ORIGINAL PAPER

How Virtual are We? Introducing the Team Perceived Virtuality Scale

Lisa Handke[1] · Patrícia Costa[2] · Jennifer Feitosa[3]

Abstract

With the strong proliferation of virtual teams across various organizations and contexts, understanding how virtuality affects teamwork has become fundamental to team and organizational effectiveness. However, current conceptualizations of virtuality rely almost exclusively on more or less fixed, structural features, such as the degree of technology reliance. In this paper, we take a socio-constructivist perspective on team virtuality, focusing on individuals' experience of team virtuality, which may vary across teams and time points with similar structural features. More specifically, we develop and validate a scale that captures the construct of Team Perceived Virtuality (Handke et al., 2021). Following a description of item development and content validity, we present the results of four different studies that demonstrate the construct's structural, discriminant, and criterion validity with an overall number of 2,294 teams. The final instrument comprises 10 items that measure the two dimensions of Team Perceived Virtuality (collectively-experienced distance and collectively-experienced information deficits) with five items each. This final scale showed a very good fit to a two-dimensional structure both at individual and team levels and adequate psychometric properties including aggregation indices. We further provide evidence for conceptual and empirical distinctiveness of the two TPV dimensions based on related team constructs, and for criterion validity, showing the expected significant relationships with leader-rated interaction quality and team performance. Lastly, we generalize results from student project teams to an organizational team sample. Accordingly, this scale can enhance both research and practice as a validated instrument to address how team virtuality is experienced.

Keywords Virtual Teams · Virtuality · Scale Development · Emergent States

Teamwork in modern organizations is unequivocally shaped by the use of virtual tools and by work arrangements that do not require (permanent) physical presence from their employees. This includes employees working from home, teams maximizing expertise by including professionals from different locations, or organizations enabling 24/7 productivity by using different time zones to their advantage (e.g., Dulebohn & Hoch, 2017; Raghuram et al., 2019).

✉ Lisa Handke
lisa.handke@fau.de

1 School of Business, Economics, and Society, Friedrich-Alexander-Universität Erlangen-Nürnberg, Lange Gasse, 20, 90403 Nuremberg, Germany

2 Instituto Universitário de Lisboa (ISCTE-IUL), BRU-IUL, Lisbon, Portugal

3 Claremont McKenna College, Claremont, CA, USA

In the wake of the COVID-19 pandemic, large-scale polls indicate that a majority of employees want to continue to work remotely at least on a part-time basis (e.g., Brenan, 2020; IBM, 2020; Iometrics and Global Workplace Analytics, 2020; Ozimek, 2020), perpetuating virtual and hybrid forms of collaboration as the new normal way of teaming (e.g., Gilson et al., 2021; Gilstrap et al., 2022). Accordingly, understanding the impact of team virtuality (i.e., the extent to which teams engage in technology-mediated communication to accomplish their shared work goals, see e.g., Dixon & Panteli, 2010; Kirkman & Mathieu, 2005) is crucial to promoting team effectiveness now and in the future. Since all teams—despite how much they use technology to communicate—are, to some extent, virtual (Gilson et al., 2021), understanding how team members actually perceive and experience their teamwork under these conditions becomes extremely relevant for scholars and practitioners alike.

However, the mixed findings uncovered in prior research suggest that the way we currently conceptualize and measure team virtuality is insufficient. Specifically, the effects of

team virtuality have been shown to depend on a range of factors that determine how and under which conditions teams work together (e.g., team type, study setting, work design, team duration, Carter et al., 2019; Gibbs et al., 2017; Handke et al., 2020; De Guinea et al., 2012; Purvanova & Kenda, 2022). For instance, studies where team members have enough time to get to know one another, their communication technologies, and the task at hand generally reveal more positive effects on team performance or satisfaction than studies with adhoc laboratory groups, even if team virtuality levels are comparable (e.g., Fuller & Dennis, 2009; van der Kleij et al., 2009). These findings suggest that the type and extent of technology-mediated communication itself is insufficient when it comes to defining team virtuality and explaining its effects on team outcomes.

From a sociomaterial perspective (e.g., Leonardi, 2012; Orlikowski, 1992; Orlikowski & Scott, 2008), the effects of a technology depend less on fixed, structural features (e.g., the technology's capacity to transmit sound in real-time) but on the way it is used in practice (e.g., whether team members use audioconferences for simple information exchanges or in-depth discussions, see also adaptive structuration theory, DeSanctis & Poole, 1994). From this perspective, technology reliance is only problematic if the team's use of this technology causes them to experience impairments in their interactions. The actual problem thus lies in the team's experience of its interaction, which is not a direct result of technology reliance per se but of technology that is poorly used. Moreover, these negative experiences can arise through technology use but could just as well occur in face-to-face interaction, such as when members forget to update each other on what they are currently working on or misinterpret a certain non-verbal cue (see also Watson-Manheim et al., 2002, 2012). Accordingly, to achieve a better understanding of team virtuality, we need to extend (and potentially replace) static, structural indicators (e.g., the overall degree of technology reliance) of virtuality with those that capture how team virtuality is experienced in practice.

Following this rationale, Handke et al. (2021) recently proposed to separate structural indicators of team virtuality (i.e., structural virtuality) from subjective perceptions of team virtuality that reflect the team's experiences. Structural virtuality reflects objective properties of the technology used by teams to communicate and/or objective distance between team members and has been operationalized in numerous ways (see also Table 1). In contrast, Handke et al.'s (2021) conceptual work highlights the importance of considering how team members *experience* virtuality. Specifically, the authors introduced the concept of *Team Perceived Virtuality* (TPV)—a cognitive-affective team emergent state that arises through team interactions. This emergent state is composed of two dimensions: collectively-experienced distance (i.e., team members' collective feelings of being distant from one another) and collectively-experienced information deficits (i.e., team members' collective perceptions of poor information exchange). The two TPV dimensions are proposed to emerge through a sensemaking process triggered by impaired team processes and to subsequently impact team outcomes. Accordingly, whereas prior measures of (structural) virtuality concentrate primarily on technology use and characteristics (e.g., media richness), TPV thus describes collective experiences during teamwork more generally and can thus be applied to any type of team setting.

Accordingly, we have three goals for this article. First, we describe the development of multi-item scales for Handke et al.'s (2021) TPV dimensions and provide evidence of their content validity (Study 1). Second, we test the construct validity of the TPV construct with its two underlying dimensions on both individual and team levels (Studies 2 and 3). Third, we demonstrate the conceptual and empirical distinctiveness of the two TPV dimensions based on related constructs (Studies 4a and 4b). Fourth, we establish criterion validity by linking the TPV subscales to affective (leader-rated team interaction quality) as well as a performance-related (leader-rated team performance) outcomes. Fifth and finally, we generalize our results from individual-level full-time workers and team-level student semester projects to an organizational sample (Study 5). We conclude with recommendations for future research and application of the TPV measure.

Team Perceived Virtuality

Handke et al. (2021) define TPV as a "shared affective-cognitive emergent state that is characterized by team members' co-constructed and collectively-experienced (1) distance and (2) information deficits, thereby capturing the unrealized nature of the team as a collective system" (p. 626). As an emergent state, TPV arises from team interaction processes—which are shaped by teams' technology use (i.e., their structural virtuality) but also by a range of other input factors, such as team familiarity or work design (Handke et al., 2021). Conceptually, TPV is positioned as a mediator in the input-mediator-output-input model of team effectiveness, translating team processes into team outcomes. Accordingly, TPV captures the perception of team virtuality in practice, meaning that TPV can change and evolve depending on how team members experience their interactions with one another and their embedding environment. As most of today's teams will display a certain degree of structural virtuality (for instance, even co-located team members often use emails to share documents, Gilson et al., 2021), the degree of structural virtuality is an important element of teams' work environments. However, the effects of structural virtuality on actual team experiences are not straightforward, regardless of how team virtuality is being operationalized.

Table 1 Overview of Related Constructs and their Difference to TPV

Construct	Definition(s)	Exemplary items	Main conceptual distinction from TPV
Structural Virtuality	"Objective features of (1) communication technology (such as the number and type of social cues that a technology can transport) and/or (2) the distance between team members (such as the number of kilometres or time zones between different sites)" (Handke et al., 2021)	Geographic dispersion/co-location: Virtuality as a dichotomous measure (co-located [0] or distributed teams [1]) (De Guinea et al., 2012) How many times did the team met FtF (Kirkman, et al., 2004) Distance in miles among team members (Hoch & Kozlowski, 2014) Technology use/reliance: "Allocate the percentage of communication that occurs via various computer-medicated communication (CMC) tools" (Rapp et al., 2010) "How often did team members use technology to coordinate the efforts of your team's members?" (Brown et al., 2020)	TPV refers to the collective *experience* of how "virtual" (i.e., distant and deficient in information exchange) the interaction is between team members. Structural virtuality focuses on objective properties of the media used to communicate and/or on the objective distance or interaction frequency of its members
Structural Virtuality	Informational value/media richness: "Extent to which virtual tools send or receive communication or data that are valuable for team effectiveness" (Kirkman & Mathieu, 2005, p. 703) "Channels' relative abilities to convey messages that carry rich information (Carlson & Smud, 1999, p. 154, see also Daft & Lengel, 1986) Synchronicity: Exchanging information (via virtual tools) in real time (Kirkman & Mathieu, 2005) "State in which individuals are working together at the same time with a common focus" (Dennis e al., 2008, p. 581)	"The communication technologies that we used provided enough information to prevent misunderstandings between team members." (Brown et al., 2020) "[Medium] allows my communication partner and me to tailor our messages to our own personal requirements" (Carlson & Zmud, 1999) "My communication with other team members was delayed because of the technology we used" (Brown et al., 2020) "How often did team members use technology to rapidly exchange information between each other?" (Brown et al., 2020)	TPV is independent of technology/communication media usage and, more importantly, of its *features* (e.g., capacity to transmit voice) and/or *character-istics* (e.g., informational value; many of which are socially constructed) Measures of media richness, informational value or synchronicity focus on the technology that is being used and its characteristics and/or properties and not how the exchange is experienced by team members
Affective Nature			
Social Presence	"The degree of salience of the other person in the interaction and the consequent salience of the inter-personal relationships" (Short et al., 1976, p. 65)	unsociable-sociable, insensitive-sensitive, cold-warm, and impersonal-personal (Sallnäs et al., 2000; Wegge et al., 2007)	Both social presence and co-presence imply the aware-ness of another person without necessarily consider-ing the quality of the relationship to this person (or group of persons). The collectively-experienced dis-tance dimension of TPV, in turn, includes affection, friendliness, or emotional availability from others
(Co-)presence	"Sense of being together with others in a mediated – either remote or virtual – environment" (Zhao, 2003, p. 445)	"To what extent were there times, if at all, during which the computer interface seemed to vanish, and you were directly working with the other person?" (Ho et al., 1998)	
(Social) Cohesion/cohesiveness	"Overall attraction or bond amongst members of a group" (Wellen & Neale, 2006, p. 168, see also Festinger, 1950)	"Our team would like to spend time together outside of work hours" (Carless and Paola, 2000)	Social cohesion translates into a desire to spend time with the group. TPV, in turn, refers to the feeling of closeness/distance during the interaction and may, or may not, transfer to wanting to spend more time with the team beyond the necessary job-related interac-tions

Table 1 (continued)

Construct	Definition(s)	Exemplary items	Main conceptual distinction from TPV
Perceived Proximity	"One person's perception of how close or how far another person is" (Wilson et al., 2008, p. 983)	"Even when we are not working in the same place, I still feel close to [distant colleague]" (O'Leary et al., 2014)	Whereas perceived proximity and propinquity reflect (and are measured as) *individual* perceptions, TPV's dimension of collectively experienced distance is conceptualized as a *collective* feeling. Hence, TPV is located at the team level, while perceived proximity and propinquity are located at the individual level. Similarly, both constructs refer to feelings of closeness towards specific individuals, whereas TPV's collectively-experienced distance dimension refers to the distance felt towards the team members as a group
Propinquity	"Subjective perception an individual holds that he or she is functionally, if not physically, close to someone else" (Walther & Bazarova, 2008, see also Korzenny, 1978)	distant-nearby, close-far, together-separate, proximal-remote, and disconnected-connected (Walther & Bazarova, 2008)	
Belonging*	"Frequent and positive interactions and to feel accepted in a stable group" (Baumeister & Leary, 1995, cit. in., Chung et al., 2020, p. 77)	"I am treated as a valuable member of my work group" "I feel that people really care about me in my work-group" (Chung et al., 2020)	Belonging and liking refer to (and are measured as) *individual* perceptions of/attitudes to the individual's relationship with/attitude to the group, whereas TPV is defined as a collective state. More specifically, different team members may perceive different levels of belonging and liking, whereas, conceptually, feelings of TPV are *shared* (i.e., perceived similarly) by team members
Liking*	"[Favourable] attitude towards other group members" (Jehn, 1995, p. 266)	"I generally like the other members of my team" "My fellow team members are satisfied with being a member of this work unit" (Jehn, 1995)	
Cognitive Nature			
Information Sharing*	"central process through which team members collectively utilize their available informational resources" (Mesmer-Magnus & DeChurch, 2009)	"In your team, how much was the information used to make decisions freely shared among the members of the team" (Bunderson & Sutcliffe, 2002) "In my team, members were kept "in the loop" about key issues affecting our business unit" (Bunderson & Sutcliffe, 2002, cit. in Marlow et al., 2018)	Information sharing is a *process* (i.e., refers to something team members *do*, to the nature of their interaction), and is usually measured in terms of quality of information or frequency of exchanges (Marlow et al., 2018), reflecting a focus on the *utility* of *conveying* information for task-related purposes; TPV, on the other hand, is an *emergent state* (i.e., refers to a collective state of the team) that arises through (poor) information exchange
Team Coordination*	"Emergent phenomenon involving the use of strategies and behavior patterns aimed at integrating and aligning the actions, knowledge, and objectives of interdependent members, with a view to attaining common goals." (Rico et al., 2008, p. 163)	"Our team worked together in a well-coordinated fashion." (Lewis, 2003) "There was much confusion about how we would accomplish the task" (Lewis, 2003)	Coordination is *task related* and relates intrinsically to actions towards the task goal the team has. TPV, however, includes both experiences of distance and information deficits that may come from taskwork, but also develop through non-task related interactions/teamwork more generally

Table 1 (continued)

Construct	Definition(s)	Exemplary items	Main conceptual distinction from TPV
Effective Use of Technology for Virtual Communication*	"[Communicating] virtually with distributed teammates in a way that reduces the increased potential for misunderstandings, negative attributions, and adverse impact on the development of shared understanding (Hill & Bartol, 2016, p. 165)	"Our team communicates virtually with other team members in a way that is clear and easily understood" (Hill & Bartol, 2016) "Our team sends virtual communications with a positive and encouraging tone" (Hill & Bartol, 2016)	Closely linked to the collective-experienced information deficits dimension of TPV, effective use of technology for virtual communication focuses on *technology-mediated communication*. Information deficits, in turn, while likely influenced by the degree of technology usage, are not restricted to the moments where teams interact through technology but refer to a holistic experience of teamwork
Shared Mental Models (SMM)	"Organized understanding or mental representation of knowledge that is shared by team members" (Mohammed et al., 2000, p. 123)	"I know how the tasks of my team members are related to each other" (taskwork SMM; Konradt et al., 2015) "I know which team members have expertise in specific areas" (teamwork SMM; Ellwart et al., 2014) "Each team member knows the characteristics and functionalities of the used digital media " (ICT-SMM; Müller & Antoni, 2020)	SMM refer to a common understanding of taskwork, teamwork, and other aspects such as the use of information and communication technology (e.g., how/when/for what to use digital media). Although TPV's dimension of collectively-experienced information deficits also reflects a shared understanding, it is not about the *usage of technology* per se, but about information exchange more generally

*Scale included in Study 4a

Specifically, the effects of team virtuality depend neither on how much teams use technology to communicate, nor on specific features of these technologies (e.g., the capacity to transmit sound). Indeed, from a sociomateriality lens (Orlikowski, 2007), structural features of technology per se do not univocally lead to specific outcomes. Individuals' agency in using technology will result in distinct use and appropriation of technology features that will impact how teams experience distance and information deficits (Costa & Handke, 2023). At the individual level, for example, the use of electronic monitoring systems depends on the attitudes that individual hold towards them (Abraham et al., 2019; Alge & Hansen, 2014). Because how individuals use certain features of technology is constrained by, yet distinct from the built-in properties of that technology, capturing the subjective rather than structural nature of team virtuality is critical. Moreover, a range of contextual factors (e.g., team familiarity, team design) can impact team experiences and specifically the impact that structural virtuality has on these experiences (see Handke et al., 2021). Structural virtuality and these other factors can thus be considered as additive in their effect on team experiences, meaning that high levels of structural virtuality be compensated with other factors, such as high levels of team familiarity.

The motivation behind the development of TPV is thus to offer a continuous construct that can capture teamwork experiences regardless of how much teams rely on technology to interact, and regardless of what type of media supports these interactions. As a result, and given the many other factors that can contribute to teamwork experiences, TPV is not expected to be directly explained or predicted by structural virtuality. Based on the extant perspective on (structural) team virtuality in the literature, TPV is a deficit-oriented state, meaning that higher levels of its two constituting dimensions are associated with more negative consequences for team functioning. The differentiation between feelings of distance and perceptions of information deficits aligns with the general distinction into affective and cognitive team (emergent) states (e.g., Bell et al., 2018; Marks et al., 2001) and reflects two central themes in the team virtuality literature, as we will describe in the following.

Collectively-Experienced Distance

The affective dimension of TPV is characterized by collective feelings of distance, stemming from team members' mutual awareness of emotional inaccessibility, disconnectedness, and estrangement. The more distant team members feel, the less their relationship with one another will be characterized by warmth, affection, friendship, or intimacy. Importantly, feeling distant is not the same as being physically distant or interacting through technology rather than face-to-face. Accordingly, team members can work from different locations but still feel close to each other ("far-but-close") just as they can be physically co-located but feel distant from one another ("close-but-far";

see O'Leary et al., 2014; Wilson et al., 2008). Similarly, the concept of electronic propinquity (Korzenny, 1978; see also Walther & Bazarova, 2008) suggests that psychological feelings of nearness or presence are influenced not only by the physical properties of a communication medium but by the people using it. For instance, a team of geographically dispersed software engineers can feel close to one another because they are very expressive in their communication, adept at recognizing each other's emotional states, and excellent at giving each other support and guidance. At the same time, a team of financial analysts working at the same site can feel very distant from one another, with team members rarely popping by at each other's desks or offices to engage in spontaneous, informal conversation, and generally knowing very little about each other's actions, experiences, or preferences.

Collectively-Experienced Information Deficits

The cognitive dimension of TPV, in turn, is characterized by collective perceptions of poor information exchange. Handke et al. (2021) describe information exchange as poor when it "does *not* (1) enable timely feedback, (2) meet team members' personal requirements (e.g., by allowing to alter messages to enhance specific team members' understanding), (3) combine a variety of different cues (e.g., by conveying both the content of a message as well as its emotional tone), and (4) use rich and varied language (e.g., by enabling the use of symbol sets close to natural language; see also Carlson & Zmud, 1999)" (p. 170). In a team where information deficits are high, team members will have difficulties in hearing and being heard, such that they cannot reach a joint understanding of what everyone in the team is thinking, feeling, and doing. Whereas related terms such as information richness or synchronicity have largely been used in conjunction with physical media properties (e.g., Dennis et al., 2008; Kirkman & Mathieu, 2005), Handke et al. (2021) note that the experience of information deficits arises not from the communication medium per se but by how its use transpires into team interactions. For instance, a team communicating exclusively through communication technology may experience very low levels of information deficits because team members know each other's communication needs and styles and have learned to anticipate and react toward potential misunderstandings (see channel expansion theory, Carlson & Zmud, 1999). At the same time, a team relying primarily on face-to-face communication can be very high in information deficits as team members fail to react to each other's inquiries on time or to seek and give constructive feedback.

Measuring TPV

Despite the conceptual importance, there is still no scale to measure TPV. Measuring TPV requires considering existing structural indicators of team virtuality, as well as related

constructs. Table 1 gives an overview of measurements of structural virtuality as well as related constructs and measurements on both cognitive and affective dimensions. In the following, we will describe how the two dimensions collectively-experienced distance and information deficits differ from these other constructs and their measurement and which implication this has for the TPV scale development.

Structural virtuality indicators range from dichotomous measures (i.e., team members are either co-located or geographically distributed) to continuous measures of objective distance (i.e., how many miles team members are apart, Hoch & Kozlowski, 2014) or technology usage (i.e., how often members use technology to communicate, Rapp et al., 2010). In other studies, authors approach structural virtuality by assessing media characteristics, such as media richness/informational value (Daft et al., 1987; Kirkman & Mathieu, 2005; for applied examples see e.g., Brown et al., 2020; Ganesh & Gupta, 2010) or synchronicity (Dennis et al., 2008; Kirkman & Mathieu, 2005; for applied examples see e.g., Brown et al., 2020; Rico & Cohen, 2005). Studies drawing on these approaches either assign values of media richness or synchronicity based on the type of technologies teams use (e.g., face-to-face communication receives a higher weight than emails, see e.g., Ganesh & Gupta, 2010; Rico & Cohen, 2005) or alternatively ask for team members' perceptions of these characteristics (Brown et al., 2020). As described above, these operationalizations of virtuality refer either to geographic dispersion, objective technology usage, or media characteristics. What they do not capture is the holistic *experience* of working together virtually, as conceptualized through TPV.

Several existing constructs measure subjective experiences which are closely related to the two TPV dimensions. Specifically, on the affective dimension (i.e., collectively-experienced distance), TPV overlaps with constructs such as perceived proximity (O'Leary et al., 2014; Wilson et al., 2008), (electronic) propinquity (Walther & Bazarova, 2008), social presence (Short et al., 1976), co-presence (Zhao, 2003), cohesion (e.g., Carron et al., 1985; Festinger, 1950), belonging (Baumeister & Leary, 1995), and liking (e.g., Jehn, 1995). On the cognitive dimension (i.e., collectively-experienced information deficits), in turn, there are similarities to constructs such as information sharing (Mesmer-Magnus & DeChurch, 2009), team coordination (Rico et al., 2008), effective use of technology for virtual communication (Hill & Bartol, 2016), and shared mental models (e.g., Klimoski & Mohammed, 1994; Mohammed et al., 2000). In a first empirical attempt to capture TPV, Costa et al. (2023) draw on propinquity and effective use of technology for virtual communication as proxies for the two TPV dimensions. However, despite being a valid first approach to demonstrate the two-dimensional and subjective nature of team virtuality, the proxies employed in their study do not fully capture the construct of TPV. Specifically,

these proxies exhibit two noteworthy shortcomings, which also apply to the other related constructs named above and displayed in Table 1.

First, collectively-experienced distance and information deficits can be perceived by all teams, not solely those who are physically dispersed and/or interact predominantly through technology (i.e., are structurally virtual). As such, TPV cannot be adequately captured through measurements that directly tie these perceptions to dispersion and/or technology use, such as in the case of effective use of technology for virtual communication, perceived proximity, or co-presence. Second, both TPV dimensions are group-level properties reflecting team members' collective experiences, rather than perceptions on a dyadic level between individual communication partners, such as in the case of propinquity. The remaining constructs, which can be considered as group-level properties that are not restricted to dispersion/technology use, in turn, generally fail to capture the nature of TPV in its entirety. For instance, in terms of similar affective constructs, cohesiveness/cohesion does not adequately reflect the psychological nature of feeling estranged from and cold towards other team members. Similarly, shared mental models, which capture group-level representations of knowledge, do not extend to the general lack of information and inability to convey and converge on meaning that defines collectively-experienced information deficits.

Hence, although some of the existing constructs are closely linked to TPV, none of them address both the affective and cognitive states that are tied to the *team-level experience of virtuality*. In Table 1 we provide a comprehensive overview of related constructs and how these differ from the TPV. With this newly conceptually-grounded measure, we respond to current research and practice needs with (1) the assessment of shared teamwork experiences that are becoming commonplace in organizations, (2) the simultaneous appraisal of affective and cognitive components of such experiences, and (3) a psychometrically sound measure that is consistent with the latest literature on virtual teams. The following sections therefore describe the development and validation of the TPV construct. The datasets as well as R and Mplus output files for all five studies can be found here: https://osf.io/ag9m7/.

Study 1: Item Generation and Content Validity

Initially, two of the authors constructed a pool of 20 items (10 per dimension) to capture the two TPV dimensions (see Table 2). To do so, we closely attended Handke et al.'s (2021) definition of collectively-experienced distance and information deficits and further reviewed related measures of affective and cognitive team states (as depicted in Table 1).

In a first step, to explore whether these adequately reflected the two underlying dimensions—collectively-experienced distance and information deficits—we asked 22 raters (graduate students in management, specializing in human resources management) to allocate the 20 items to two different dimensions and to subsequently label each of these dimensions. The 20 items were presented in random order, so that distance and information deficits items were not presented together in two distinct blocks. Using R Statistical Software (v4.3.2, R Core Team, 2023) and the smacof package (v2.1–5, Mair et al., 2022), we performed a multidimensional scaling (MDS) procedure, which constructs a two-dimensional map that captures the distance between items based on the likelihood of these items being assigned to the same dimension. In the resulting MDS map, items with higher similarity (i.e., that were more often assigned to the same dimension), are located closer together than items that were less likely to be assigned to the same dimension. As can be seen in Fig. 1, the items intended to capture collectively-experienced distance and information deficits, respectively, clustered closely together at opposite ends of the MDS map. We further calculated Kruskal's (1964) stress statistic—which ranges from 0 (perfectly stable) to 1 (perfectly unstable)—to assess the stability of the MDS solution, i.e., the extent to which a particular matrix is not random (Sturrock & Rocha, 2000). We obtained a stress value of 0.071, which is below the 1% cut-off value for this number of objects (items) in a two-dimensional matrix (Sturrock & Rocha, 2000). This allowed us to conclude that the assignment of the 20 items to the two TPV dimensions was not random across the 22 raters, emphasizing that the items clearly differentiate between an affective and a cognitive dimension of TPV. The most common label given to the dimension consisting of the items developed to reflect information deficits was "(team/group) communication", whereas the most common labels for the items developed to reflect the collectively-experienced distance were "relationship" and "closeness/distance/proximity". Accordingly, raters were not only clearly able to differentiate between the items but also developed labels reflective of the two TPV dimensions.

In a second step, items were reviewed by a panel of well-known virtual team experts. The 11 reviewers were scholars in the field of teams, and specifically of virtual teams, who agreed to rate each of the items as *Essential, Useful,* or *Not necessary*. The Content Validity Ratio (CVR, Lawshe, 1975; Wilson et al., 2012) was used to measure the raters' agreement. CVR values ranged from 1 to 0.27 considering the relevant ratings: essential and useful. Using the critical value of 0.78 (for 11 panelists

Table 2 Study 1 and Study 2: Content Validity Ratio, Descriptives, Reliabilities, and Factor Loadings

		Study 1	Study 2			FL	
#		CVR	M	SD	α	1	2
Collectively-experienced distance					.93		
1	In my team, we feel at ease with each other. [R][a]	0.45					
2	In my team, members are emotionally inaccessible to each other.[a]	0.64					
3	In my team, we can confide in each other. [R][a]	0.27					
4	In my team, we feel that we like each other. [R][a]	0.45					
5	In my team, we feel detached from each other.	1	2.69	1.59		.89	
6	In my team, we feel that our relationship is cold.	1	2.61	1.69		.90	
7	In my team, we feel like we are far away from each other.	0.82	2.81	1.73		.90	
8	In my team, we feel estranged from each other.	0.82	2.59	1.65		.93	
9	In my team, we feel like we cannot get through to each other.	0.82	2.59	1.57		.90	
10	In my team we feel close to each other. [R][b]	1	2.84	1.38		0.41	
Collectively-experienced information deficits					.89		
	When we exchange information in my team …						
11	…it's difficult to ensure we can go back and check it.[a]	0.27					
12	… it has deficits.[a]	0.45					
13	…we do not react to each other's contributions quickly enough.[a]	0.64					
14	…we have trouble understanding each other.[a]	0.64					
15	…the ways in which we can express ourselves are limited.	0.82	2.84	1.58			.85
16	…it's hard to convey the actual meaning of what we are saying.	1	2.60	1.52			.91
17	…it's difficult to understand if we are on the same page or not.	1	2.68	1.56			.88
18	…we are unable to convey the necessary information in its entirety.	1	2.68	1.56			.85
19	…we don't know whether everyone has had access to the same information.	1	2.87	1.61			.78
20	…we can transmit the emotional tone of our message. [R][b]	0.82	2.89	1.29			.21

Study 1 N=11, Study 2 N=393. CVR=content validity ratio; FL=Standardized factor loading; R=Items recoded prior to analysis

[a] Items deleted after Study 1. [b] Items deleted after Study 2. [c] Level-1 (i.e., individual-level) factor loadings based on multilevel CFA
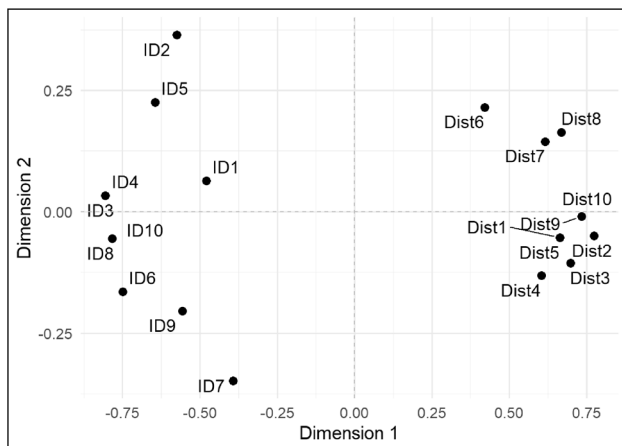
Fig. 1 MDS Map (Study 1). *Note.* N=22. ID=Information deficits, Dist=Distance

and $p = 0.01$, Lawshe, 1975), six items from each dimension could be retained, therefore leading to a final list of 12 items.

Study 2: Structural Validity (Individual-Level)

In this study, we tested the two-factor structure (i.e., collectively-experienced distance and collectively-experienced information deficits) of the 12-item TPV scale developed in Study 1 in a sample of 393 individuals. We thus performed confirmatory factor analysis (CFA) to assess model fit and further compared the two-factor model against an alternative one-factor model assuming TPV has no subdimensions. Finally, given that TPV has been conceived as a measure applicable to teams irrespective of their level of

structural virtuality, we further conducted measurement invariance analyses across low and high structural virtuality subsamples.

Sample and Procedure

We recruited 447 employees working in the United States via MTurk to participate in an online survey. Ethics approval for this research was granted by the Institutional Review Board of Claremont McKenna College under the project title *Development of the Team Perceived Virtuality Scale* (protocol number 2022-01-013A1). Four hundred and twenty-three participants completed the survey. To ensure data quality, we followed the recommendation for using Mturk samples to filter out employees who fail to correctly answer two attention check items in the first survey (Cheung et al., 2017). The first item asked employees whether they had ever had a fatal heart attack while watching television (Paolacci et al., 2010), whereas the second asked them to name their dream job (Peer et al., 2014).

This quality check resulted in a final sample of 393 employees. Employees' age ranged from 20 to 69 years ($M = 39.46$, $SD = 10.61$). All participants worked full-time and were employed in various industry sectors (e.g., IT/data analysis = 18.3%, retail/sales = 11.2%, finance/accounting = 8.7%). Employees' average team was 4.21 years ($SD = 7.48$), with the average size of their immediate team being 10.79 members (SD = 24.21). On average, they spent 57.78% ($SD = 29.43$, range: 0–100%) of their time interacting face-to-face with their teams.

Measures and Data Analysis

Participants rated all 12 TPV items (see Table 2) on a 7-point response scale from 1 *(strongly disagree)* to 7 *(strongly agree)*. To validate the factorial structure of the TPV scale developed in Study 1, we conducted a CFA with the MLR estimator using Mplus version 8.4 (Muthén & Muthén, 1998 – 2017),[1] with the 12 items loading onto two factors (i.e., collectively-experienced distance and information deficits). To evaluate model fit, we examined the χ^2, the standardized root mean square residual (SRMR), the root mean square of error of approximation (RMSEA), and the comparative fit index (CFI). The following cutoff values

___

[1] For open source software users, see also our analyses using the R package lavaan (0.6–16, Rosseel, 2012), which yield nearly identical results (apart from minor differences in selected model fit criteria, such as AIC, BIC and SRMR and minor differences in some estimates, both of which may be explained through different optimization methods used by the respective software). This also applies to the other SEM analyses in this paper, with the exception of the more complex MCFA which led to improper solutions in lavaan and was thus performed with Mplus.

were used to indicate acceptable model fit: CFI > 0.95 and RMSEA and SRMR < 0.08 (Browne & Cudeck, 1992; Hu & Bentler, 1999). We further tested the two-factor solution against a model with all items loading only on one global factor (i.e., assuming that TPV is a one-dimensional construct). Finally, for measurement invariance analyses, we divided the sample into two groups based on the percentage of time that participants engaged in face-to-face interactions with their teams: low structural virtuality (50–100% face-to-face interactions; $M = 76.05$, $SD = 15.08$, $n = 259$) versus high structural virtuality (0–49% face-to-face interactions; $M = 22.46$, $SD = 14.31$ $n = 134$).

Results

The two-factor solution exhibited a very good model fit ($\chi^2 = 93.21$, df = 53, p < 0.001, CFI = 0.981, RMSEA = 0.044, SRMR = 0.032), which was superior to the one-factor model ($\chi^2 = 271.47$, df = 54, $p < 0.001$, CFI = 0.897, RMSEA = 0.101, SRMR = 0.048), as shown by a significant Satorra-Bentler scaled (Satorra & Bentler, 2001) χ^2 difference ($\Delta\chi^2 = 49.71$, Δdf = 1 $p < 0.001$). However, factor loadings of the two reverse-coded items (see Table 2, item #10 and item #20) were much smaller than for all the other items loading onto the respective factor. Moreover, a further model in which these two reversed-coded items loaded onto an error method factor exhibited a superior model fit over the two-factor model ($\chi^2 = 75.62$, df = 51, $p < 0.001$, CFI = 0.988, RMSEA = 0.035, SRMR = 0.027; Satorra-Bentler scaled $\Delta\chi^2 = 16.60$, Δdf = 2, $p < 0.001$). We therefore decided to omit the two reverse-coded items from further analyses. Accordingly, the final TPV scale consisted of 10 items, loading onto two factors and showing an excellent model fit ($\chi^2 = 41.60$, df = 34, $p = 0.174$, CFI = 0.996, RMSEA = 0.024, SRMR = 0.017).

Given that the TPV subscales should be applicable at various levels of structural virtuality, we further wanted to confirm invariance of the TPV measurement model across individuals working under conditions of high versus low structural virtuality. To do so, we examined configural, metric, and scalar multigroup invariance of the two TPV dimensions (for an overview of measurement invariance conventions and reporting, see Putnick & Bornstein, 2016) based on the final 10-item scale. Configural invariance considers whether the same pattern of factor loadings holds across different subsamples/groups. Adequacy of the configural model is given by typical CFA guidelines. Accordingly, the model fit of the configural model in this analysis can be considered as very good ($\chi^2 = 87.31$, $df = 68$, p = 0.057, CFI = 0.99, RMSEA = 0.04, SRMR = 0.02). Given configural invariance, we then tested for metric invariance. Metric invariance constrains respective factor loadings to equality across groups.

As the overall model fit of the metric invariance model was not significantly worse than for the configural invariance model (metric invariance model: $\chi^2 = 94.63$, $df = 76$; Satorra-Bentler scaled $\Delta\chi^2$ 7.19, $\Delta df = 8$, $p = 0.516$), we found support for metric invariance, suggesting that relations between the items and their respective TPV dimensions were equivalent across both groups. In a final step, we tested for scalar invariance, which places additional equality constraints on item intercepts across groups. Once again, the model fit was not significantly worse than for the less constrained (i.e., metric) model (scalar invariance model: $\chi^2 = 105.17$, $df = 84$; Satorra-Bentler scaled $\Delta\chi^2$ 10.82, $\Delta df = 8$, $p = 0.212$), thereby providing support for full scalar invariance across low and high levels of structural virtuality. This means that participants view the TPV construct in a similar manner regardless of their structural virtuality. Next to measurement invariance, there was also no overall correlation between the degree of virtuality and the distance ($r = 0.02$, $p = 0.693$) or information deficits dimension ($r = 0.01$, $p = 0.853$). The latent factor correlation between the two dimensions was $\rho_{CFA} = 0.89$, 95% CI [0.84; 0.93].

Study 3: Structural Validity (Individual and Team Levels)

As TPV is a team-level construct, the purpose of Study 3 was to validate the 10-item solution derived from Study 2 on a sample of 1,087 teams. Accordingly, we performed multi-level CFA (MCFA) to assess model fit and further calculated measures of interrater reliability (ICC) and agreement (rwg) to justify the use of our scale at the team level.

Sample

Data was pulled from the ITPmetrics.com database[2] from 122 classes involving 5,121 student team members organized into 1,144 teams. These assessments were conducted between November 2022 to December 2022. Courses were in a multitude of disciplines, institutions, and countries. A total of 38.4% of the respondents identified as female, 59.5% as male, and the remaining were undisclosed. The mean participant age was 21.88 ($SD = 4.47$) and the average group size was 4.48 ($SD = 1.34$). To ensure data quality, we filtered out all participants that had failed to correctly answer two attention checks built into the online assessments. The

attention checks instructed participants to respond either with *strongly agree* (attention check 1) or *strongly disagree* (attention check 2). Eight hundred and twenty-four participants were filtered out based on failing these attention checks. We further deleted all teams consisting of only one respondent ($n = 48$), resulting in a final sample of 4,249 individuals, nested in 1,087 teams and 119 classes. In this final sample, participants' mean age was 21.75 ($SD = 4.27$), and 38.2% identified as female, 59.5% as male, and 2.3% were undisclosed. The average team size was 3.91 ($SD = 1.22$).

Measures and Data Analysis

Participants rated all 10 TPV items (see Table 2) on a 7-point response scale ranging from 1 *(strongly disagree)* to 7 *(strongly agree)*. To validate the factorial structure of the TPV scale on the individual and team levels, we first determined whether the data empirically justified aggregating the two TPV dimensions to the team level. To do so, we calculated intraclass correlations (ICCs) for the two TPV dimensions. ICC(1) represents the reliability of an individual's rating of the group mean (i.e., the proportion of total variance attributable to team membership), ICC(2) represents the reliability of the group average rating (Chen et al., 2005). ICC(1) values of 0.01 are considered as small, > 0.10 as medium, and > 0.25 as large (LeBreton & Senter, 2008). In their review of cut-off values for common aggregation indices, Woehr et al (2015) reported an average ICC(1) value of 0.22 ($SD = 0.15$) and an average ICC(2) value of 0.64 ($SD = 18$) for group-level constructs across 290 (for ICC(1)) and 239 (for ICC(2)) studies. Moreover, Woehr et al. (2015) also suggest comparing ICC indices obtained for actual team data to those of "pseudo" teams formed by randomly combining individual responses into "teams". This random group resampling procedure (RGR) provides estimates of the amount of within-unit reliability one would expect when individuals are not rating a common target (i.e., the same team). These estimates are then compared to those from the actual teams, with larger ICC estimates for observed teams relative to values obtained for pseudo teams (i.e., RGR results) providing support for team-level aggregation.

However, although ICC values provide important information on the relative between-team variance, the aggregation of individuals' referent-shift measures (i.e., item wordings aligned with team-level processes or states, as in the case of the TPV scale) needs to be supported through agreement indices (Chen et al., 2005; James, 1982). Accordingly, to support aggregating team members' individual responses to the team level, we further calculated James et al.'s (1984) rwg(j) agreement index using both a uniform (rectangular) null distribution and a slightly skewed null distribution (Biemann et al., 2012; LeBreton & Senter, 2008). Whereas the

---

uniform null distribution assumes that team members use the entire response scale, a slightly skewed null distribution corresponds to a leniency bias (meaning that individuals tend to choose lower values on the response scale, which may be likely in the case of TPV's deficit orientation) and are thus necessarily smaller than those based on a uniform null distribution (Woehr et al., 2015). LeBreton and Senter (2008) suggest to evaluate rwg(j) values based on the following categorization: lack of agreement (0.00–0.30), weak agreement (0.31–0.50), moderate agreement (0.51–0.70), strong agreement (0.71–0.90), or very strong agreement (0.91–1.00). Woehr et al. (2015) further suggest evaluating interrater agreement (rwg) relative to estimates reported in the literature. The authors report average values of 0.84 ($SD = 0.11$) for a uniform null and average values of 0.67 ($SD = 0.19$) for a slightly skewed null distribution across both group- and organizational-level constructs. We calculated all indices of agreement using the R package multilevel (version 2.7, Bliese et al., 2022).

In a second step, we performed MCFA with the MLR estimator using Mplus version 8.4 (Muthén & Muthén, 1998—2017) to account for the nested structure in our dataset (i.e., individuals as part of teams). MCFA models individual and team-level constructs simultaneously at both levels (see Dyer et al., 2005; Hirst et al., 2009). Following the results of Study 2, we assumed two latent factors representing the two TPV dimensions. We used the same criteria to assess model fit as for single-level CFA (see Study 2). Even though the team-level data (Level 2) was further nested in classes, ICC(1) values were so low for this level of analysis (i.e., < 0.05) that we did not consider class as a further level of analysis.

Results

Collectively-experienced distance showed the following team-level psychometric properties: ICC(1) = 0.25, ICC(2) = 0.56, median $r_{wg(j)uniform}$ = 0.96, median $r_{wg(j)skewed}$ = 0.94, α = 97. Team-level psychometric properties for collectively-experienced information deficits were: ICC(1) = 0.13, ICC(2) = 0.38, median $r_{wg(j)uniform}$ = 0.93, median $r_{wg(j)skewed}$ = 0.82, α = 0.89. The RGR procedure produced average ICC(1) values of 0.005 (collectively-experienced distance; $SD = 0.007$) and 0.004 (collectively-experienced information deficits; $SD = 0.007$) and average ICC(2) values of 0.019 (collectively-experienced distance; $SD = 0.026$) and 0.019 (collectively-experienced information deficits; $SD = 0.026$). Accordingly, even though the ICC(2) value for collectively-experienced information deficits was below 1 SD from the average ICC(2) values reported by Woehr et al. (2015), it was still substantially larger than the average (even larger than 99%) of the ICC(2) values obtained

for the pseudo teams. All other ICC values were near the average values or at least within 1 SD from the average values reported by Woehr et al. (2015). Finally, the rwg(j) values of both TPV subscales indicated strong to very strong agreement, even under the slightly skewed null distribution. In sum, these results suggest that aggregation to the team level is appropriate.

Furthermore, the MCFA showed an excellent model fit ($\chi^2 = 772.50$, $df = 68$, $p < 0.001$, CFI = 0.970, RMSEA = 0.049, SRMR$_{within}$ = 0.031, SRMR$_{between}$ = 0.058). This suggests that the two-factor solution fit the data well and consistently at both levels. Moreover, the two-factor model fit the data significantly better than the one-factor model ($\chi^2 = 4119.68$, $df = 70$, $p < 0.001$, CFI = 0.827, RMSEA = 0.117, SRMR$_{within}$ = 0.129, SRMR$_{between}$ = 1.137), as shown by a significant Satorra-Bentler scaled (Satorra & Bentler, 2001) χ^2 difference ($\Delta\chi^2 = 1,110.64$, $\Delta df = 2$, $p < 0.001$). Similar to Study 2, we observed high factor correlations between the two dimensions at the individual ($\rho_{CFA} = 0.70$, 95% CI [0.66; 0.73]) and team ($\rho_{CFA} = 0.99$, 95% CI [0.92; 0.1.06]) levels of analysis. Descriptive statistics as well as factor loadings at individual and team levels of analysis are displayed in Table 3. Although the specific technologies used were not measured, work with students highlights that approximately 66% of them rely on technological tools in the classroom (Brasca et al., 2022). Examples of these virtual tools include video conferencing platform (e.g., Zoom, Microsoft Teams) for synchronous communication, emails, and text messages for asynchronous communication, and Google drive or related software for project and information management (Horvat et al., 2021). Therefore, an array of structural virtuality options is available and was likely used by our sample.

Study 4: Conceptual and Empirical Distinctiveness & Criterion Validity

The purpose of Study 4 was to assess the conceptual distinctiveness (Sample 4a) as well as the discriminant (Subsample 4b) and criterion validity (Subsample 4c) of the two TPV dimensions. For conceptual distinctiveness, we followed the guidelines initially posed by Anderson and Gerbing (1991), recently updated by Colquitt et al. (2019). To further support the content validity of TPV, we sought out naïve judges (i.e., individuals without substantial experience in the field). We selected six orbiting constructs: three cognitive (i.e., *effectiveness of the use of technology for virtual communication, information sharing,* and *team coordination*) and three affective (i.e., *belonging, social cohesion,* and *liking*) constructs to compare to our focal dimensions collectively experienced

Table 3 Study 3: Descriptives, Reliabilities, and Factor Loadings, Study 4a: Substantive Agreement & Validity Coefficients

| | | Study 3 | | | | | | | Study 4a | | | |
		M	SD	α	FL$_i$ 1	FL$_i$ 2	FL$_t$ 1	FL$_t$ 2	p_{sa}	Avg	c_{sv}	Avg
#												
Collectively-experienced distance				.97						0.87		0.74
1	In my team, we feel detached from each other	2.18	1.48		.872		.992		0.93		0.86	
2	In my team, we feel that our relationship is cold	1.93	1.27		.898		.997		0.80		0.61	
3	In my team, we feel like we are far away from each other	2.02	1.38		.938		.991		0.89		0.78	
4	In my team, we feel estranged from each other	1.90	1.27		.920		1.00		0.88		0.76	
5	In my team, we feel like we cannot get through to each other	1.81	1.19		.831		.971		0.85		0.69	
Collectively-experienced information deficits				.93						0.72		0.45
	When we exchange information in my team …											
1	…the ways in which we can express ourselves are limited	2.07	1.25			.756		.990	0.63		0.26	
2	…it's hard to convey the actual meaning of what we are saying	1.98	1.18			.850		.972	0.77		0.54	
3	…it's difficult to understand if we are on the same page or not	2.17	1.38			.833		.939	0.70		0.40	
4	…we are unable to convey the necessary information in its entirety	1.98	1.24			.818		.958	0.76		0.52	
5	…we don't know whether everyone has had access to the same information	2.09	1.36			.698		.9724	0.76		0.52	

Study 3: Individual-level $N = 4,249$, team-level $N = 1,087$. Study 4a: $N = 117$. FL$_i$ = Standardized factor loading at individual level of analysis. FL$_t$ = Standardized factor loading at team level of analysis. p_{sa} = proportion of substantive agreement. c_{sv} = substantive validity coefficient

Table 4 Study 4b & 4c: Team-level Descriptive Statistics and Intercorrelations

		Subsample 4b						Subsample 4c					
		M	SD	1	2	3	4	M	SD	1	2	5	6
1	TPV – distance	1.93	0.94	-	.86**	-.69**	-.71**	1.82	0.85	-	.80**	-.40**	-.40**
2	TPV – information deficits	1.98	0.82	.84**	-	-.69**	-.69**	1.94	0.73	.78**	-	-.33**	-.35**
3	Team monitoring and backup	4.51	0.47	-.65**	-.65**	-	-.90**						
4	Coordination	4.47	0.50	-.69**	-.66**	.84**	-						
5	Team interaction quality							4.28	0.79	-.38**	-.32**	-	.67**
6	Team performance							4.28	0.85	-.39**	-.31**	.63**	-

$N = 643$ teams (Subsample 4b); $N = 400$ teams (Subsample 4c). $**p < .001$. Lower diagonal represents scale score correlations, upper diagonal represents latent factor correlations

information deficits and collectively experienced distance, respectively.

For discriminant validity, we compared the two TPV dimensions against two team action process measures: *team monitoring and backup* and *coordination*. Based on Marks et al.'s (2001) seminal team process framework, emergent states such as TPV need to be differentiated from other mediators in team effectiveness models, namely team processes. Whereas emergent states describe affective, cognitive, and motivational states of teams, team processes denote team members' "interdependent acts that convert inputs to outcomes through cognitive, verbal, and behavioral activities directed toward organizing taskwork to achieve collective goals" (Marks et al., 2001, p. 357). As Handke et al.'s (2021) theoretical model sees TPV arising from team interactions (with both TPV dimensions being negatively related to team action processes), the two TPV dimensions should thus be highly related, yet distinguishable from team action processes, such as team monitoring and backup and coordination. Common approaches to establish discriminant validity are to construct a measurement model in which latent variables that represent conceptually similar constructs are allowed to freely covary with each other (unconstrained model) to then a) calculate the factor correlations between these latent variables and b) compare this unconstrained model against constrained ones where the covariance of the latent variables is set to equal one or any self-defined cut-off reflecting a high correlation (see e.g., Rönkkö & Cho, 2022; Shaffer et al., 2016).

We performed these analyses using team-level CFA, similar to other group-level construct validation procedures (e.g., Mathieu et al., 2020; Tannenbaum et al., 2024). Accordingly, using team-level data, we first conducted a CFA with the MLR estimator using Mplus version 8.4 (Muthén & Muthén, 1998—2017) in which collectively-experienced distance, collectively-experienced information deficits, team monitoring and backup, and coordination were modelled as four first-order factors that were allowed to freely covary. Based on recommendations by Rönkkö and Cho (2022) we further

calculated 95% confidence intervals for the factor correlations between the two TPV constructs and the two action processes, whereby the absolute value of the upper limit should not exceed 0.80. In a next step, we compared this unconstrained model to four constrained models where the respective factor correlations had been set to -0.85 (based on a cut-off value suggested by Shaffer et al., 2016, and which leads to stricter assessments of discriminant validity than the typically employed cut-off value of 1, see Rönkkö & Cho, 2022). Significantly higher χ^2 values of the unconstrained model are considered to support discriminant validity. Please note that factor correlations were set to negative values given the deficit-oriented nature of the TPV subscales. The factor correlations can be seen in the upper diagonal in Table 4.

Criterion validity, in turn, is achieved when the TPV subscales relate to an external criterion that seems to be a result of TPV. Based on Handke et al.'s (2021) theoretical model, we focused on an affective (leader-rated team interaction quality) as well as a performance-related (leader-rated team performance) outcome. Specifically, we expect that collectively-experienced distance would be negatively related to leader-rated team interaction quality and that collectively-experienced information deficits would be negatively related to leader-rated team performance.

Sample 4a

We recruited 124 employees working in the United States via Prolific to participate in an online survey. All participants indicated to work in a team context more or less regularly. Participants were compensated £2.50 (approx. $3) for their time. Ethics approval for this research was granted by the Institutional Review Board of Claremont McKenna College under the same protocol number as for Study 2. First, participants were presented with a brief training to understand how to drag items to the construct boxes as proposed in Colquitt et al. (2019). To ensure data quality, we removed participants who did not pass one of the two attention checks, which

asked them to drag an item to a specific box. A final sample of 117 individuals, who took average of 12 min, completed the task. In this final sample, 40% of participants identified as female, the mean age was 38.61 years ($SD = 11.19$), and 80.87% worked full-time. A Q-sort task included reading 34 randomly ordered items and classifying them as best representing one of the defined constructs. These items came from either the 10-item TPV scale as our focal construct or one of the six orbiting constructs (e.g., information sharing, belonging, and liking). See Table 1 for the complete list of the orbiting constructs as well as their definitions and item sources. All employed items can also be found in the supplemental materials uploaded to OSF.

Data Analysis

We calculated the two recommended indices, the proportion of substantive agreement (p_{sa}) and the substantive validity coefficient (c_{sv}) as indicators of the TPV's distinctiveness in comparison to closely related constructs. The former ranges from 0 to 1, where higher values indicate the extent to which items were properly matched to the intended construct, whereas the latter ranges from -1 to 1 and it compares the matching to the appropriate construct versus all other orbiting constructs; therefore, a higher and negative value would indicate an item is more often linked to another construct than the expected one.

Results

Utilizing Colquitt et al.'s (2019) evaluation criteria of highly correlated constructs, all 10 items met the criteria for being moderate to very strong. None of our items represented any alternative construct over the expected dimensions of TPV. This indicates the appropriateness of our items as representing the content domain of TPV according to its dimensions, and better than any other orbiting scale. See Table 3 for the TPV results and the Supplemental Materials on OSF for the orbiting constructs' indices. In summary, we found that the dimension of experienced information deficits ($p_{sa} = 0.72$; $c_{sv} = 0.45$) differs from effectiveness of the use of technology for virtual communication, information sharing, and team coordination items while experienced distance ($p_{sa} = 0.87$; $c_{sv} = 0.74$) differs from belonging, social cohesion, and liking items. As a result, this finding significantly strengthens our confidence in the content validation of our scale when compared to established and validated scales.

Subsamples 4b and 4c

As with Study 3, both subsamples were pulled from the ITP-metrics.com database (but did not overlap with the sample used in Study 3). Similarly, we excluded all individuals who

failed the two attention checks, resulting in an overall sample of 4,213 (previously 5,030) individuals, nested in 1,143 teams (average team size: 3.69, SD = 1.27) and 121 classes. Assignment to the two subsamples was determined based on whether they contained leader ratings (Subsample 4c) or not (Subsample 4b). Six hundred and eighty teams had taken part in assessments that did not contain leader evaluations, a further 37 teams were omitted from further analyses as they were represented by only one respondent. Accordingly, the final Subsample 4b consisted of 2,402 individuals, nested in 643 teams. In Subsample 4b, participants' mean age was 22.20 ($SD = 4.18$) and 42.6% identified as female, 55.4% as male, and the remainder were undisclosed. The average team size was 3.74 ($SD = 1.17$). Four hundred and sixty-three teams had designated team leaders, and in some teams (less than 8%) there was more than one team leader (in this case, external ratings reflected mean ratings across team leaders). We eliminated all teams with only one (or no) team member other than the team leader, resulting in a final Subsample 4c of 400 teams, comprised of 1,272 team members (excluding team leaders). Team members' mean age in the final Subsample 24c was 22.61 ($SD = 4.61$) and 35.2% identified as female, 63.1% as male, and the remainder were undisclosed. The average team size was 3.05 ($SD = 0.95$).

Measures

TPV Dimensions

Collectively-Experienced Distance Collectively-experienced distance was captured with five items, as employed in Study 3 (see also Table 3), and rated on a 7-point response scale ranging from 1 *(strongly disagree)* to 7 *(strongly agree)*. In Subsample 4b, team-level psychometric properties were ICC(1) = 0.32, ICC(2) = 0.64, median $r_{wg(j)uniform} = 0.97$, median $r_{wg(j)skewed} = 0.95$, $\alpha = 0.98$. The RGR procedure produced average ICC(1) values of 0.007 ($SD = 0.010$) and average ICC(2) values of 0.023 ($SD = 0.034$). In Subsample 4c, team-level psychometric properties were ICC(1) = 0.26, ICC(2) = 0.51, median $r_{wg(j)uniform} = 0.98$, median $r_{wg(j)skewed} = 0.96$, $\alpha = 0.97$. The RGR procedure produced average ICC(1) values of 0.010 ($SD = 0.015$) and average ICC(2) values of 0.028 ($SD = 0.042$). In sum, ICC values were near the mean values reported by Woehr et al. (2015), and ICC values for the pseudo teams were substantially lower. Moreover, rwg(j) values indicated very strong agreement for collectively-experienced distance in both studies.

Collectively-Experienced Information Deficits Collectively-experienced information deficits were captured with five items, as employed in Study 3 (see also Table 3), and rated

on a 7-point response scale ranging from 1 *(strongly disagree)* to 7 *(strongly agree)*. In Study 4b, team-level psychometric properties were ICC(1) = 0.29, ICC(2) = 0.61, median $r_{wg(j)uniform}$ = 0.95, median $r_{wg(j)skewed}$ = 0.93, α = 0.96. The RGR procedure produced average ICC(1) values of 0.011 (*SD* = 0.016) and average ICC(2) values of 0.022 (*SD* = 0.034). In Subsample 4c, team-level psychometric properties were ICC(1) = 0.18, ICC(2) = 0.39, median $r_{wg(j)uniform}$ = 0.95, median $r_{wg(j)skewed}$ = 0.93, α = 0.93. The RGR procedure produced average ICC(1) values of 0.010 (*SD* = 0.015) and average ICC(2) values of 0.030 (*SD* = 0.044). In sum, ICC values were near the mean values reported by Woehr et al. (2015), with the exception of the ICC(2) values in Study 4c. However, as with Study 3, all ICC values in the two studies were substantially larger than the values obtained for the pseudo teams. Moreover, rwg(j) values indicated very strong agreement for collectively-experienced information deficits in both studies.

Team Action Processes

Team Monitoring and Backup Team monitoring and backup was measured using three items each, adapting items from Mathieu et al.'s (2020) team process survey measure (for an overview of the employed items, see O'Neill et al., 2020). An exemplary item is "We seek to understand each other's strengths and weaknesses". Items were rated on a five-point response scale ranging from 1 *(strongly disagree)* to 5 *(strongly agree)*. Team-level psychometric properties were ICC(1) = 0.26, ICC(2) = 0.57, median $r_{wg(j)uniform}$ = 0.98, median $r_{wg(j)skewed}$ = 0.97, α = 0.91.

Coordination Coordination was measured using three items each, adopting items from Mathieu et al.'s (2020) team process survey measure (for an overview of the employed items, see O'Neill et al., 2020). An exemplary item is "Our team smoothly integrates our work efforts". Items were rated on a five-point response scale ranging from 1 *(strongly disagree)* to 5 *(strongly agree)*. Team-level psychometric properties were ICC(1) = 0.27, ICC(2) = 0.58, median $r_{wg(j)uniform}$ = 0.98, median $r_{wg(j)skewed}$ = 0.97, α = 0.93.

Leader-Rated Team Outcomes

Leader-Rated Team Interaction Quality Leader-rated team interaction quality was based on four items from Wageman et al.'s (2005) Team Diagnostic Survey. An exemplary item is "There is a lot of unpleasantness among members of this team". Items were rated on a five-point response scale ranging from 1 *(strongly disagree)* to 5 *(strongly agree)*. All items were recoded, so that higher values reflected a higher quality of team interaction. Internal consistency was α = 0.81.

Leader-Rated Team Performance Leader-rated team performance was assessed with three items reflecting established performance criteria, namely efficiency, quality, and overall achievement (as employed by Van der Vegt & Bunderson, 2005). An exemplary item is "The team produces high-quality work". Items were rated on a five-point response scale ranging from 1 *(strongly disagree)* to 5 *(strongly agree)*. Internal consistency was α = 0.91.

Measurement Invariance Analyses

To confirm measurement invariance across the two subsamples, we examined configural, metric, and scalar multigroup invariance of the two TPV subdimensions. Analyses were performed at the team level. Given the very good fit of the configural model (χ^2 = 257.74, *df* = 68, *p* < 0.001, CFI = 0.98, RMSEA = 0.07, SRMR = 0.03), we tested first for metric, then for scalar invariance, both of which found full support (metric invariance model: χ^2 = 269.99, *df* = 76; Satorra-Bentler scaled $\Delta\chi^2$ 12.89, Δdf = 9, *p* = 0.116; scalar invariance model: χ^2 = 285.63, df = 84; Satorra-Bentler scaled $\Delta\chi^2$ 8.84, Δdf = 8, *p* = 0.356).

Results

Team-level descriptives and intercorrelations for both subsamples can be found in Table 4. For discriminant validity (Subsample 4b), we tested both factor correlations between collectively-experienced distance/ information and team monitoring and backup/coordination in the unconstrained model[3] and then compared this model to four constrained models in which the respective factor correlations had been fixed at -0.85. The unconstrained model yielded an excellent model fit (χ^2 = 253.89, df = 98, *p* < 0.001, CFI = 0.98, RMSEA = 0.050, SRMR = 0.023). Factor correlations for the latent variables were: ρ_{CFA} = -0.69 (distance with team monitoring and backup, 95% CI [-0.60; -0.78]), ρ_{CFA} = -0.71 (distance with coordination, 95% CI [-0.63; -0.80]), ρ_{CFA} = -0.69 (information deficits with team monitoring and backup, 95% CI [-0.58; -0.79]), and ρ_{CFA} = -0.69 (information deficits with coordination, 95% CI [-0.59; -0.79]). As can be seen in Table 4, the factor correlations between the two TPV (ρ_{CFA} = 0.86; 95% CI [0.83; 0.90]) as well the two action process subscales (ρ_{CFA} = 0.90, 95% CI [0.87; 94]) were substantially higher. Model comparisons showed

---

[3] For comparability across studies, we also performed MCFA for the unconstrained model but encountered an improper solution most likely attributable to the high correlation between the two action processes. We encountered no problems when running the MCFA with just the two TPV dimensions (see online supplemental material).

significant Satorra-Bentler scaled χ^2 differences between the unconstrained and all four constrained models (fixed correlation distance-team monitoring and backup: $\Delta\chi^2$ 46.72, $\Delta df = 1$, $p < 0.001$; fixed correlation distance-coordination: $\Delta\chi^2$ 38.01, $\Delta df = 1$, $p < 0.001$; fixed correlation information deficits-team monitoring and backup: $\Delta\chi^2$ 50.79, $\Delta df = 1$, $p < 0.001$; fixed correlation information deficits-coordination: $\Delta\chi^2$ 66.78, $\Delta df = 1$, $p < 0.001$). In sum, these results support the discriminant validity of the two TPV dimensions as emergent states versus team processes.

In terms of criterion validity (Subsample 4c), the intercorrelations shown in Table 4 support the assumed negative relationship between collectively-experienced distance and leader-rated team interaction quality as well as between collectively-experienced deficits and leader-rated team performance. The two TPV dimensions also correlated highly at $\rho_{CFA} = -0.80$, 95% CI [0.74; 86].

To complement the findings from the aforementioned samples, we gathered field data from organizational teams to gather further evidence for the generalizability of our TPV scale (Study 5).

Study 5: Generalizability to Organizational Teams

Even though student projects teams (such as in Study 3, Study 4b, and Study 4c) show effects of structural virtuality that are similar to organizational teams (Purvanova & Kenda, 2022), meta-analysis has also revealed that student samples do not significantly differ from one another in their demographics or even general observed correlations (Wheeler et al., 2014), meaning that it is possible that the variability in how students experience virtuality may be narrower than in other groups. Consequently, the aim of Study 5 was to test the generalizability of the TPV scale to organizational teams. Specifically, we sought to confirm the structural validity of TPV and its two dimensions on both individual and team levels in a sample of 1,063 individuals nested in 164 teams from a German organization specialized in international business services.

Sample and Procedure

Employees ($N = 2,820$) were contacted to participate in the online survey, which had been developed in collaboration with the Worker's Council of the participating organization. In accordance with the Ethics Code of the American Psychological Association, participants were first presented with an informed consent form that included information about the purpose of the study, the right to decline and withdraw their participation, prospective research benefits (here: potential improvement to hybrid working conditions implemented by

the Worker's Council and the human resource department), and how data was to be stored, used, and made accessible for research purposes. Work unit membership and demographic information was obtained via human resource data and connected to raw survey data through a unique identifier embedded into the URL of the survey link sent to participants' work email address. Of the 1,351 participants that participated in the survey, 1,219 fully completed it and 1,073 of these could also be clearly assigned to a team.[4] Deleting all teams with only one participant resulted in a final sample size of 1,063 individuals, nested in 164 teams. The average team size was 6.48 ($SD = 3.11$, range: 2 – 16). Sixty-three and a half percent of respondents identified as female, the mean age was 36.15 years ($SD = 8.64$), and the mean organizational tenure was 4.75 years ($SD = 5.02$). There were 65 different nationalities represented in the sample, the majority being German (61.34%). The majority of the sample worked full-time (86.64%), On average, participants worked remotely for 88.16% ($SD = 21.09$, range: 0 – 100%) of their working time.

Measures and Data Analysis

Participants rated all 10 TPV items on a 7-point response scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). Team-level psychometric properties for collectively experienced distance were ICC(1) = 0.14, ICC(2) = 0.51, median $r_{wg(j)uniform} = 0.88$, median $r_{wg(j)skewed} = 0.81$, $\alpha = 0.92$. The RGR procedure produced average ICC(1) values of 0.006 ($SD = 0.010$) and average ICC(2) values of 0.036 ($SD = 0.055$). Team-level psychometric properties for collectively experienced information deficits were ICC(1) = 0.08, ICC(2) = 0.37, median $r_{wg(j)uniform} = 0.90$, median $r_{wg(j)skewed} = 0.84$, $\alpha = 0.98$. The RGR procedure produced average ICC(1) values of 0.007 ($SD = 0.010$) and average ICC(2) values of 0.038 ($SD = 0.057$). In sum, ICC values were within the range of 1 SD around the mean values reported by Woehr et al. (2015)—once again with the exception of ICC(2) values for information deficits—and ICC values for the pseudo teams were substantially lower. Moreover, rwg(j) values indicated strong agreement for both TPV subscales, even under the slightly skewed distribution.

---

[4] To ensure that participants had the same team in mind when answering the TPV items, we asked them to think of the members who reported to the same disciplinary leader as them, which in this case corresponded to a work unit in the provided HR data. However, in some cases, work units had no designated leader (due to turnover, restructuring, or because they belonged to special status groups, e.g., apprentices, worker's council members), making it unclear whether the respective participants would still have the same team in mind or think of e.g., higher-level work units, such as their department. To avoid any misalignments in referent shift, we thus decided to exclude these individuals from further analysis.

Table 5 Study 5: Descriptives, Reliabilities, and Team + Individual-level Item Factor Loadings

#			M	SD	α	FL_i 1	FL_i 2	FL_t 1	FL_t 2
Collectively-experienced distance					.95				
1	In my team, we feel detached from each other.		3.29	1.56		.621		.988	
	[In meinem Team fühlen wir uns voneinander losgelöst.]								
2	In my team, we feel that our relationship is cold.		2.50	1.41		.863		.957	
	[In meinem Team fühlen sich unsere Beziehungen untereinander unterkühlt an.]								
3	In my team, we feel like we are far away from each other.		2.74	1.51		.877		.999	
	[In meinem Team fühlen wir uns weit voneinander entfernt.]								
4	In my team, we feel estranged from each other.		2.38	1.40		.916		.985	
	[In meinem Team fühlen wir uns voneinander entfremdet.]								
5	In my team, we feel like we cannot get through to each other.		2.17	1.27		.847		.987	
	[In meinem Teamfühlt es sich so an, als könnten wir nicht zueinander finden.]								
Collectively-experienced information deficits					.92				
	When we exchange information in my team …								
	[Wenn wir in meinem Team Informationen austauschen…]								
1	…the ways in which we can express ourselves are limited.		2.37	1.29			.769		.868
	[… sind die Wege, über die wir uns ausdrücken, eingeschränkt.]								
2	…it's hard to convey the actual meaning of what we are saying.		2.27	1.21			.913		.989
	[…ist es schwer, die tatsächliche Bedeutung von dem, was wir sagen, zu vermitteln.]								
3	…it's difficult to understand if we are on the same page or not.		2.30	1.25			.898		.966
	[… ist es schwer zu verstehen, ob wir dasselbe meinen oder nicht.]								
4	…we are unable to convey the necessary information in its entirety.		2.20	1.25			.835		.999
	[… sind wir nicht fähig, die nötigen Informationen in ihrer Gänze zu vermitteln.]								
5	…we don't know whether everyone has had access to the same information.		2.75	1.57			.706		.989
	[… wissen wir nicht, ob alle Zugang zu denselben Informationen hatten.]								

Brackets enclose German items. FLi = Standardized factor loading at individual level of analysis. FLt = Standardized factor loading at team level of analysis. Individual-level N = 1063, Team-level N = 164

In sum, we thus considered aggregation to the team-level as justified. We then performed MCFA with the MLR estimator using Mplus version 8.4 (Muthén & Muthén, 1998—2017). Given the international composition of the participating organization, the survey could be taken in either German (with items developed through back translation with native speakers, see Table 5) or English. We thus also tested for measurement invariance across the two languages (German: n = 813; English: n = 250).

Results

The MCFA showed an excellent model fit (χ^2 = 300.80, df = 68, p < 0.001, CFI = 0.966, RMSEA = 0.057, SRMR$_{within}$ = 0.035, SRMR$_{between}$ = 0.053), suggesting that the two-factor solution fit the data well and consistently at both levels. Moreover, the two-factor model fit the data significantly better than the one-factor model (χ^2 = 1687.83, df = 70, p < 0.001, CFI = 0.761, RMSEA = 0.147, SRMR$_{within}$ = 0.098, SRMR$_{between}$ = 0.738), as shown by

a significant Satorra-Bentler scaled (Satorra & Bentler, 2001) χ^2 difference ($\Delta\chi^2$ = 5,591.57, Δdf = 2, p < 0.001). Factor correlations were ρ_{CFA} = 0.71 (95% CI [0.66; 76]) at the individual level and ρ_{CFA} = 0.85 (95% CI [0.67; 1.03]) at the team level. Descriptive statistics as well as factor loadings at individual and team levels of analysis are displayed in Table 5. Given intra-team variations in language settings, measurement invariance analyses were performed at the individual level, but variables were group-mean centered prior to analysis to account for the nestedness of the data. Given the very good fit of the configural model (χ^2 = 253.22, df = 68, p < 0.001, CFI = 0.96, RMSEA = 0.07, SRMR = 0.03), we then tested for metric invariance. While we could not find support for full metric invariance (χ^2 = 277.49, df = 76, p < 0.001; compared to configural model: Satorra-Bentler scaled $\Delta\chi^2$ 22.15, Δdf = 8, p = 0.005) we did find support for partial metric invariance, in which four of the loadings of the distance factor were allowed to vary across languages (χ^2 = 258.28, df = 72, p < 0.001; compared to configural model: Satorra-Bentler scaled $\Delta\chi^2$ 1.92, Δdf = 4, p = 0.750). We further tested team-level correlations

between the TPV subscale means and degree of structural team virtuality (i.e., team levels of remote work). Both correlations were small and insignificant (structural virtuality – distance: $r = $ -0.05, $p = 0.553$; structural virtuality – information deficits: $r = $ -0.15, $p = 0.051$).

Discussion

With the proliferation of work arrangements that allow individuals to work from various locations and the technological developments to facilitate information exchange and interaction at a distance, teams are now working virtually on a regular (if not daily) basis. Therefore, understanding how teams (regardless of how much, how often or how rich/synchronous their use of technology is) experience their virtual collaboration is a necessary condition for designing their work in a way that promotes effectiveness while catering to individual team members' needs and intervening in times when teamwork is impaired. The existing literature on team virtuality has almost exclusively focused on structural features, such as the amount of time individuals interact via technology, whether these technologies allow for synchronous communication (e.g., Dennis et al., 2008; Kirkman & Mathieu, 2005), or how much they approximate face-to-face interaction (e.g., Daft et al., 1987; Kock, 2004). However, the heterogeneous findings linked to team virtuality effects (e.g., Handke et al., 2020; Purvanova & Kenda, 2022; De Guinea et al., 2012) suggest that current virtuality conceptualizations and measurements cannot fully capture the actual practice of virtual teamwork. Specifically, they do not capture technology use in practice and thus do not acknowledge how teams actually experience virtual interactions. Accordingly, this paper addressed how teams *perceive* their level of virtuality, based on the conceptual proposal by Handke et al. (2021) of Team Perceived Virtuality.

More specifically, we developed and validated the TPV scale, an instrument that accounts for the subjective experience of distance and information deficits in teams with different degrees of structural virtuality. The results of our five studies provide support for the main conceptual propositions around TPV put forward by Handke et al. (2021). First, we found support for a bi-dimensional structure of the construct across distinct samples and levels of analysis, with one dimension reflecting a more affective nature (distance), and another reflecting a more cognitive nature (information deficits) of TPV. Second, we provide evidence for the team-level essence of TPV, supporting its definition as a team emergent state. Third, our data reflects the predictive power of TPV on both affective and performance-related team outcomes (based on leader ratings to avoid same source bias) and its significant connection to team processes. Fourth, we provide evidence for the need of the TPV construct, which

is distinct from structural virtuality (as evident from its non-significant correlations) as well as from neighboring and related constructs that capture the perceived outcomes of structural virtuality (e.g., belonging, information sharing, etc.). Taken together, our results corroborate the multidimensional and team-level nature of TPV and its relationship with team outcomes.

We do want to note that albeit support for the multidimensional nature of TPV, we also observed very high factor correlations between collectively-experienced distance and information deficits at the team level of analysis. This pattern of high factor correlations between dimensions of the same scale resembles many other validated multidimensional scales at both team (e.g., team processes, Mathieu et al., 2020; team resilience, Tannenbaum et al., 2024) and individual levels (e.g., work engagement, Schaufeli et al., 2006; psychological empowerment, Spreitzer, 1995). However, these results also suggest that researchers may want to consider computing an overall TPV score that would correspond to Handke et al.'s (2021) framing of TPV as an aggregate of the two dimensions. Alternatively, as we do find support for a (conceptual) distinctiveness between the two dimensions as affective versus cognitive facets of TPV, researchers may also opt for one of the two dimensions depending on their outcome of interest.

Theoretical Contribution

Even though task-media-fit theories (e.g., media richness theory, Daft et al., 1987; task-media fit hypothesis, McGrath & Hollingshead, 1993; media synchronicity theory, Dennis & Valacich, 1999; Dennis et al., 2008) have greatly contributed to our understanding of technology use, they paint a fairly static picture of team virtuality and its effects. Particularly the more traditional "cues-filtered-out theories" (see e.g., Walther & Parks, 2002) generally assume that media have inherent characteristics (i.e., information richness), which will be experienced similarly by all individuals and remain fixed over time. Moreover, a vast number of studies analyzing the effects of team virtuality are still based on experiments concentrating on short-term use and effects, based on isolated tasks performed by ad-hoc groups or dyads, who have no shared history or future of working together (e.g., Hambley et al., 2007; Rico & Cohen, 2005).

In reality, however, organizational teams are likely to experience more variety and freedom in their taskwork and will gain experience in working together over the course of their interaction. This not only makes individuals more adept at using technology in different ways, it also makes them perceive a technology as being richer (see e.g., Carlson & Zmud, 1999; Fuller & Dennis, 2009). For instance, given enough message exchange, technology-mediated

communication has been shown to exhibit high degrees of relational communication (Walther & Tidwell, 1995; Walther, 1992, 1994). Moreover, even experimental research shows no differences between face-to-face and technology-mediated groups' performance when these worked on multiple/successive tasks (Fuller & Dennis, 2009; Simon, 2006; van der Kleij et al., 2009). These findings are supported by recent meta-analyses (Carter et al., 2019; Purvanova & Kenda, 2022), which show no effect of structural team virtuality on team performance outcomes in organizational/longer-tenured teams and thereby emphasize the shortcomings of structural virtuality operationalizations.

Accordingly, the validated TPV construct highlights how important it is to capture technology use in practice and subjective teamwork experiences to explain the differential effect of structural team virtuality. Specifically, it is less the physical properties of technology (e.g., the transmission of voice) or its built-in features (e.g., whether they allow group messaging or notifications) but how each team uses and acts upon these (e.g., who is invited to the group chat, which notifications are turned off) that will influence teamwork experiences (Costa & Handke, 2023). This is also reflected in the low and insignificant correlations between TPV and structural virtuality we found in Studies 2 and 5, likely signaling the stronger impact of other contextual factors (e.g., team familiarity, team design) or team processes on teamwork experiences. We do note, however, that these low/insignificant correlations could also be due to the operationalization of structural virtuality (here, as the team-level percentage of remote work time). Structural virtuality has been operationalized on many different dimensions (e.g., dispersion, reliance on communication technology, informational value of communication technology) and considering only one of these dimensions as a measure of structural virtuality may be an incomplete operationalization of the construct, leading to lower relationships with TPV. Finally, the rise of hybrid work practices likely increases intra-team variation in structural virtuality. Hence aggregating structural virtuality to the team level through a compositional (rather than compilational) approach can often be a less accurate picture of how teams collectively use technology to communicate and coordinate their actions.

To overcome the shortcomings of structural virtuality and its extant operationalization, we thus aim to contribute to the broader discussion on the subjective rather than objective measurement of constructs, as it has been debated in other literatures, such as organizational culture and climate (Glisson & James, 2002; Yammarino, & Dansereau, 2011). In sum, this work contributes to newer conceptualizations of technology use and team virtuality through highlighting its social construction and subjectivity, as well as the key role interactions between technology features and human agency (Orlikowski, 2007).

Practical Implications

The present work provides a robust instrument to address teams' subjective virtuality experiences. As such, it can guide practitioners and team leaders when managing teams, regardless of their structural level of virtuality. More specifically, it calls their attention to three main aspects. First, teams need to work on both distance and information deficit perceptions. For instance, leaders should provide time and space for enough interpersonal exchanges between members to reduce distance perceptions as well as continuously monitor team members' ability to achieve a joint understanding in order to manage information deficits perceptions. Second, the co-construction of meaning around technology use needs to be done proactively, rather than solely as a reaction to interaction impairments. For example, leaders can promote functional sensemaking (Morgeson et al., 2010) about technology usage and its consequences by letting the team reflect on their technology usage, its impact on task and relational function, and how it may be optimized. This regular practice can also be coupled with the intentional development of shared mental models around technology usage (Müller & Antoni, 2020), by which members develop a common understanding of how, when, and for what they are expected to use technology, therefore reducing misinterpretations and communication impairments. Third and finally, acknowledging the importance of subjective virtuality experiences calls attention to the eventual unintended consequences of technology usage (Soga et al., 2021) that can influence team perceptions. Understanding that, for example, allowing team members to email others on a Saturday can result in increased technostress (e.g., Salanova et al., 2012) can prompt leaders to use other technology features (such as email scheduling tools) to circumvent potential pitfalls. It is through the assessment of TPV that managers can flag when socialization activities and communication interventions are needed. In this post-pandemic era in which workers are co-located, hybrid, and in-person, the ability to optimize face-to-face interactions among workers and identify potential communication gaps holds significant value. Being able to make the most of employees' time by fostering quality interactions and pinpointing any communication deficiencies can greatly benefit organizations.

Limitations and Future Research

Parallel to the contributions highlighted above, the present work has some limitations that constrain the generalizability of its findings. First, we show different types of validation studies across samples and contexts, but the specific influence of technology has not yet been tested. Although conceptually TPV can be applied across teams, there may be some empirical nuances regarding technology. Future

research can accommodate this limitation by manipulating the technology used and measuring TPV at the same time. It can also enhance our understanding of the generalizability of the criterion-related validity by utilizing various team outcomes, such as objective team performance. Future research should, therefore, continue to validate this TPV measure across contexts and sample types to properly expand its nomological network. Moreover, we included leader-rated team outcomes in Study 4c but were unable to define or evaluate how these leaders were designated. Hence, future studies with organizational team samples should look into specific leadership details when studying TPV.

Second, four of our studies employed the original English version of the scale, with English-speaking samples mostly from North America. When applying this to a mostly German sample (Study 5), we uncovered the potential for measurement nuances across cultures. The construct can, therefore, gain from cross-cultural validation with individuals from other geographic regions and cultures. Although culture might influence the relative importance of each dimension on their relationship to team's outcomes, exploring how different cultural contexts shape these team processes and contribute differently to the emergence of TPV would be promising.

Third, although TPV seems promising in revealing "red flags" in team dynamics, we have not empirically established the contingencies in how TPV relates to team outcomes. To properly understand the complexity of perceived virtuality in real teams, having a psychometrically sound measure is the first step. Researchers can now continue to validate this measure including more contextual and temporal elements. Specifically, future research can enhance complexity in two ways: longitudinally and with multi-level contingencies. There are individual and organizational characteristics that can be further explored as suppressors or accelerators in the relationship between TPV and team outcomes. For instance, other empirical work in virtual teams has shown how cognitive reappraisal, an emotion regulation strategy (Theodorou et al., 2023), and empowerment climate (Nauman et al., 2010) can function as moderators that may change how TPV relates to important team outcomes. Some other contingencies could include team size, team design, and team evolution and maturation (e.g., Handke et al., 2021; Kirkman & Mathieu, 2005).

Fourth, in the realm of multi-level measurement, our scale was crafted to align with the referent-shift composition model (Chan, 1998), and our results show that team members tend to agree to a shared perception of TPV (cf. rwg and ICC values). Nonetheless, it is worth noting that compilation models, where certain team members may offer nuanced perspectives of TPV that challenge the extent of sharedness necessary for the emergence of this construct, are worth exploring in future studies. This will help to better understand what can influence the degree of team members' agreement on TPV. Again, within the climate literature, the concept of climate strength (i.e., the degree of within-unit agreement on climate perceptions), has revealed potential for considering different composition models of team constructs. Thus, we hope future research continues to use this instrument and uncovers more of its properties within a more complex environment.

Lastly, further analyses need to be conducted to assess the incremental validity of the two TPV dimensions. In their conceptual model, Handke et al. (2021) propose that whereas perceived distance would be more predictive of affective outcomes, perceived information deficits would contribute to understanding teams' task-related performance. As such TPV should ideally help us understand why (virtual) team performance and satisfaction often do not align (e.g., Simon, 2006; Van der Kleij et al., 2009). Recent research with proxies (i.e., related constructs) of TPV has provided initial support for these differential relationships (Costa et al., 2023). However, in our scale validation, Study 4c showed higher correlations between the two outcomes and perceived distance than between the respective outcomes and perceived information deficits. As elaborated earlier, these relationships would need to be tested in more complex models with organizational field data in order to assess whether the differential relationship between the two TPV dimensions and team outcomes truly exists.

Conclusion

We put forth a measure of TPV that is conceptually distinct to structural virtuality, allowing for a more dynamic and subjective approach to virtuality. Through five studies, we find support for content, construct, and criterion-related validity for the operationalization of the construct coined by Handke et al. (2021). Specifically, virtual team experts provided support for face validity to the TPV measure and naïve judges were able to discriminate between items that measured TPV and those that measured related constructs. Moreover, confirmatory factor analyses across multiple samples showed a very good model fit at both individual and team levels of the final 10-item measure, consisting of two dimensions each: collectively-experienced distance and collectively-experienced information deficits. Furthermore, our measure shows discriminant validity to important team processes and criterion-related validity to leader-rated interaction quality and team performance. Taken together, this newly developed TPV measure offers substantive evidence that shows its relevance to science and practice for addressing team virtuality as it is experienced.

Declarations

Conflict of Interest The authors declare that they have no conflict of interest.

References

Abraham, M., Niessen, C., Schnabel, C., Lorek, K., Grimm, V., Möslein, K., & Wrede, M. (2019). Electronic monitoring at work: The role of attitudes, functions, and perceived control for the acceptance of tracking technologies. *Human Resource Management Journal, 29*(4), 657–675.

Alge, B. J., & Hansen, S. D. (2014). Workplace monitoring and surveillance research since "1984": A review and agenda. In M. D. Coovert & L. F. Foster Thompson (Eds.), *The psychology of workplace technology* (pp. 209–237). Routledge.

Anderson, J. C., & Gerbing, D. W. (1991). Predicting the performance of measures in a confirmatory factor analysis with a pretest assessment of their substantive validities. *Journal of Applied Psychology, 76*(5), 732–740. https://doi.org/10.1037/0021-9010.76.5.732

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin, 117*(3), 497–529. https://doi.org/10.1037/0033-2909.117.3.497

Bell, S. T., Fisher, D. M., Brown, S. G., & Mann, K. E. (2018). An approach for conducting actionable research with extreme teams. *Journal of Management, 44*(7), 2740–2765. https://doi.org/10.1177/0149206316653805

Biemann, T., Cole, M. S., & Voelpel, S. (2012). Within-group agreement: On the use (and misuse) of rWG and rWG (J) in leadership research and some best practice guidelines. *The Leadership Quarterly, 23*(1), 66–80. https://doi.org/10.1016/j.leaqua.2011.11.006

Bliese, P., Gilad, C., Downes, P., Schepker, D., & Lang, J. (2022*). multilevel: Multilevel Functions.* R package version 2.7. https://CRAN.R-project.org/package=multilevel. Accessed 15 Nov 2023.

Brasca, C., Kirshan, C., Marya, V., Owen, K., Sirois, J., & Ziade, S. (2022). *How technology is shaping learning in higher education.* McKinsey & Company. https://www.mckinsey.com/industries/education/our-insights/how-technology-is-shaping-learning-in-higher-education#/.

Brenan, M. (2020). *U.S. Workers Discovering Affinity for Remote Work.* Gallup. https://news.gallup.com/poll/306695/workers-discovering-affinity-remote-work.aspx.

Brown, M. I., Prewett, M. S., & Grossenbacher, M. A. (2020). Distancing ourselves from geographic dispersion: An examination of perceived virtuality in teams. *Group Dynamics: Theory, Research, and Practice, 24*(3), 168–185. https://doi.org/10.1037/gdn0000120

Browne, M. W., & Cudeck, R. (1992). Alternative Ways of Assessing Model Fit. *Sociological Methods & Research, 21*(2), 230–258. https://doi.org/10.1177/0049124192021002005

Bunderson, J. S. & Sutcliffe, K. M. (2002). Comparing alternative conceptualizations of functional diversity in management teams: Process and performance effects. *Academy of Management Journal, 45*(5), 875–893. https://doi.org/10.5465/3069319

Carless, S. A., & De Paola, C. (2000). The measurement of cohesion in work teams. *Small Group Research, 31*(1), 71–88. https://doi.org/10.1177/104649640003100104

Carlson, J. R., & Zmud, R. W. (1999). Channel expansion theory and the experiential nature of media richness perceptions. *Academy of Management Journal, 42*(2), 153–170. https://doi.org/10.2307/257090

Carron, A. V., Widmeyer, W. N., & Brawley, L. R. (1985). The development of an instrument to assess cohesion in sport teams: The Group Environment Questionnaire. *Journal of Sport and Exercise Psychology, 7*(3), 244–266. https://doi.org/10.1123/jsp.7.3.244

Carter, K. M., Mead, B. A., Stewart, G. L., Nielsen, J. D., & Solimeo, S. L. (2019). Reviewing work team design characteristics across Industries: Combining meta-analysis and comprehensive synthesis. *Small Group Research, 50*(1), 138–188. https://doi.org/10.1177/1046496418797431

Chan, D. (1998). Functional relations among constructs in the same content domain at different levels of analysis: A typology of composition models. *Journal of Applied Psychology, 83*(2), 234–246.

Chen, G., Mathieu, J. E., & Bliese, P. D. (2005). A framework for conducting multi-level construct validation. *In Multi-level issues in organizational behavior and processes* (pp. 273–303). Emerald Group Publishing Limited. https://doi.org/10.1016/S1475-9144(04)03013-9

Cheung, J. H., Burns, D. K., Sinclair, R. R., & Sliter, M. (2017). Amazon Mechanical Turk in organizational psychology: An evaluation and practical recommendations. *Journal of Business and Psychology, 32*(4), 347–361. https://doi.org/10.1007/s10869-016-9458-5

Chung, B. G., Ehrhart, K. H., Shore, L. M., Randel, A. E., Dean, M. A., & Kedharnath, U. (2020). Work Group Inclusion: Test of a Scale and Model. *Group & Organization Management, 45*(1), 75–102. https://doi.org/10.1177/1059601119839858

Colquitt, J. A., Sabey, T. B., Rodell, J. B., & Hill, E. T. (2019). Content validation guidelines: Evaluation criteria for definitional correspondence and definitional distinctiveness. *Journal of Applied Psychology, 104*(10), 1243–1265. https://doi.org/10.1037/apl0000406

Costa, P. L., & Handke, L. (2023). Virtuality and the eyes of the beholder: beyond static relationships between teams and technology. In L. L. Gilson, T. O'Neill, & M. T. Maynard (Eds.), *Handbook of Virtual Work* (pp. 199–215). Edward Elgar Publishing. https://doi.org/10.4337/9781802200508.00020

Costa, P., Handke, L., König, M., & Thieme, O. (2023). *Team perceived virtuality: Empirical exploration of its two dimensions.* Theory, Research, and Practice. Advance online publication. https://doi.org/10.1037/gdn0000202

Daft, R. L., Lengel, R. H., & Trevino, L. K. (1987). Message equivocality, media selection, and manager performance: Implications for information systems. *MIS Quarterly, 11*(3), 355–366. https://doi.org/10.2307/248682

De Guinea, A. O., Webster, J., & Staples, D. S. (2012). A meta-analysis of the consequences of virtualness on team functioning.

Information & Management, 49(6), 301–308. https://doi.org/10.1016/j.im.2012.08.003

Dennis, A. R., Fuller, R. M., & Valacich, J. S. (2008). Media, tasks, and communication processes: A theory of media synchronicity. *MIS Quarterly, 32*(3), 575–600. https://doi.org/10.2307/25148857

Dennis, A. R., & Valacich, J. S. (1999). Rethinking media richness: Towards a theory of media synchronicity. In *Proceedings of the 32nd Annual Hawaii International Conference on System Sciences* (pp. 1–10). IEEE. https://doi.org/10.1109/HICSS.1999.772701

DeSanctis, G., & Poole, M. S. (1994). Capturing the complexity in advanced technology use: Adaptive structuration theory. *Organization Science, 5*(2), 121–147. https://doi.org/10.1287/orsc.5.2.121

Dixon, K. R., & Panteli, N. (2010). From virtual teams to virtuality in teams. *Human Relations, 63*(8), 1177–1197. https://doi.org/10.1177/0018726709354784

Dulebohn, J. H., & Hoch, J. E. (2017). Virtual teams in organizations. *Human Resource Management Review, 27*(4), 569–574. https://doi.org/10.1016/j.hrmr.2016.12.004

Dyer, N. G., Hanges, P. J., & Hall, R. J. (2005). Applying multilevel confirmatory factor analysis techniques to the study of leadership. *The Leadership Quarterly, 16*(1), 149–167. https://doi.org/10.1016/j.leaqua.2004.09.009

Ellwart, T., Konradt, U., & Rack, O. (2014). Team mental models of expertise location: Validation of a field survey measure. *Small Group Research, 45*(2), 119–153. https://doi.org/10.1177/1046496414521303

Festinger, L. (1950). Informal social communication. *Psychological Review, 57*(5), 271–282. https://doi.org/10.1037/h0056932

Fuller, R. M., & Dennis, A. R. (2009). Does fit matter? The impact of task-technology fit and appropriation on team performance in repeated tasks. *Information Systems Research, 20*(1), 2–17. https://pubsonline.informs.org/doi/abs/10.1287/isre.1070.0167.

Ganesh, M. P., & Gupta, M. (2010). Impact of virtualness and task interdependence on extra-role performance in software development teams. *Team Performance Management: An International Journal, 16*(3/4), 169–186. https://doi.org/10.1108/13527591011053250

Gibbs, J. L., Sivunen, A., & Boyraz, M. (2017). Investigating the impacts of team type and design on virtual team processes. *Human Resource Management Review, 27*(4), 590–603. https://doi.org/10.1016/j.hrmr.2016.12.006

Gilson, L., Costa, P., O'Neill, T., & Maynard, T. (2021). Putting the TEAM back into virtual teams. *Organizational Dynamics, 50*(1), 100847. https://doi.org/10.1016/j.orgdyn.2021.100847

Gilstrap, C. A., Srivastava, S., & Gilstrap, C. M. (2022). Making sense of teamwork in mobile hybrid teams: A lexical analysis. *Team Performance Management, 28*(5/6), 382–397. https://doi.org/10.1108/TPM-11-2021-0080

Glisson, C., & James, L. R. (2002). The cross-level effects of culture and climate in human service teams. *Journal of Organizational Behavior, 23*(6), 767–794. https://doi.org/10.1002/job.162

Hambley, L. A., O'Neill, T. A., & Kline, T. J. (2007). Virtual team leadership: The effects of leadership style and communication medium on team interaction styles and outcomes. *Organizational Behavior and Human Decision Processes, 103*(1), 1–20. https://doi.org/10.1016/j.obhdp.2006.09.004

Handke, L., Klonek, F. E., Parker, S. K., & Kauffeld, S. (2020). Interactive effects of team virtuality and work design on team functioning. *Small Group Research, 51*(1), 3–47. https://doi.org/10.1177/1046496419863491

Handke, L., Costa, P. L., Klonek, F. E., O'Neill, T. A., & Parker, S. K. (2021). Team perceived virtuality: An emergent state perspective. *European Journal of Work and Organizational Psychology, 30*(5), 624–638. https://doi.org/10.1080/1359432X.2020.1806921

Hill, N. S., & Bartol, K. M. (2016). Empowering leadership and effective collaboration in geographically dispersed teams. *Personnel Psychology, 69*(1), 159–198. https://doi.org/10.1111/peps.12108

Hirst, G., Van Dick, R., & Van Knippenberg, D. (2009). A social identity perspective on leadership and employee creativity. *Journal of Organizational Behavior, 30*(7), 963–982. https://doi.org/10.1002/job.600

Ho, C., Basdogan, C., Slater, M., Durlach, N., & Srinivasan, M. A. (1998, June). An experiment on the influence of haptic communication on the sense of being together. In *Proceedings of the British telecom workshop on presence in shared virtual environments* (pp. 10–11).

Hoch, J. E., & Kozlowski, S. W. J. (2014). Leading virtual teams: Hierarchical leadership, structural supports, and shared team leadership. *Journal of Applied Psychology, 99*(3), 390–403. https://doi.org/10.1037/a0030264

Horvat, N., Becattini, N., & Škec, S. (2021). Use of information and communication technology tools in distributed product design student teams. *Proceedings of the Design Society, 1*, 3329–3338. https://doi.org/10.1017/pds.2021.594

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

IBM. (2020). *COVID-19 Is Significantly Altering U.S. Consumer Behavior and Plans Post-Crisis*. IBM Newsroom. https://newsroom.ibm.com/2020-05-01-IBM-Study-COVID-19-Is-Significantly-Altering-U-S-Consumer-Behavior-and-Plans-Post-Crisis.

Iometrics and Global Workplace Analytics (2020). *Global Work-from-Home Experience Survey*. https://globalworkplaceanalytics.com/global-work-from-home-experience-survey. Accessed 3 June 2024.

James, L. R. (1982). Aggregation bias in estimates of perceptual agreement. *Journal of Applied Psychology, 67*(2), 219–229. https://doi.org/10.1037/0021-9010.67.2.219

James, L. R., Demaree, R. G., & Wolf, G. (1984). Estimating within-group interrater reliability with and without response bias. *Journal of Applied Psychology, 69*(1), 85–98. https://doi.org/10.1037/0021-9010.69.1.85

Jehn, K. A. (1995). A Multimethod Examination of the Benefits and Detriments of Intragroup Conflict. *Administrative Science Quarterly, 40*(2), 256–282. https://doi.org/10.2307/2393638

Kirkman, B. L., & Mathieu, J. E. (2005). The dimensions and antecedents of team virtuality. *Journal of Management, 31*(5), 700–718. https://doi.org/10.1177/0149206305279113

Kirkman, B. L., Rosen, B., Tesluk, P. E., & Gibson, C. B. (2004). The impact of team empowerment on virtual team performance: The moderating role of face-to-face interaction. *Academy of Management Journal, 47*(2), 175-192. https://doi.org/10.5465/20159571

Klimoski, R., & Mohammed, S. (1994). Team mental model: Construct or metaphor? *Journal of Management, 20*(2), 403–437. https://doi.org/10.1016/0149-2063(94)90021-3

Kock, N. (2004). The psychobiological model: Towards a new theory of computer-mediated communication based on Darwinian evolution. *Organization Science, 15*(3), 327–348. https://doi.org/10.1287/orsc.1040.0071

Konradt, U., Schippers, M. C., Garbers, Y., & Steenfatt, C. (2015). Effects of guided reflexivity and team feedback on team performance improvement: The role of team regulatory processes and cognitive emergent states. *European Journal of Work and Organizational Psychology, 24*(5), 777–795. https://doi.org/10.1080/1359432X.2015.1005608

Korzenny, F. (1978). A Theory of Electronic Propinquity: Mediated Communication in Organizations. *Communication Research, 5*(1), 3–24. https://doi.org/10.1177/009365027800500101

Kruskal, J. B. (1964). Nonmetric multidimensional scaling: A numerical method. *Psychometrika, 29*(2), 115–129. https://doi.org/10.1007/BF02289694

Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology, 28*(4), 563–575. https://doi.org/10.1111/j.1744-6570.1975.tb01393.x

LeBreton, J. M., & Senter, J. L. (2008). Answers to 20 questions about interrater reliability and interrater agreement. *Organizational Research Methods, 11*(4), 815–852. https://doi.org/10.1177/1094428106296642

Leonardi, P. M. (2012). *Car crashes without cars: Lessons about simulation technology and organizational change from automotive design*. MIT Press.

Lewis, K. (2003). Measuring transactive memory systems in the field: Scale development and validation. *Journal of Applied Psychology, 88*(4), 587–604. https://doi.org/10.1037/0021-9010.88.4.587

Mair, P., Groenen, P. J. F., & de Leeuw, J. (2022). More on multidimensional scaling in {R}: smacof version 2. *Journal of Statistical Software, 102*(10), 1-47 10.18637/jss.v102.i10.

Marks, M. A., Mathieu, J. E., & Zaccaro, S. J. (2001). A temporally based framework and taxonomy of team processes. *Academy of Management Review, 26*(3), 356–376. https://doi.org/10.5465/amr.2001.4845785

Marlow, S. L., Lacerenza, C. N., Paoletti, J., Burke, C. S., & Salas, E. (2018). Does team communication represent a one-size-fits-all approach?: A meta-analysis of team communication and performance. *Organizational Behavior and Human Decision Processes, 144*, 145–170. https://doi.org/10.1016/j.obhdp.2017.08.001

Mathieu, J. E., Luciano, M. M., D'Innocenzo, L., Klock, E. A., & LePine, J. A. (2020). The development and construct validity of a team process survey measure. *Organizational Research Methods, 23*(3), 399–431. https://doi.org/10.1177/1094428119840801

McGrath, J. E., & Hollingshead, A. B. (1993). Putting the "group" back in group support systems: Some theoretical issues about dynamic processes in groups with technological enhancements. In L. M. Jessup & J. S. Valacich (Eds.), *Group support systems: New perspectives* (pp. 78–96). Macmillan.

Mesmer-Magnus, J. R., & DeChurch, L. A. (2009). Information sharing and team performance: A meta-analysis. *Journal of Applied Psychology, 94*(2), 535–546. https://doi.org/10.1037/a0013773

Mohammed, S., Klimoski, R., & Rentsch, J. R. (2000). The measurement of team mental models: We have no shared schema. *Organizational Research Methods, 3*(2), 123–165. https://doi.org/10.1177/109442810032001

Morgeson, F. P., DeRur, S., & Kram, E. P. (2010). Leadership in teams: A functional approach to understanding leadership structures and processes. *Journal of Management, 36*(1), 5–39. https://doi.org/10.1177/0149206309347376

Müller, R., & Antoni, C. H. (2020). Individual perceptions of shared mental models of information and communication technology (ICT) and virtual team coordination and performance – The moderating role of flexibility in ICT use. *Group Dynamics: Theory, Research, and Practice, 24*(3), 186. https://doi.org/10.1037/gdn0000130

Muthén, L. K., & Muthén, B. (1998–2017). *Mplus User's Guide*. Eighth Edition. Muthén & Muthén.

Nauman, S., Khan, A. M., & Ehsan, N. (2010). Patterns of empowerment and leadership style in project environment. *International Journal of Project Management, 28*(7), 638–649.

O'Leary, M. B., Wilson, J. M., & Metiu, A. (2014). Beyond being there. *MIS Quarterly, 38*(4), 1219–1244. https://doi.org/10.25300/MISQ/2014/38.4.13

O'Neill, T. A., Deacon, A., Gibbard, K., Larson, N., Hoffart, G., Smith, J., & Donia, B. L. M. (2018). Team dynamics feedback for postsecondary student learning teams. *Assessment & Evaluation in Higher Education, 43*(4), 571–585. https://doi.org/10.1080/02602938.2017.1380161

O'Neill, T. A., Pezer, L., Solis, L., Larson, N., Maynard, N., Dolphin, G. R., ... & Li, S. (2020). Team dynamics feedback for postsecondary student learning teams: introducing the "Bare CARE" assessment and report. *Assessment & Evaluation in Higher Education, 45*(8), 1121–1135. https://doi.org/10.1080/02602938.2020.1727412

Orlikowski, W. J. (1992). The duality of technology: Rethinking the concept of technology in organizations. *Organization Science, 3*(3), 398–427. https://doi.org/10.1287/orsc.3.3.398

Orlikowski, W. J. (2007). Sociomaterial practices: Exploring technology at work. *Organization Studies, 28*(9), 1435–1448. https://doi.org/10.1177/0170840607081138

Orlikowski, W. J., & Scott, S. V. (2008). 10 Sociomateriality: Challenging the separation of technology, work and organization. *Academy of Management Annals, 2*(1), 433–474. https://doi.org/10.1080/19416520802211644

Ozimek, A. (2020). *The Future of Remote Work*. Upwork.

Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on amazon mechanical turk. *Judgment and Decision Making, 5*(5), 411–419. https://doi.org/10.1017/S1930297500002205

Peer, E., Vosgerau, J., & Acquisti, A. (2014). Reputation as a sufficient condition for data quality on Amazon Mechanical Turk. *Behavior Research Methods, 46*, 1023–1031. https://doi.org/10.3758/s13428-013-0434-y

Purvanova, R. K., & Kenda, R. (2022). The impact of virtuality on team effectiveness in organizational and non-organizational teams: A meta-analysis. *Applied Psychology, 71*(3), 1082–1131. https://doi.org/10.1111/apps.12348

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review, 41*, 71–90. https://doi.org/10.1016/j.dr.2016.06.004

R Core Team (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/. Accessed 15 Nov 2023.

Raghuram, S., Hill, N. S., Gibbs, J. L., & Maruping, L. M. (2019). Virtual work: Bridging research clusters. *Academy of Management Annals, 13*(1), 308–341. https://doi.org/10.5465/annals.2017.0020

Rapp, A., Ahearne, M., Mathieu, J., & Rapp, T. (2010). Managing sales teams in a virtual environment. *International Journal of Research in Marketing, 27*(3), 213–224. https://doi.org/10.1016/j.ijresmar.2010.02.003

Rico, R., & Cohen, S. G. (2005). Effects of task interdependence and type of communication on performance in virtual teams. *Journal of Managerial Psychology, 20*(3/4), 261–274. https://doi.org/10.1108/02683940510589046

Rico, R., Sánchez-Manzanares, M., Gil, F., & Gibson, C. (2008). Team implicit coordination processes: A team knowledge–based approach. *Academy of Management Review, 33*(1), 163–184. https://doi.org/10.5465/amr.2008.27751276

Rönkkö, M., & Cho, E. (2022). An Updated Guideline for Assessing Discriminant Validity. *Organizational Research Methods, 25*(1), 6–14. https://doi.org/10.1177/1094428120968614

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*(2), 1–36. https://www.jstatsoft.org/article/view/v048i02.

Salanova, M., Llorens, S., Cifre, E., & Martínez, I. M. (2012). We need a hero! Toward a validation of the healthy and resilient organization (HERO) model. *Group & Organization Management, 37*(6), 785–822. https://doi.org/10.1177/1059601112470405

Sallnäs, E. L., Rassmus-Gröhn, K., & Sjöström, C. (2000). Supporting presence in collaborative environments by haptic force feedback. *ACM Transactions on Computer-Human Interaction (TOCHI), 7*(4), 461–476. https://doi.org/10.1145/365058.365086

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika, 66*(4), 507–514. https://doi.org/10.1007/BF02296192

Schaufeli, W. B., Bakker, A. B., & Salanova, M. (2006). The Measurement of Work Engagement With a Short Questionnaire: A Cross-National Study. *Educational and Psychological Measurement, 66*(4), 701–716. https://doi.org/10.1177/0013164405282471

Shaffer, J. A., DeGeest, D., & Li, A. (2016). Tackling the problem of construct proliferation: A guide to assessing the discriminant validity of conceptually related constructs. *Organizational Research Methods, 19*(1), 80–110. https://doi.org/10.1177/1094428115598239

Short, J., Williams, E., & Christie, B. (1976). *The social psychology of telecommunications*. Wiley.

Simon, A. F. (2006). Computer-mediated communication: Task performance and satisfaction. *The Journal of Social Psychology, 146*(3), 349–379. https://doi.org/10.3200/SOCP.146.3.349-379

Soga, L. R., Vogel, B., Graça, A. M., & Osei-Frimpong, K. (2021). Web 2.0-enabled team relationships: An actor-network perspective. *European Journal of Work and Organizational Psychology, 30*(5), 639–652. https://doi.org/10.1080/1359432X.2020.1847183

Spreitzer, G. M. (1995). Psychological empowerment in the workplace: Dimensions, measurement, and validation. *Academy of Management Journal, 38*(5), 1442–1465. https://doi.org/10.5465/256865

Sturrock, K., & Rocha, J. (2000). A multidimensional scaling stress evaluation table. *Field Methods, 12*(1), 49–60. https://doi.org/10.1177/1525822X0001200104

Tannenbaum, S., Mathieu, J., Levy, J., Watson, D., Maynard, T., Beard, R., ... & Blue, S. (2024). The development and validation of an Army team resilience measure. *Military Psychology, 36*(1), 83–95. https://doi.org/10.1080/08995605.2022.2065154

Theodorou, A., Spano, G., Bratman, G. N., Monneron, K., Sanesi, G., Carrus, G., ... & Panno, A. (2023). Emotion regulation and virtual nature: cognitive reappraisal as an individual-level moderator for impacts on subjective vitality. *Scientific Reports, 13*(1), 5028. https://doi.org/10.1038/s41598-023-30287-7

Van der Vegt, G. S., & Bunderson, J. S. (2005). Learning and performance in multidisciplinary teams: The importance of collective team identification. *Academy of Management Journal, 48*(3), 532–547. https://doi.org/10.5465/AMJ.2005.17407918

Van der Kleij, R., Schraagen, J. M., Werkhoven, P., & De Dreu, C. K. W. (2009). How conversations change over time in face-to-face and video-mediated communication. *Small Group Research, 40*(4), 355–381. https://doi.org/10.1177/1046496409333724

Wageman, R., Hackman, J. R., & Lehman, E. (2005). Team diagnostic survey: Development of an instrument. *The Journal of Applied Behavioral Science, 41*(4), 373–398. https://doi.org/10.1177/0021886305281984

Walther, J. B. (1992). Interpersonal effects in computer-mediated interaction: A relational perspective. *Communication Research, 19*(1), 52–90. https://doi.org/10.1177/009365092019001003

Walther, J. B. (1994). Anticipated ongoing interaction versus channel effects on relational communication in computer-mediated interaction. *Human Communication Research, 20*(4), 473–501. https://doi.org/10.1111/j.1468-2958.1994.tb00332.x

Walther, J. B., & Bazarova, N. N. (2008). Validation and application of electronic propinquity theory to computer-mediated communication in groups. *Communication Research, 35*(5), 622–645. https://doi.org/10.1177/0093650208321783

Walther, J. B., & Parks, M. R. (2002). Cues filtered out, cues filtered. In: Computer mediated communication and relationships. In M. L. Knapp & J. A. Daly (Eds.), *Handbook of interpersonal communication* (3rd ed., pp. 529–561). Sage.

Walther, J. B., & Tidwell, L. C. (1995). Nonverbal cues in computer-mediated communication, and the effect of chronemics on relational communication. *Journal of Organizational Computing and Electronic Commerce, 5*(4), 355–378. https://doi.org/10.1080/10919399509540258

Watson-Manheim, M. B., Chudoba, K. M., & Crowston, K. (2002). Discontinuities and continuities: A new way to understand virtual work. *Information Technology & People, 15*(3), 191–209. https://doi.org/10.1108/09593840210444746

Watson-Manheim, M. B., Chudoba, K. M., & Crowston, K. (2012). Perceived discontinuities and constructed continuities in virtual work. *Information Systems Journal, 22*(1), 29–52. https://doi.org/10.1111/j.1365-2575.2011.00371.x

Wegge, J., Vogt, J., & Wecking, C. (2007). Customer-induced stress in call centre work: A comparison of audio-and videoconference. *Journal of Occupational and Organizational Psychology, 80*(4), 693–712. https://doi.org/10.1348/096317906X164927

Wellen, J. M., & Neale, M. (2006). Deviance, self-typicality, and group cohesion: The corrosive effects of the bad apples on the barrel. *Small Group Research, 37*(2), 165–186. https://doi.org/10.1177/1046496406286420

Wheeler, A. R., Shanine, K. K., Leon, M. R., & Whitman, M. V. (2014). Student-recruited samples in organizational research: A review, analysis, and guidelines for future research. *Journal of Occupational and Organizational Psychology, 87*(1), 1–26. https://doi.org/10.1111/joop.12042

Wilson, J. M., Boyer O'Leary, M., Metiu, A., & Jett, Q. R. (2008). Perceived proximity in virtual work: Explaining the paradox of far-but-close. *Organization Studies, 29*(7), 979–1002. https://doi.org/10.1177/0170840607083105

Wilson, F. R., Pan, W., & Schumsky, D. A. (2012). Recalculation of the critical values for Lawshe's content validity ratio. *Measurement and Evaluation in Counseling and Development, 45*(3), 197–210. https://doi.org/10.1177/0748175612440286

Woehr, D. J., Loignon, A. C., Schmidt, P. B., Loughry, M. L., & Ohland, M. W. (2015). Justifying Aggregation With Consensus-Based Constructs: A Review and Examination of Cutoff Values for Common Aggregation Indices. *Organizational Research Methods, 18*(4), 704–737. https://doi.org/10.1177/1094428115582090

Yammarino, F. J., & Dansereau, F. (2011). Multilevel issues in organizational culture and climate research. In N. M. Ashkanasy, C. P. M. Wilderom, & M. F. Peterson (Eds.), *The handbook of organizational culture and climate* (2nd ed., pp. 50–76). https://doi.org/10.4135/9781483307961

Zhao, S. (2003). Toward a taxonomy of copresence. *Presence: Teleoperators and Virtual Environments, 12*(5), 445–455. https://doi.org/10.1162/105474603322761261