity evaluation of assessment instruments through formal content validity analysis, *Psychological methods*, 01/2023.

Sundstrom, E., De Meuse, K.P. & Futrell, D. (1990) Work Teams: Applications and Effectiveness, *The American psychologist*, vol. 45, no. 2, pp. 120-133.

Zafft, C., Perez, M.A. & Adams, S. (2006) Development Of A Scoring System For The Team Effectiveness Questionnaire (Teq), *American Society for Engineering Education-ASEE*, Atlanta, pp. 11.460.1.

Vagias, Wade M. (2006) Likert-type scale response anchors. Clemson International Institute for Tourism & Research Development, Department of Parks, Recreation and Tourism Management. Clemson University.

Wiley (2014) The Five Behaviors of a Cohesive Team - Research Report, Version 2
<https://www.fivebehaviors.co.uk/FiveBehaviors/media/SiteFiles/assets/Five-Behaviors-Research-Report-v2.pdf> (Accessed: 30 December 2022)

Wiley (2023) The Five Behaviors <https://www.fivebehaviors.com/> (Accessed: 30 December 2022)

Willits, F.K., Theodori, G.L. & Luloff, A. (2016) ANOTHER LOOK AT LIKERT SCALES, *Journal of rural social sciences*, vol. 31, no. 3, pp. 126.