# DEVELOPMENT AND VALIDATION OF THE TEAM PERCEIVED VIRTUALITY SCALE

LISA HANDKE
Friedrich-Alexander-Universität Erlangen-Nürnberg, Lange Gasse 20, 90429 Nuremberg, Germany

PATRÍCIA COSTA
Iscte – IUL, Portugal

JENNIFER FEITOSA
Claremont McKenna College, CA, USA

## INTRODUCTION

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, drawing 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.

Accordingly, 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).

In this work, we develop a scale that captures the construct of Handke et al.'s (2021) TPV dimensions amd validate this scale in five studies. First, we describe the development of multi-item scales for the two 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 generalize our results from individual-level full-time workers and team-level student semester projects to an organizational sample (Study 5). Taken together, we put forth a newly developed 10-item measure of TPV that is conceptually distinct to structural virtuality, allowing for a more dynamic and subjective approach to virtuality (see Table 1 for a description of the final items). We conclude with recommendations for future research and application of the TPV measure.

## STUDY 1: ITEM GENERATION

Two of the authors constructed a pool of 20 items (10 per dimension) to capture the two TPV dimensions. Items were reviewed by a panel of 11 well-known virtual team experts. The Content Validity Ratio (CVR, Lawshe, 1975; Wilson et al., 2012) was used to measure the raters' agreement. CVR values ranged from 1 to .27 considering the relevant ratings: essential and useful. Using the critical value of .78 (for 11 panelists and $p$ = 0.01, Lawshe, 1975), six items from each dimension could be retained, therefore leading to a list of 12 items.

## STUDY 2: STRUCTURAL VALIDITY (INDIVIDUAL-LEVEL)

We recruited 447 employees working in the United States via MTurk to participate in an online survey. Participants rated all 12 TPV items on a 7-point response scale from 1 *(strongly disagree)* to 7 *(strongly agree)*. We conducted a CFA with the MLR estimator using Mplus version 8.4 (Muthén & Muthén, 1998 – 2017), with the 12 items loading onto two factors (i.e., collectively-experienced distance and information deficits). The two-factor solution exhibited a very good model fit ($\chi^2$ = 93.21, df = 53, p < .001, CFI = .981, RMSEA = .044, SRMR = .032), which was superior to the one-factor model ($\chi^2$ = 271.47 , df = 54, $p$ < .001, CFI = .897, RMSEA = .101, SRMR = .048), as shown by a significant Satorra-Bentler scaled (Satorra & Bentler, 2001) $\chi^2$ difference ($\Delta\chi^2$ = 49.71, $\Delta$df = 1 $p$ < .001). However, factor loadings of the two reverse-coded items 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, $p$ < .001, CFI = .988, RMSEA = .035, SRMR = .027; Satorra-Bentler scaled $\Delta\chi^2$ = 16.60, $\Delta$df = 2, $p$ < .001). We therefore decided to omit the two reverse-coded items from further analyses. Accordingly, the final TPV scale consisted of 10 items (cf. table 1), loading onto two factors and showing an excellent model fit ($\chi^2$ = 41.60, df = 34, $p$ = .174, CFI = .996, RMSEA = .024, SRMR = .017).

------------------------
Table 1 about here
------------------------

## 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, pulled from the ITPmetrics.com database[1]. 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). Participants rated all 10 TPV items on a 7-point response scale ranging from 1 *(strongly disagree)* to 7 *(strongly agree*). We calculated intraclass correlations (ICCs) for the two TPV dimension and the rwg(j) (collective-experienced distance: ICC(1) = .25, ICC(2) = .56, median $r_{wg(j)uniform}$ = .96, median $r_{wg(j)skewed}$ = .94, $\alpha$ = 97.; collectively-experienced information deficits: ICC(1) = .13, ICC(2) = .38, median $r_{wg\ (j)uniform}$ = .93, median).

The MCFA showed an excellent model fit ($\chi^2$ = 772.50, *df* = 68, $p$ < .001, CFI = .970, RMSEA = .049, $SRMR_{within}$ = .031, $SRMR_{between}$ = .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* < .001, CFI = .827, RMSEA = .117, $SRMR_{within}$ = .129, $SRMR_{between}$= 1.137), as shown by a significant Satorra-Bentler scaled (Satorra & Bentler, 2001) $\chi^2$ difference ($\Delta\chi^2$ = 1,110.64, $\Delta df$ = 2, *p* < .001).

## STUDY 4: CONCEPTUAL AND EMPIRICAL DISTINCTIVENESS & CRITERION VALIDITY

### Sample 4a

We recruited 124 employees working in the United States via Prolific to participate in an online survey. 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). 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 dimension of experienced information deficits ($p_{sa}$= .72; $c_{sv}$= .45) differs from effectiveness of the use of technology for virtual communication, information sharing, and team coordination items while experienced distance ($p_{sa}$= .87; $c_{sv}$= .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

Both subsamples were pulled from the ITPmetrics.com database (but did not overlap with the sample used in Study 3). 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). The 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).

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 and then compared this model to four constrained models in which the respective factor correlations had been fixed at -.85. Analyses were carried out at the team level. The unconstrained model yielded an excellent model fit ($\chi^2$ = 253.89, df = 98, p < .001, CFI = .98, RMSEA = .050, SRMR= .023). Factor correlations for the latent variables were: -.69 (distance with team monitoring and backup, 95% CI [-.60; -.78]), -.71 (distance with coordination, 95% CI [-.63; -.80]), -.69 (information deficits with team monitoring and backup, 95% CI [-.58; -.79]), and -.69 (information deficits with coordination, 95% CI [-.59; -.79]). Model comparisons showed significant Satorra-Bentler scaled $\chi^2$ 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* <.001; fixed correlation distance-coordination: $\Delta\chi^2$ 38.01, $\Delta df$ = 1, *p* <.001; fixed correlation information deficits-team monitoring and backup: $\Delta\chi^2$ 50.79, $\Delta df$ = 1, *p* <.001; fixed correlation information deficits-coordination: $\Delta\chi^2$ 66.78, $\Delta df$ = 1, *p* <.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 support the assumed negative relationship between collectively-experienced distance and leader-rated team interaction quality ($r$ = -.38, $p$ <. 001) as well as between collectively-experienced deficits and leader-rated team performance ($r$ = -.31, $p$ <.001). 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

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, resulting 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.

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 justified aggregation to the team-level as justified. The MCFA showed an excellent model fit ($\chi^2$ = 300.80, $df$ = 68, $p$ < .001, CFI = .966, RMSEA = .057, $SRMR_{within}$ = .035, $SRMR_{between}$ = .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 ($\chi^2$ = 1687.83, $df$ = 70, $p$ < .001, CFI = .761, RMSEA = .147, $SRMR_{within}$ = .098, $SRMR_{between}$= .738), as shown by a significant Satorra-Bentler scaled (Satorra & Bentler, 2001) $\chi^2$ difference ($\Delta\chi^2$ = 5,591.57, $\Delta df$ = 2, $p$ < .001We 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 non-significant (structural virtuality – distance: $r$ = -.05, $p$ = .553; structural virtuality – information deficits: $r$ = -.15, $p$ =.051).

## DISCUSSION

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 distinctiveness of the two TPV dimensions, the team-level nature of this construct, and its relationship with team outcomes.

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., 2009) 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. Third and finally, acknowledging the importance of subjective virtuality experiences calls attention to the eventual unintended consequences of technology usage (Soga et al., 2020) 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.

Parallel to the contributions highlighted above, the present work has some limitations that constrain the generalizability of its findings. While we show different types of validation studies across samples and contexts, the specific influence of technology has not yet been tested. 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. 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. Furthermore, 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.

## ENDNOTES

1. ITPmetrics.com offers a range of 100% freely accessible assessments, including the instantaneous creation of user-friendly feedback reports. The assessment platform currently has over 600,000 assessments taken. For more information, see e.g., O'Neill et al. (2017; 2020)

## REFERENCES AVAILABLE FROM THE AUTHORS

**TABLE 1**

**Final TPV scale**

| | |
|---|---|
| Collectively-experienced distance | |
| 1 | In my team, we feel detached from each other |
| 2 | In my team, we feel that our relationship is cold |
| 3 | In my team, we feel like we are far away from each other |
| 4 | In my team, we feel estranged from each other |
| 5 | In my team, we feel like we cannot get through to each other |
| Collectively-experienced information deficits | |
| | When we exchange information in my team … |
| 1 | ...the ways in which we can express ourselves are limited. |
| 2 | ...it's hard to convey the actual meaning of what we are saying |
| 3 | ...it's difficult to understand if we are on the same page or not |
| 4 | ...we are unable to convey the necessary information in its entirety |
| 5 | ...we don't know whether everyone has had access to the same information |