Papers on Anthropology XXVII/2, 2018, pp. 57–72

RELIABILITY AND VALIDITY OF ANALYTICAL CONSTRUCT IN THE CASE OF VIRTUAL WORK

Kandela Õun

Pärnu College, University of Tartu, Estonia

ABSTRACT

Social scientists are interested in understanding social reality as accurately as possible, trying to measure observable and indirectly measurable phenomena as well as the phenomena that still need to be constructed for measurement. Constructing social phenomena so that they can be measured is the everyday work of social scientists as the development of societies constantly brings forth new phenomena that attract authors' interest to investigate and interpret them. Technology-enhanced telework as a research theme has become topical because of advances in information and communication technology and the fact that they have become widespread. The biggest obstacle to the investigation of technology-enhanced telework (also called *virtual work*) is the huge range of different definitions and approaches; virtual work mostly described rather than measured in surveys.

The current article focuses on the reliability and validity of the analytical construct of virtual work previously developed by the author on the basis of service sector organisations.

Reliability is quite simple to measure – Chronbach's alpha is well known to measure it, but there are several approaches to the measurement validity beginning with the first authors Cronbach and Meehl [14] to Messick [32], Kane [29], Bornstein [7] and Slocum et al [48]. Every researcher adds a new point of view that makes understanding and using the methods of validity more complicated, but it is important to find a proper and suitable way to analyse the validity of construct of virtual work. In the article, the author will analyse the reliability and validity of two models which measure technology-enhanced telework.

Keywords: *reliability; validity; virtual work; technology-enhanced; analytical construct*

https://doi.org/10.12697/poa.2018.27.2.07

INTRODUCTION

Based on theory, a theoretical construct is created in the process of conceptualisation, followed by a measuring instrument which helps to compile an analytical construct, measuring a phenomenon by combining answers to different questions (e.g. compiling indexes) [52]. Creating new social-scientific knowledge brings up conceptualisation and operationalisation of the matter of research as an important question and evaluating the reliability of the instrument that has been created [46]. The results of conceptualisation and operationalisation depend to a great extent, on the one hand, on how well the term has been defined, how well it has been given sense theoretically and published as an instrument and, on the other hand, on what was actually measured. According to Baert [4 p. 6], social-scientific research constantly strives to give the reality as accurate a sense as possible; in parallel with creating new theories, it is important whether they can be refuted (falsification) using, for example, empirical experiments [40]. Technically, the process of measurement or operationalisation involves "rules of assigning numbers to objects to represent quantities of attributes" [37]. By Fawcett and Garity [20], the conceptual (C) component of a conceptual-theoretical-empirical (CTE) structure is the conceptual model that is selected to guide the research or practice. Researchers spend a great deal of time searching for instruments, including interview guides and questionnaires, that are appropriate measures of the middle-range theory concepts that they plan to study [20, p.189].

There are several terms and constructs in social sciences which cannot be measured directly for different reasons. For example, such latent characteristics as alienation, lack of social regulation, social neglect, intelligence, quality of life, cost of living, sovereignty, and business secrets, are characteristics the essence of which is known, but it is complicated to measure them because there is no specific scale or question. Since the time of Spearman [49] and Thurstone [53], a common factor has been considered to be an unmeasured variable whose variation (a) generates variation in, and (b) explains the correlations or covariances among two or more dependent variables that it predicts [35]. According to Cronbach and Meehl [14, p 283], "A construct is some postulated attribute of people, assumed to be reflected in test performance." Conceptually, unmeasured variables that operationalise a construct in a "formative" measurement model also are latent variables under appropriate conditions [54]. Tarling [52, p 125–127] points out three main methods – proxy variables, indices, and factor analysis. These methods use as original characteristics latent manifest variables that form the basis of the study as well as characteristics that enable

measurement. Factor analysis and, more recently, structural equations are frequently used for this purpose. Covariance-based (e.g., techniques implemented in statistical packages such as LISREL, Amos, EQS, etc) and component-based (e.g., PLS) SEM allow researchers to simultaneously examine measurement and structural models [24], yet researchers tend to focus on the structural model rather than fully consider the relationship between measures and their relevant latent constructs [27]. Treiblmaier *et al* [54] discussed the usage of formative constructs and simple factors in models and found that formative constructs are difficult to identify and thus not easy to be evaluated and tested, so they proposed a two-step approach to the operationalisation of the construct being measured formatively. They concluded that even though a formative construct can have a scientific meaning conceptually, in practice the reflective measurement approach that is well considered is still the best way of applying the formative approach [54]. None of these methods is without problems. Thus, researchers are constantly looking for possible new approaches.

Measurements, which in social sciences often include getting answers from people to questions or verbal expressions [50, p 148], are more specific than constructs; they enable the gathering of data for developing analytical constructs, besides there is a rich choice between instruments, which in turn also includes several approaches. Cicourel [11] claims that the data of interviews and queries have been gained in the context of wider interactive influences, which includes both cognitive and linguistic aspects in the complex of existing and developing socio-cultural trends. Thus, the questions in questionnaires are usually in the verbal form, and such indicators as: intelligence, attitude towards survey and reading capability influence the answers. Quantitative research emphasizes quantification on the collection and analysis of data that entails a deductive approach focusing on the testing of theories. Though quantitative research analyses and tests relationships between different phenomena, the meaning of the connections studied may remain undiscovered [43].

RELIABILITY AND VALIDITY OF MEASUREMENTS

No measurement system is perfect; there is always some slippage (and sometimes a lot) between the complex construct that we are interested in and the results of measurement. We can talk about two sources of error in any measurement: (1) construct under-representation and (2) construct-irrelevant variance [12, 32, 48]. Scientific approach requires that all results of studies are reliable and valid [3, 18, 47]. In the practice of measuring constructs, an analytical

construct or measurement corresponding to a theoretical construct is usually created. According to Groves [26, p 261], reliability is the measure of variability of respondent's answers when comparing several measurement results. Sarantakos [45, p 88] adds that reliability is the indicator of objectivity, stability, integrity, and exactness, which measures the quality of indicators and instruments. The term reliability stems from classical testing theory and refers to the correlation between two or more measurement trials of the same object, using measurements which are as similar as possible [30]. One way of checking the reliability of study results is to measure twice, which creates the need for conducting of pilot studies [46, p 16]. Reliability refers to replicable studies; comparing the results gained by several authors is called inter-rater reliability. According to Sarantakos [45, p 88–89], there are at least three different types of reliability in social sciences: stability, representativeness, and reliability of equivalency, which examine the use of measurements in different time periods, between different groups or individuals, and different equal indicators or instruments during the process of operationalisation. An indicator of the reliability of wide-spread analytical constructs (e.g. indexes created on the basis of several characters or factors compiled in the process of analysis) is the reliability coefficient Cronbach's α [13]. When Cronbach's Alpha Coefficient is less than 0.40, it is unreliable; between 0.40–0.56, the test is said to have low reliability; between 0.60–0.79, the test is fairly reliable, and between 0.80–1.00, it is highly reliable [5, 38]. One among the used methods is the test-retest method where the same indicator or index measures the same respondent twice, and the same result shows the great reliability of the measurement. Other methods are the split-half method where the questions in the questionnaire are divided into two groups (e.g. paired and not paired questions) and the results' correlation is found; applying an alternative indicator to the same data and finding the correlation between the results gained. [45, p 88]

A major aspect of every search for research instruments and practice tools is determining measurement validity, which refers to the appropriateness of the instruments and tools, which in turn refers to what middle-range theory concept the research instrument or practice tool *really* measures [20]. A largely neglected aspect of measurement validity is determining whether the definition of the middle-range theory concept measured by the research instrument or practice tool is congruent with the focus of the conceptual model that was selected to guide a study or practice. In addition to reliability, validity, which studies the connection (correlation) between the measured values and the theoretical construct (e.g. intelligence, poverty etc.) is also an indicator of

the good analytical construct [39]. The term validity was first published by Cronbach and Meehl [14], and four years later similar thoughts were published by Campell and Fiske [9]. Cronbach and Meehl [15] came out with four types of validity (predictive, concurrent, content and construct validity) and different approaches how to measure those. They suggested that the first two validity types should be analysed by investigating group differences, correlation matrices and factor analysis. If two measures are presumed to measure the same construct, a correlation between them is predicted. If the obtained correlation departs from the expectation, however, there is no way to know whether the fault lies in test A, test B, or the formulation of the construct. A matrix of intercorrelations often points out profitable ways of dividing the construct into more meaningful parts, factor analysis being a useful computational method in such studies.

Bornstein [7] created a new approach which he called the process-focused (PF) model. This model conceptualises validity as the rate in case respondents go through psychological processes during testing; those processes are caused directly by the nature of the instruments used in measuring and the context of testing. The new aspect of this model is using the experimental method, manipulation with characters, which could influence the result of an analytical construct and its connections with the criteria that show how good it is. It enables the author to draw clearer conclusions about psychological processes being measured.

By Slocum et al [48], modern definitions of validity assert that measurement validity is a single thing (i.e., a unitary construct) rather than a collection of separate types of validity (e.g., content validity, criterion validity, etc.). This view treats all evidence about validity as evidence of how well the results produced by the measure reflect the construct. One of the first tasks in thinking about validity is to clearly define and analyse the construct to be measured [1, 29]. The structural aspect of validity asks questions about the ways in which scores on individual items are combined into component scores, and how these are combined into broader scores. That means, validity is based on both empirical findings and logical analysis [48].

TECHNOLOGY-ENHANCED TELEWORK – VIRTUAL WORK

Gilson et al [25, p 1317] found that "a first step is to derive a more unified measurement and treatment of the construct, which would allow for comparisons to be made across studies. Accordingly, in addition to more carefully considering how team virtuality is measured."

Herein the concept of *virtual work* (more commonly termed telework) is defined as the working environment where completing assignments of work is implemented by using different communication channels mediated by technology, and where separation from co-workers may be the issue. Virtual work as a research theme has become topical because of advances in information and communications technology, and the fact that they have become widespread. The first articles dealing with the new form of work and work virtuality could be found from the Emerald database since 1984. Technology-enhanced telework and its opportunities have been studied in Estonia since 2000. The Prime Minister's office commissioned a study to map whether it was possible to introduce telework into public administration [51] and the impact of telework on employment [2]. A number of bachelor's and master's theses have been written, for example, on the feasibility of telework in local authorities [41], employment of physically disabled people through telework [36] and telework as a means of motivating employees [42]. In addition, a survey was conducted in Estonian enterprises on telework as a flexible opportunity to combine work and family life [55], and there is a research paper relating to the use of information and communication technology in client communication rather than employees communicating with each other in the true sense of teleworking [17].

The biggest obstacle in work virtuality investigation is the huge range of different definitions and approaches, and it is mostly described rather than measured in surveys. In conclusion, work virtuality as a construct and the research into creating an instrument for measuring it are largely absent. Attempts to measure virtual teamwork have been made by several researchers, but few [10, 23] have actually offered a model to do it. The author of this paper elaborates on the process of developing the construct of work virtuality and creating the corresponding instrument building on the communication richness theory [15, 16, 21].

DATA AND METHOD

The empirical database for this article is the questionnaire-based research that was conducted from February to June in 2006 and 2007[1]. The questionnaire takes into account, among other questions, also the respondents' answers about using communication channels and their impact on job satisfaction.

[1] This research was supported by the Estonian Ministry of Education target fund SF0180037s08 and the Estonian Science Foundation Grant 7018.

The answers were collected from 3156 respondents (see also Table 1) from 323 different service sector enterprises, including public sector organisations; in 2006 the number of respondents was 2207 and in 2007 – 949. The sample was random, and every county in Estonia was covered. The service sector companies were chosen for study, as it is more common to use virtual work in the service sector than in production.

Table 1. Sample characteristics.

Category		Number	Category		Number
Organisations		323	Male respondents		958
Respondents		3156	Female respondents		2149
Average work tenure (years)	less than 3	1218	Highest level of education achieved	Tertiary education	1385
	3–5	479		Vocational education	950
	6–9	432		Secondary education	710
	10–20	413		Elementary education	67
	more than 20	131		Unanswered	44

During the operationalisation process, two measurements (also called indices) were calculated. The first index (A) of the construct "*Virtuality*" uses a second-order model method developed by John [28], which in principle means that first-order constructs are formed as *richness*, *frequency* and *time* (also operationalised), and they serve as a basis for the second-order construct *virtuality*. Because there are three first-order constructs, the developed analytical construct is handled as a three-dimensional model [33], and each first-order construct contributes to the value of the second-order construct. To calculate the second-order construct ***Virtuality***, the three first-order constructs are summed, and the three-dimensional model is derived as a sum index. The development procedure was quite long and includes different re-coding techniques, so there might be reliability and validity problems. The need for an improved construct (second index B) originated from the analysis [34] where the problems respondents experienced were not as highly related to degrees of virtuality as reported in literature. The fact that large statistical differences did not emerge during the group comparisons, raised doubts as to whether the correct basis was used in the operationalisation of the constructs and, in particular, the process of bringing together the results of the two different questionnaires seemed to require improvement. The reason might have been value laden data or the operationalisation procedure, and it was considered necessary to change the operational process to calculate a new virtuality index.

The empirical results of the second index (B) are based on a sample of 781 respondents from 93 service sector organisations, and the 2007 database only was used in order to eliminate the differences between the two databases aroused by improving the questionnaire. Index B was calculated using a different methodology [56] in which the three first-order constructs (significance, frequency and time) for each of the 8 channels of communication were multiplied and the results merged into one index using the same equation which takes into account "richness" [16]. Another aspect is that index B aims to identify the differences of virtual work in the case of easy (EW) and hard work (HW) tasks, which were differentiated according to the respondents' subjective opinions in the improved questionnaire. Another essential result is that the new index is more convenient to use since its calculation methodology is considerably more user-friendly and, because of the quick development of ICT, this model allows to change communication channels not used so often any more for newer ones. Now, there are two indices of work virtuality, and there is a need to analyse the reliability and validity of both of them.

RESULTS

Social scientists attempt to ensure that the research process is as error-free as possible, beginning with how constructs are measured. Researchers look for consistency of measurement over time, in relationship with other related measurements, or in measurements or observations made by two or more researchers [22, p 16]. According to Lyman [31] a test's reliability could be influenced by five sources of error: content of the test, its persistence over the time, the person who carries it out, respondents, and the situation where it is carried out, the first three of which are the responsibility of the person undertaking the test. Reliability is a prerequisite for validity: we cannot measure a phenomenon if the measure we are using yields inconsistent results. A measure may not be valid due to individual errors (individuals may want to provide socially desirable responses) or method errors (questions may be unclear or poorly written) [22, p 16].

Two analytical constructs *A* and *B* were described above, and the principles by which they were created can be found in Figure 1.

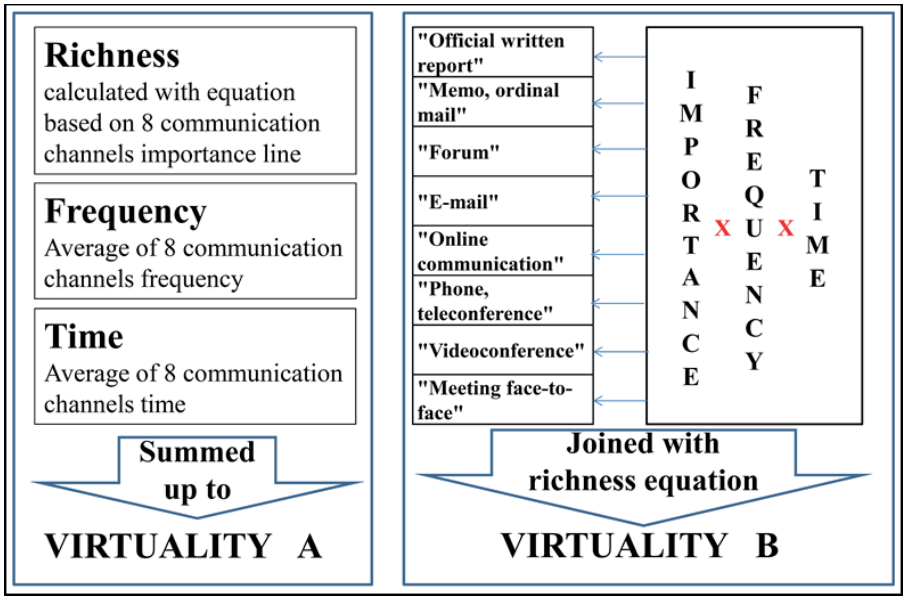

Figure 1. Two approaches to the structure of a virtual construct

To test the reliability of two analytical constructs of technology-enhanced telework, the values of analytical construct *A* were found based on the data of the second analytical construct to avoid the effects of combinations of data being gathered in two different years on the value of the constructs. In creating construct *B* in an improved questionnaire, easy and hard work were taken into account as characteristics. Following these findings, the comparison of constructs *A* and *B* is presented as two different indexes which describe work virtuality in fulfilling easy and hard work tasks. To find the reliability of analytical construct *A*, the sub-constructs *richness*, *time*, and *frequency* were calculated as well the Cronbach's alphas of the original characteristics of construct *B*, which are presented in Table 2.

The reliability values allow assumptions that shortcomings in the construct development still exist. The value of Cronbach's alpha of reliability must be higher than 0.8 – between 0.40–0.79, the test is said to have low or fair reliability [5, 38]. A more detailed examination shows that the lowest reliability is the richness of construct *A* and "meeting face-to-face" of construct *B*. An examination was also made whether any original characteristics could be excluded in order to increase the reliability of richness. It was discovered that if "meeting face-to-face" were excluded, this would increase the reliability in the case of easy work up to 0.39 and hard work to 0.37. In the case of frequency and

time of the constructs, there would be no improvement in terms of reliability if any of the original characteristics were excluded. Second-order construct *A* reliability remains a little under the minimum level. Construct *B* first-order constructs had better reliability values of original characteristics than construct *A* first-order constructs, which were higher than the minimum level of 0.7 in more than half of the cases. When calculating the reliability values for second-order constructs using first-order constructs, the results were lower than for construct *A*. However, finding the reliability based on original characteristics (24 all together), the results were higher than for construct *A*. Hence the question, what is the cause of this difference? Is it due to the fact that construct *B* was calculated using the richness model's first-order constructs?

Table 2. The reliability values of analytical constructs of A and B

Cronbach's alpha		Easy work		Hard work	
Second-order construct	First-order construct	First-order construct's alpha	Second-order construct's alpha	First-order construct's alpha	Second-order construct's alpha
A	Richness	0,31	0,61	0,31	0,65
	Frequency	0,55		0,55	
	Time	0,56		0,58	
B	"Official written report"	0,63	0,66	0,63	0,68
	"Memo, ordinary mail"	0,7		0,71	
	"Forum"	0,7		0,71	
	"E-mail"	0,75		0,76	
	"Online communication"	0,76		0,76	
	"Phone, teleconference"	0,52		0,59	
	"Videoconference"	0,67		0,65	
	"Meeting face-to-face"	0,41		0,41	

Following the above, Pearson's correlation coefficient between constructs *A* and *B* was calculated. In the case of easy work, the correlation between the two constructs was 0.495 ($p < 0.001$) and hard work 0,541 ($p < 0.001$), but since the variations of the constructs are different, then this could be a reason for such medium values. The correlation between easy and hard work was 0.82 for construct *A* ($p < 0.001$) and 0.77 for construct *B* ($p < 0.001$). In conclusion, there are no significant differences in reliability between those two constructs, and the differences in correlation are mostly due to the differences in scales.

Practical application of the index can be differentiated, and both indexes have their own theoretical explanations as well.

The validity of the two analytical constructs of technology-enhanced telework is difficult to measure, but, in the operationalisation process described shortly in this paper, a clear logical reasoning is shown. According to Slocum [48], this is also important for construct validity. In this operationalisation process also Chronbach and Meehl's [14] suggestions to study group differences, correlation matrices and factor analysis were followed, and these results are published in [34] and [56].

Based on construct *A*, technology-enhanced telework intensity can be described and classified in a similar manner to that described in the literature, but there are still some differences stemming from the process of construct creation. Comparing the values of the construct, one can conclude that virtual jobs are not wide-spread in those enterprises that were examined, and thus the results cannot be compared to a real construct of reality. The operationalisation process was complicated, but the values of necessary sub-constructs can be calculated, and the usage of virtual work in enterprises can be evaluated.

In creating construct *B*, the situation caused by the data was also considered, which had not been paid attention to at the beginning because the idea to create the construct originally arose from the literature. The respondents had clear preferences, they used some means more often, and those that were not used so often could be considered to be removed from the index. In their 2003, study Chudoba *et al* [10, p 12] also mentioned that, in the case of instant messaging (MSN, Skype) and video conferences as less used means, the reliability of factor analysis was too low (alpha = 0.42) and was left out of the study. These communication channels (video conferences, using forums, instant messaging) were not left out, as information technology means have developed since the original study was carried out, and, at the time of this paper, they are being used more and more often. The idea to deal with work virtuality based on communication means is important in the future, since it enables adaptation of the calculation method by adding or removing communication means.

THE GOODNESS OF THE MODEL

In social sciences, it is essential not only to describe connections but also to study their causality. There must be a connection between the characteristics studied; to use the term causal connections, the cause must appear before the consequence, and the connection between the characteristics should not

disappear, taking into account the effect of other characteristics as well [52, p 111]. Using confirmatory factor analysis, the path way model and the structural equation model facilitates such an approach. According to Tarling [52, p 154], descriptive factor analysis enables one to study a number of factors and their content based on measured characteristics; in the case of confirmative factor analysis, the author decides which measured characteristics are correlated to corresponding latent phenomena. The structural equation models method has become increasingly popular recently; it is a combination of statistical techniques, such as exploratory factor analysis and multiple regression, and it describes causal connections between latent characters, and its purpose is to chart those connections [6, p 5], which is often the most interesting result for an author in social sciences. The results gained are usually presented as a scheme which also provides a visual overview of the connections of which there are often many.

A construct (model) is created in social sciences based on values that could be found in literature or society, but often there is no opportunity to compare it with a real analytical construct. In social sciences, there is a tendency to create norms, but this depends on too many factors that exist in a specific society, and, thus, it is not possible to conduct cross-national comparisons.

As a result of second operationalisation, an analytical construct was created, which is more comfortable to use but may not be final, as the analysis of reliability showed that reliability is there but not good enough to be considered excellent. So, the analytical construct of technology-enhanced telework is not ready, and it seems that it takes some further improvement to be closer to the theoretical construct. The SEM analysis with AMOS showed that there is a need for new data because existing data are not gathered for this kind of analysis, and model specification in SEM needs normally distributed data. The existing database includes too many zero values and that is not helpful either.

CONCLUSIONS

Since the beginning of sociology there has been constant debate over whether it is a science or not, which features make it a science etc. For years researchers have been replacing simple analytical methods with more and more complicated ones – natural scientists and technology scientists have followed a similar course of action. During the past decade it was rare to find a scientific article published in an internationally recognised journal and based on questionnaire data using a quantitative research method, which not include

multi-dimensional data analysis. The most frequently used methods are multinomial regression and logistic regression analysis, which enable social sciences to emulate the research traditions of the so-called real sciences. In the case of sociological data, one pre-requisite of regression analysis which requires the persistence and normal distribution of characteristics is frequently not met. Respondents' attitudes and opinions are mostly presented on the Likert scale in which case a distribution similar to normal distribution can be seen if there are large quantities of data.

According to Bradley and Schafer [8, p 147], social scientists are often unable to replicate the results of their studies using the same conditions, and thus they have to model the "unsameness" of events, adding the existence of possible accidental "white noise". Studying social phenomena and not being able to measure them directly because they are frequently latent, the analytical model can only partially explain the phenomena that are studied. In addition, the data represent subjective reality of the data source (respondents), which is value laden and influenced by the data creation context. This is justified if the characteristics are correct and the model is specified, in case this "noise" (accidental effects, being judgmental) does not affect connections between characters.

The construct of technology-enhanced telework is also a latent characteristic. In economics (especially in management theory), there are many such characteristics, and in this field of science, researchers aim to start using methods which would evaluate then adequately. It is important that there is lot of qualitative and almost no quantitative research in virtual work literature, and that gap needs to be filled. It is also important that the development of ICT is so rapid that there is a need to change the model. Gilson et al [25], for example, have noted that the use of social media is the newest channel of communication, and it is already used in workplaces as a communication channel, especially among younger employees who can access information through technology quickly and who often expect instantaneous access.

REFERENCES

1. AERA, APA, & NCME (1999). Standards for educational and psychological testing (Vol. 552). Washington: American Educational Research Association.
2. Ariko Marketing (2002). Infotehnoloogia, telekommunikatsiooni ning kaugtöö kasutuse mõju tööhõivele. Retrieved from http://www.sm.ee/fileadmin/meedia/Dokumendid/Toovaldkond/uuringud/kaugtoo_2002.pdf

3. Babbie E. (2007). The Practice of Social Research. Belmont, CA: Thomson & Wadsworth Publishing Company.
4. Baert P. (2005). Philosophy of the Social Sciences (1st ed.). Polity Press.
5. Beser A., Bahar Z., Aydogdu G.N., Ersin F., Kissal A. (2012). Validity and reliability study for the self-efficacy scale. Journal of Society for Development in New Net Environment in B&H, 6(1), 195–200.
6. Blunch N. (2008). Introduction to Structural Equation Modelling Using SPSS and Amos. SAGE.
7. Bornstein R.F. (2011). Toward a process-focused model of test score validity: improving psychological assessment in science and practice. Psychological Assessment, 23(2), 532–44.
8. Bradley J., Schaefer K.C. (1998). The uses and misuses of data and models: the mathematization of the human sciences. Sage Publications.
9. Campell D.T., Fiske D.W. (1959). Convergent and discriminant validation by the multitrait-multhimethod matrix. Psychological Bulletin, 56(2), 257–72.
10. Chudoba K.M., Wynn E., Lu M., Watson-Manheim M.B. (2005). How virtual are we? Measuring virtuality and understanding its impact in a global organization. Information Systems Journal, 15(4), 279–306.
11. Cicourel A.V. (2007). Interviews, Surveys, and the Problem of Ecological Validity. In Social Surveys 2 (pp. 238–246). Sage Publications.
12. Cook T.D., Campell D.T. (1979). Quasi -experimentation: Design & analysis issues for field settings. Boston, MA: Houghton Mifflin.
13. Cronbach L. (1951). Coefficient alpha and the internal structure of tests. Psychometrika, 16, 297–334.
14. Cronbach L.J., Meehl P.E. (1955). Construct Validity in Psychological Tests. Psychological Bulletin, 52(4), 281–302.
15. Daft R.L. (1999). Leadership: Theory and Practice. Dryden Press.
16. Daft R.L., Lengel R.H. (1986). Organizational Information Requirements, Media Richness and Structural Design. Management Science, 32(5), 554–571.
17. Eilmann S., Solakivi T. (2007). ICT Survey in Estonia. LogOn Baltic Regional Reports.
18. Elmes D., Kantowitz B., Roediger I.H. (2006). Research Methods in Psychology. (8th Ed.). Belmont, CA: Wadsworth Thomson Learning.
19. Fawcett J. (2013). Thoughts about conceptual models and measurement validity. Nursing Science Quarterly, 26(2), 189–91.
20. Fawcett J., Garity J. (2009). Evaluating Research for Evidence-Based Nursing Practice. Philadelphia: F. A. Davis Company.
21. Ferry D.L., Kydd C.T., Sawyer J.E. (2001). Measuring facts of media richness. Journal of Computer Information Systems, 41(4), 69–79.
22. Frankfort-Nachmias C., Leon-Guerrero A. (2011). Social Statistics for a Diverse Society. Pine Forge Press.

23. Ganesh M.P., Gupta M. (2010). Impact of virtualness and task interdependence on extra-role performance in software development teams. Team Performance Management, 16(3/4), 169–186.
24. Gefen D., Boudreau M.-C., Straub D.W. (2001). Structural Equation Modelling and Regression: Guidelines for Research Practice. MIS Quarterly, 25(1), 1–16.
25. Gilson L.L., Maynard M.T., Young N.C.J., Vartiainen M. (2015). Virtual Teams Research: 10 Years, 10 Themes, and 10 Opportunities, 41(5), 1313–1337.
26. Groves M.R., Fowler J.F.J., Couper P.M., Lepkowski M.J., Singer E., Tourangerou R. (2004). Survey Methodology (1st Ed.). Hoboken (N.Y): Wiley-Interscience.
27. Jarvis C.B., Mackenzie S.B., Podsakoff P.M. (2003). A Critical Review of Construct Indicators and Measurement Model Misspecification in Marketing and Consumer Research, 30 (September).
28. John G. (1984). An Empirical Investigation of Some Antecedents of Opportunism in a Marketing Channel. Journal of Marketing Research, 21(3), 278–290.
29. Kane M.T. (2006). Validation. In R.L. Brennen (Ed.), Educational measurement (4th Ed., pp. 17–64). Westport, CT: American Council on Education/Praegar.
30. Lord F.M., Novick M.R. (1968). Statistical theories of mental test scores. Addison-Wesley Pub. Co.
31. Lyman H.B. (1991). Test Scores and what They Mean. Prentice Hall PTR.
32. Messick S. (1989). Validity. In R.L. Linn (Ed.), Educational measurement (3rd Ed., pp. 13–103). New York, NY: American Council on Education/Macmillan.
33. Mihhailova G., Õun K., Türk K. (2009). Virtual work and its challenges and types. The Business Review, Cambridge, 2, 96–103.
34. Mihhailova G., Õun K., Türk K. (2011). Virtual work usage and challenges in different service sector branches. Baltic Journal of Management, 6(3), 342–356.
35. Mulaik S.A. (2009). Foundations of factor analysis. CRC Press.
36. Mölder K. (2004). Tallinna Pedagoogikaülikool .Tallinn.
37. Nunnally J.C. (1967). Psychometric theory (1st Ed.). Michigan: McGraw-Hill.
38. Ozdamar K. (2002). Statistical data analysis with package programs. Eskisehir: Kaan Publishing House.
39. Pennington D.C. (2003). Essential Personality. Arnold.
40. Popper K.R. (1966). The Open Society and Its Enemies (5th Ed.). London: Princeton University Press.
41. Põllumäe S. (2003). Kaugtöö rekendamine riigi ja kohaliku omavalitsuse asutustes. Tallinna Tehnikaülikool.
42. Rebane K. (2004). Kaugtöö töötajate motiveerimise vahendina (võimaluste analüüs AS EMT näitel). Tartu Ülikool.
43. Reino A. (2009). Manifestations of organizational culture based on the example of Estonian organizations. University of Tartu.
44. Sapsford R. (2007). Survey Research. SAGE.
45. Sarantakos S. (2005). Social research (3rd Ed.). New York: Palgrave Macmillan.

46. Saris W.E., Gallhofer I.N. (2007). Design, Evaluation and Analysis of Questionnaires for Survey Research (1st Ed.). John Wiley & Sons.
47. Schaefer R.T., Haaland B. (2003). Sociology: A Brief Introduction. McGraw-Hill Ryerson Ltd.
48. Slocum T.A., Detrich R., Spencer T.D. (2012). Evaluating the Validity of Systematic Reviews to Identify Empirically Supported Treatments. Education and Treatment of Children, 35(2), 201–233.
49. Spearman C. (1927). The abilities of man. New York, NY: Macmillan.
50. Sullivan T.J. (2001). Methods of social research. Fort Worth: Harcourt College Publisher.
51. Tammo E., Kokovkin R., Lüsi M., Telvik A. (2000). Kaugtöö avalikus halduses, Raport Peaministribüroole.
52. Tarling R. (2009). Statistical Modelling for Social Researchers. Principles and practice. Taylor & Francis Ltd.
53. Thurstone L.L. (1947). Multiple factor analysis. Chicago: University of Chicago Press.
54. Treiblmaier H., Bentler P.M., Mair P. (2011). Formative Constructs Implemented via Common Factors. Structural Equation Modeling: A Multidisciplinary Journal, 18(1), 1–17.
55. Vaher B., Seeder K. (Eds.). (2007). Töö ja pere: paindlik töökorraldus ja lastevanemate tööhõive. Eesti Tööandjate Keskliit.
56. Õun K., Mihhailova G. (2011). An improved approach to the construct of virtual work based on Estonian service sector organisations. Lithuanian Journal of Statistic, 50(1), 14–22.

Address for correspondence:
Kandela Õun
Pärnu College, University of Tartu
e-mail: kandela.oun@ut.ee