Turkish Journal of Psychology, December 2018, *33*(82), 65-66 DOI: 10.31828/tpd.13004433.2018.82.02.04

Summary
Within-Person Reliability of Positive and Negative Affect in Repeated Measurements

Ali Can Gök
Middle East Technical University

Emre Selçuk
Middle East Technical University

Tülin Gençöz
Middle East Technical University

Affect is an important construct in understanding psychological functioning utilized as repeated measurements in both scientific research and clinical settings (Merz & Roesch, 2011; Röcke, Li, & Smith, 2009). The Positive and Negative Affect Schedule (PANAS) (Watson, Clark, & Tellegen, 1988) is a widely used instrument in order to measure both positive and negative affect (Schmukle, Egloff, & Burns, 2002; Watson & Vaitya, 2003). In short, the PANAS measures Positive Affect (PA) containing states of being enthusiastic, active, and full of energy. High scores on PA refer to high energy, ability to work with pleasure and concentration. On the other hand, low scores on PA refer to sadness or stagnation. The PANAS measures Negative Affect (NA) containing anger, fear, guilt, and subjective distress. Low scores on NA refer to calmness and tranquility (Watson, et al., 1988; Gençöz 2000).

Repeated measurements are utilized to detect within person changes in affect (Cranford et al., 2006; Eid & Diener, 1999). This enables researcher to track fluctuations in affect while minimizing the retrospective reporting bias (Cranford et al., 2006). However, the detailed picture provided by repeated measures depends on how reliable the instrument is to within person change (Bolger, Davis, & Rafaeli, 2003).

A critical point when dealing with repeated measurements is that variance may stem from two different sources: between and within-person. In cross-sectional design where the researcher takes measurement on one occasion, the variance in the measurement occasion comes from between person differences. The PANAS has been shown to capture between person differences reliably (Watson, et al., 1988; Gençöz 2000). However, in repeated measures, variance stems from not only between person, but also within person differences (Bolger et al., 2003; Cranford et al., 2006; Hu et al., 2016).

According to Fiske and Rice (1955), in order to be able to portion within-person variance, fluctuations in one person's test scores are to be partitioned from unsystematic measurement error. Thus, differences on one participant's test scores on different measurement occasions should be decomposed into unsystematic measurement error and the systematic change that occurs in the participant (Eid & Diener, 1999). Therefore, as in the case with between-person reliability, within-person reliability is to be examined before proceeding with studies with more than one measurement occasion.

In psychometric or adaptation studies, usually test-retest coefficient is reported. However, test-retest coefficient is simply the correlation coefficient of the same measurement on two different measurement occasions (Guttman, 1945; Hu et al., 2016; Shrout & Lane, 2012; Traub, 1994). In test-retest coefficient, the test score is partitioned as unsystematic measurement error and true score. Moreover, there is an assumption that if one participant scores differently on different measurement occasions, the difference is due to unsystematic measurement error. Hence, it is assumed that the true score of one participant is not changed. Therefore, the correlation between two different measurement occasions indicates how reliable the instrument measures true scores of participants on different time occasions (Shavelson, Webb, & Rowley, 1989). This may be misleading when researchers study within-person change, due to the fact that test-retest coefficient ignores within-person changes (Eid & Diener, 1999; Hu et al., 2016).

To overcome these obstacles, Cranford and his colleagues (2006) developed a solution within the framework of Generalizability Theory (Cronbach, Gleser, Nanda, & Rejaratham, 1972). In their solution, the underlying assumption is that variance comes from multiple sources. Therefore, within person change is systematic and it can be measured (Cranford et al., 2006). This approach allows for computing different indices of between-person reliability (Cranford et al., 2006), which we discuss in the following sections.

Address for Correspondence: Dr. Ali Can Gök, Middle East Technical University, Üniversiteler Mah. Dumlupınar Blv. No:1, Department of Psychology, 06800, Ankara
E-mail: alicangk@gmail.com

An investigation of PANAS's between and within-person reliability will be beneficial for its use in studies utilizing more than one measurement occasions (i.e. longitudinal designs, diary designs), and in studies aiming at differentiating between state and trait affect.

Method

In the present study, data were drawn from two different sources. First one is the data set used in PANAS's Turkish adaptation study (Gençöz 2000), in which data were gathered on four different measurement occasions (Study 1). The second one is part of the first author's doctoral dissertation study, in which data were gathered on thirteen different measurement occasions (Study 2).

Participants

Study 1. One hundred ninety-nine undergraduate students, 78 male (39.2%), 108 female (54.3%), 13 unspecified (6.5%), participated in the study. Participants' ages ranged between 18 and 34 (M = 20.90, SD = 1.88). Participants completed the PANAS at least on two measurement occasions. Measurement occasions were three weeks apart.

Study 2. In this study, participants were 46 Clinical Psychology graduate students who were performing psychotherapy under supervision. However, the study proceeded with only 41 therapists due to the fact that participants with at least two measurement occasions were included in the study. There were 35 female (85.4%), 6 male (14.6%) participants. Participants' ages ranged between 23 and 29 (M = 25.56, SD = 1.45). Participants filled out PANAS on thirteen consequent weeks. Measurement occasions were one week apart.

Instruments

Positive and Negative Affect Schedule (PANAS). The PANAS was developed by Watson and his colleagues (1988) in order to measure PA and NA. Its internal consistency coefficient was .88 for PA, and .85 for NA. Test-retest coefficient for the original form is .47. Turkish adaptation study was conducted by Gençöz (2000). Its internal consistency for the Turkish form was .86 for PA, and .85 for NA. Test-retest coefficient for the Turkish form was .47.

Procedure

As mentioned earlier, within the framework of Generalizability Theory (Cronbach et al., 1972), variance comes from different sources (Shavelson et al., 1989). Variance components may come from person, time, item variance and their interactions (Cranford et al., 2006). For example, for person j, on item i, on time k, the score P can be modeled as:

$$P_{ijk} = \mu + K_j + M_i + Z_k + (KM)_{ij} + (KZ)_{jk} + (MZ)_{ik} + (KMZ)_{ijk} + e_{ijk}$$

Grand mean for all items on all measurement occasions is μ; K_j , M_i , Z_k, are effects person, item and time, respectively (Note: Symbols were Turkified for Turkish readers, and are not the same with those on Cranford et. al., 2006). The rest of the symbols represent their two and three-way interactions, and e_{ijk} is the error term (see Cranford et al., 2006). For these effects' variance components, VARCOMP procedure in SPSS can be applied. Formulas for between and within-person reliability indices were derived from Cranford and his colleagues (2006).

Results

After performing variance components analysis, between and within-person reliability indices were calculated (Cranford et al., 2006). For study 1, within-person reliability for PA was .80, and for NA .79. Between-person reliability for one measurement occasion was .84 for PA, and .81 for NA. Between-person reliability for across measurement occasions, where one participant's score on one measurement occasion was checked against another participant's score on another measurement occasion, was .50 for PA, .47 for NA. Finally, between-person reliability for mean PA and NA across all measurement waves was .95 for PA, and .95 for NA.

Same procedure was performed on study 2. Within-person reliability in study 2 for PA was .86, for NA .87. Between-person reliability for one measurement occasion was .86 for PA, and .67 for NA. Between-person reliability for across measurement occasions, where one participant's score on one measurement occasion was checked against another participant's score on another measurement occasion, was .39 for PA, .18 for NA. Finally, between-person reliability for mean PA and NA across all measurement waves was .98 for PA, and .96 for NA.

Discussion

The Turkish version of PANAS showed satisfactory between and within-person coefficients in both studies. Moreover, due to its high level of within-person reliability, the PANAS should be considered as an instrument having good discriminative ability in assessing both state and trait affect. We recommend researchers to provide both between and within-person reliability scores when utilizing the PANAS or any other repeated measures in longitudinal and diary studies.