The Use of Different Multiple Devices for an Ecological Assessment in Psychological Research: An Experience with a Daily Affect Assessment

**Margherita Pasini, Margherita Brondino, Roberto Burro,
Daniela Raccanello and Sara Gallo**

Abstract Data gathering in psychological research is changing given the technological evolution and the availability of many different devices. However, very little research has been done to verify the validity of this kind of psychological measurements, even if a valid measure is the starting point for a valid research. Our study presents the results of a daily measurement of affect connected with the amount of time a person spends in a natural setting, considering the paradigm of restorative environments. High level of compliance was found from respondents, and also a good quality of the used measurements, in term of construct validity (using a Confirmatory Factor Analysis), invariance of the measure across days, and criterion related validity.

1   Introduction

Psychological research is often connected with the assessment of psychological constructs, and good quality of measurements is a fundamental requirement to increase internal validity, i.e. the extent to which a causal relation based on a study can be considered justifiable. The lack of systematic errors increase internal validity and

M. Pasini(✉) · M. Brondino · R. Burro · D. Raccanello · S. Gallo
Department of Philosophy, Education and Psychology,
University of Verona (Italy), Lungadige Porta Vittoria, 17, 37129 Verona, Italy
e-mail: {margherita.pasini,margherita.brondino,roberto.burro,daniela.raccanello}@univr.it,
sara.gallo@studenti.univr.it

a usual source of systematic error in behavioural sciences could be the measurement instrument (scale) used to assess a psychological construct. A good scale has good psychometric properties, specifically validity and reliability. Nowadays, given the growth in technological developments, psychological constructs can be assessed using many different devices (personal computers, laptops, tablets, mobile phones). Even if many advantages can be connected with this possibility, there is a lack of scientific literature concerning the actual uses and the real effectiveness of these applications to increase data quality and consequently psychological research quality [16].

The use of these devices in the assessment could present various advantages. First, one advantage comes from the assessment mode, the self-administration in absence of a data collector. Data quality could be higher in self-administration modes, especially when the topic of interest is in some way sensitive [11]. In addition, people may tend to give more socially desirable responses in interviewers' administration than in self-administration [4].

Another advantage concerns the possibility, in contrast to what happens in laboratory studies, to record the construct of interest within the individual's environment, increasing ecological validity [14]. This possibility is extremely important in environmental psychology, in which research designs are used to explore human behaviour in relation to physical environment. The possibility to assess the constructs of interest when participants are in their real environment (allowing also to measure response latency) can increase the validity of the measures, with respect to the same assessment using lab studies and simulated environments, as pictures or virtual reality. Furthermore, assessing constructs as they naturally occur, permits to avoid some biases connected with retrospective self-report methods, for instance "peak" and "recency" effects [8].

Another advantage in using technological devices as personal computers, tablets or mobile phones to collect data in psychological research studies seems to be connected with higher and easier participation. Axinn, Gatny, and Wagner [1] found that allowing participants to switch modes in mixed-device survey research kept involved more participants compared to a web only approach. When questionnaires are dynamically programmed and suitable for completion on small devices, more people are inclined to use them for survey completion.

Even if this kind of assessment is becoming more and more frequent, the current state of knowledge about the dynamics of taking surveys on mobile devices is not as advanced as necessary, and more scientific literature is needed to evaluate some methodological aspects. For instance, it could be interesting to verify the participants' compliance. Another critical issue concerns the psychometrical properties of scales administered using these modes. A typical psychometric property of a scale consists in construct validity, i.e. the fact that the scale behaves as it should behave, respecting the underlying theory. Another important kind of validity which is usually verified is criterion-related validity, i.e the relation between the measure and another variable theoretically connected with the construct as a criterion variable.

Therefore, the present study aims to address these issues using a daily assessment of affect considering the paradigm of environmental psychology and restorative environments, in particular the "Stress Recovery Theory" [18] and the "Attention Restoration Theory" [9]. Taking into account these theoretical approaches, many research studies report that the exposure to natural environments produces positive changes in emotional states. In particular, experiencing natural environments is connected with health, well-being, and positive mood [3, 7, 17]. The huge amount of studies which report the positive effect of natural environment on affect lead us to choose this paradigm to verify the effectiveness of a multi-device assessment method. In the psychological literature, the two most general components characterizing affective experiences are Positive Affect, PA, and Negative Affect, NA (or Positive and Negative Activation, according to more recent formulations underlying their activation component) [20, 21]. Beyond being the structural dimensions of affect more frequently characterizing English mood terms, they represent the emotional dimensions underlying subjective well-being and they are strongly connected to personality factors [6, 19].

The aim of our study was to verify measurement quality of a daily affect assessment, related to how long participants had been in a natural setting or in a built setting during the day, along one week. This aim was explored in four steps:

- exploring which device was used, the perceived usability of each used device, and the participants' compliance along the week also related to the used device;
- testing the construct validity of the scale used to assess affect, verifying the theoretically based one-factor structure;
- testing the measurement invariance of the same instruments along the seven days in the week;
- testing the criterion-related validity using how long a participant had been in a natural environment as the criterion variable.

2   Method

2.1   Participants

The sample included 108 Italian first- and second-year undergraduate students enrolled at a Psychology degree (mean age: 25 year, SD=6.1; 95% female). Informed consent forms described the potential participants of the goals of the study and that they could stop their participation at any time during the study. To be included in the study, participants had to give their e-mail address. Data were collected between November 13 (Friday) and November 19 (Thursday), 2016. After data collection, all participants were informed in a public meeting about the results.

2.2   *Material and Procedure*

During the study week, the participants received an e-mail message every morning at 10, in which they were reminded to answer to an online questionnaire between 6 p.m. and midnight.

The online questionnaire was developed and administered using the Apsym-Survey Software (ApSS). ApSS is a LAMP[1] customization of the LimeSurvey open-source project that allows to create powerful online question and answer surveys that can work for a few or many simultaneous participants, close or far away, by means of a common internet browser. ApSS follows a Branching/Skip logic procedure, in order to result self-guiding for the participants who are responding. The main operating characteristics that ApSS offers are: 30 different question types, multi-lingual surveys, WYSIWYG (acronym for 'What You See Is What You Get') online-editor, possibility to integrate pictures and movies into a survey, anonymous/not-anonymous surveys, option to buffer answers to continue a survey at a later time, cookie or session based surveys, survey expiry dates, import/export functions for many types of file-extensions, descriptive statistical analyses. In particular, it allows the use of multiple different types of devices (computers, tablets, smartphones) for an ecological assessment due to its responsive-layout, i.e. a web design-approach oriented to provide a final adjusted viewing and interaction experience, comfortable reading and navigation with a minimum of resizing and scrolling across a wide range of devices.

In order to have a better user-experience, an Apsym-Survey Software application for Smartphone/Tablet (ApSS-app, for Android and IPhone/IPad devices) was developed using Apache Flex, an open-source application framework/cross-platform runtime system that allows to develop mobile applications by ActionScript (i.e., an object-oriented programming language).

The questionnaire focused on several measures, among which Positive and Negative Affect and time spent in physical environments. The kind of device used by participants and usability of each device were also monitored with some specific questions.

2.2.1   Positive and Negative Affect Schedule

Affect was measured using the italian version of the Positive and Negative Affect Schedule, PANAS [20], validated into Italian by Terracciano, McCrae, and Costa [15]. It's a 20-item scale presenting adjectives referred to two higher-order factors, Positive Affect (e.g., *active*, *enthusiastic*, *excited*) and Negative Affect (e.g., *afraid, upset, distressed*), with ten adjectives for each factor. Participants were asked to evaluate how much they had felt the state described by each adjective on

---

[1] LAMP is an model of web service solution stacks, acronym of the names of four original open-source parts: Linux, the well-known Unix-like computer operating system (OS), Apa-che HTTP Web-Server, MySQL relational database management system (RDBMS), and PHP server-side scripting language for web development.

a 7-point Likert-type scale (1 = *not at all* and 7 = *very much*), referring to the current day. Since natural environment should increase positive mood, only the Positive Affect dimension was considered in the present study.

2.2.2   Physical Environment

One item was used to assess the amount of time a person was in a natural environment during each day. The question concerns how long participants had been in a natural setting, and it was evaluated on a 5-point Likert scale (1 = never; 2 = until 30 minutes; 3 = from 30 minutes to 1 hour; 4 = from 1 to 2 hours; 5 = more than 2 hours).

2.2.3   Devices' Use and Usability

Which device has been used and the evaluation of usability of each device were assessed in the last-day survey, with some specific items.

2.3   *Data Analyses*

First, descriptive statistics were computed, to explore the use of the different devices and the compliance. We define compliance as answering to the questionnaire all the seven days. To test construct validity, Confirmatory Factor Analysis (CFA) was performed, to verify the mono-factorial structure of the scale. Due to the structure of the data, collected for seven days on the same sample, multilevel CFA was run using complex option of Mplus, version 6.11 [12]. We took into account the Chi-square ($\chi2$), the comparative fit index (CFI), the Tucker-Lewis index (TLI), the root-mean-square error of approximation (RMSEA), and the standardized root-mean-square residuals (SRMR), with CFI $\geq$ .90, RMSEA $\leq$ .08, and SRMR $\leq$ .11 as threshold values [2]. Measurement invariance was verified with a multi-group CFA. We checked for the model invariance across the seven days assessing the factor structure simultaneously for each day. Measurement Invariance (MI) analyses examined hypotheses on the similarity of the covariance structure across days by considering: (1) configural invariance, allowing all the parameters to be freely estimated; (2) metric invariance, requiring invariant factor loadings; (3) scalar invariance, requiring also invariant intercepts; and (4) uniqueness invariance, requiring invariant item uniqueness. Due to the small size of our sample, support for non-invariance required Δ CFI $\leq$ -.005, supplemented by Δ RMSEA $\geq$ .010 or SRMR $\geq$ .025, for testing metric invariance, and .010 or .005, respectively, for testing scalar and uniqueness invariance.

   To test criterion-related validity, bivariate Pearson's correlation coefficient between positive affect and the amount of time spent in a physical environment during the seven days was used.

3   Results and Discussion

3.1   Use of the Devices and Compliance

Smartphone has been the most used device, with 77.8% of the participants using it at least one time, followed by notebook (41.7%), tablet (20.4%), PC desktop (15.7%) and mobile phone without touch screen (11.1%). When asked to estimate the percentage of use of a device during the week, mean percentage of use of smartphone was 78.8%, 34.4% for notebook, 17.2% for tablet, 14.4% for mobile phone, and 11.8% for PC desktop. Concerning usability, estimated on a 7-point Likert scale, all the devices were evaluated quite easy to use: notebooks and tablets were evaluated as the easiest devices, with 83.7% and 81.3% of the sample choosing the first two point of the scale ("not demanding at all", "low demanding"). A bit more demanding was considered the smartphone, with 52.9% of participants considering it "not demanding at all" or "low demanding".

Participants' compliance was quite high: 77.8% of participants answered each of the seven days. Considering the participants who did not comply, 70.8% did not answer once, 20.8% did not answer twice, and 8.3% of the sample did not answer for three days. Along days, compliance varied from 97.2% to 93.5%, and no week-end effect could be identified [5]. Compliance was also related to the used device: only 19% of missing was present in participants who declared to use the smartphone more frequently than the other devices.

3.2   Psychometric Properties of the PANAS Positive Affect Subscale

Since some adjectives of the scale were almost synonymous (enthusiastic and excited; concentrated and attentive), we decided to add two covariances between errors in the model. The model showed good fit indexes: $\chi^2(33) = 151.56$, $p < .001$, CFI =.96, TLI = .94, RMSEA = .070, and SRMR = .036. All factor loadings were statistically different from 0, ranging from .48 to .80. This result confirms the theoretical one-factor structure of the Positive subscale of the PANAS administered using this mode.

3.3   Invariance of Measurement across Days

The results of the sequence of gradually more restrictive tests of MI supported the higher level of invariance, in particular: metric invariance (Δ CFI = .000; Δ RMSEA = .004), scalar invariance (Δ CFI = .002; Δ RMSEA = .002), and uniqueness invariance (Δ CFI = .005; Δ RMSEA = .008), across days. This means that the same factor structure of the scale, that is the theoretically based one-factor structure, was confirmed across the seven days.

3.4 Criterion-Related Validity

Bivariate Pearson's correlations were computed, day by day, between the extent of the time spent in a natural environment and the Positive Affect. Table 1 shows these correlations.

Table 1 Pearson's correlation coefficients (*r*) between time spent in a natural setting and Positive Affect, for the seven days

	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7
Correlations between time spent in a natural setting and Positive Affect	.21*	.21*	.26*	.21*	-.11	.20*	-.01

*$p < .05$.

Except for day 5 and day 7, for which the two variables were not related, in all the other five days we found significantly positive moderate correlations, confirming that a more intense positive affect corresponds to a higher amount of time one person spends in a natural setting.

4 Conclusions

Our research study aimed at investigating the measurement quality of a daily affect self-assessment using different multiple devices, in the participants' everyday life setting and not in a lab. The advantage to use this assessment mode is connected to the possibility to assess a psychological construct in an ecological situation. This is extremely important, for instance, in environmental psychology, that study the physical environment's impact on human perceptions and behaviours. In the present study the daily affect assessment was also related to the amount of time participants have been daily in a natural setting, considering the paradigm of environmental psychology and restorative environments as the theoretical frameworks [9, 18].

First, an interesting result concerns the most used device by the sample of young-adult participants: smartphone was the favourite device, and the other devices (in order: notebook, tablet, PC desktop, mobile Phone with no touch-screen) was clearly less used. This could depend on the age and educational level of participants: younger and educated population tend to use smartphone more than other populations, and more than other devices [10]. Despite the fact that this was the favourite device, smartphone's usability for this survey was considered lower than other devices' usability, such as notebook or tablet. This result suggests that more information should be gathered about the reasons of this lower perceived usability in the present case. Data also documented a high participants' compliance along the week (highest for smartphone users), suggesting that proposing an easy way for participants to respond to a survey in their everyday life setting could lead to a higher compliance and to lower sample mortality. To verify these speculations, future works should compare directly compliance investigated with the

described modalities to compliance typical of more traditional data gathering instruments. Second, our analyses indicated the goodness of the structure of the PANAS Positive Affect subscale, supporting the construct validity of the measurement instrument to assess affect also when presented through a multi-device approach system. Third, measurement invariance of measures of positive affect along seven days in the week further supported their validity within multi-device surveys. Finally, the relation between natural environment and positive affect documents the criterion-related validity of the measurement of affect also in this mode; in addition, these findings contribute to extend current theoretical knowledge on the links between natural environments and affect for a specific sample.

A severe limitation of this study concerns the sociodemographic characteristics of the sample, which was mainly composed of female and young participants. It should be important to repeat the same research using samples more balanced for sex, and also involve older participants, to exclude a cohort effect. Future directions could also explore the measurement invariance across modes [22], to separately verify whether differences in the quality of measurements could be found considering the different devices. Also considering computer skills of participants could be interesting in future researches.

Relation between positive affect and amount of time spent in a natural setting was found, but in this study no attention has been paid to the perceived restorativeness of the natural setting, which should be the direct cause of positive affect: Assessing also this perception, using the same mode and a short instrument [13], could contribute to explain better the link between natural environment and positive affect.

Acknowledging limitations such as, for instance, the impossibility to separate the effect of different devices, or the non-balanced by sex and age sample, our work could be considered a first step to document how the growth in technological developments can be advantageous in terms of better data quality and related psychological research quality, considering the lack of scientific findings documenting it.

References

1. Axinn, W., Gatny, H., Wagner, J.: Maximizing data quality using mode switching in mixed-device survey design: Nonresponse bias and models of demographic behaviour. Methods, Data, Analyses **9**(2), 163–184 (2015)
2. Beauducel, A., Wittmann, W.W.: Simulation study on fit indexes in CFA based on data with slightly distorted simple structure. Structural Equation Modeling **12**, 41–75 (2005). doi:10.1207/s15328007sem1201_3
3. Berto, R.: The role of nature in coping with psycho-physiological stress: A literature review on restorativeness. Behavioural Sciences **4**(4), 394–409 (2014). doi:10.3390/bs4040394
4. Bowling, A.: Mode of questionnaire administration can have serious effects on data quality. Journal of Public Health **27**(3), 281–291 (2005)
5. Courvoisier, D.S., Eid, M., Lischetzke, T., Schreiber, W.H.: Psychometric properties of a computerized mobile phone method for assessing mood in daily life. Emotion **10**(1), 115–124 (2010)

6. Diener, E., Emmons, R.A., Larsen, R.J., Griffin, S.: The Satisfaction With Life Scale. Journal of Personality Assessment **49**, 71–75 (1985)
7. Hartig, T., Mitchell, R., de Vries, S., Frumkin, H.: Nature and health. Annual Review of Public Health **35**, 207–228 (2014)
8. Kahneman, D.: Objective happiness. In: Kahneman, D., Diener, E., Schwarz, N. (eds.) Well-being: The Foundations of Hedonic Psychology, pp. 3–25. Russell Sage Foundation, New York (1999)
9. Kaplan, S.: The restorative benefits of nature: Toward and integrative framework. Journal of Environmental Psychology **15**, 169–182 (1995)
10. Kim, Y., Briley, D.A., Ocepek, M.G.: Differential innovation of smartphone and application use by sociodemographics and personality. Computers in Human Behavior **44**, 141–147 (2015)
11. Kreuter, F., Presser, S., Tourangeau, R.: Social desirability bias in CATI, IVR, and web surveys: The effects of mode and question sensitivity. Public Opinion Quarterly **72**, 847–865 (2008)
12. Muthén, L.K., Muthén, B.O.: Mplus user's guide, 5th edn. Muthén & Muthén, Los Angeles (1998–2007)
13. Pasini, M., Berto, R., Brondino, M., Hall, R., Ortner, C.: How to measure the restorative quality of environments: The PRS-11. Procedia - Social and Behavioral Sciences **159**, 293–297 (2014)
14. Shiffman, S., Stone, A.A., Hufford, M.R.: Ecological momentary assessment. Annual Review in Clinical Psychology **4**, 1–32 (2008)
15. Terracciano, A., McCrae, R.R., Costa, P.T.: Factorial and construct validity of the Italian Positive and Negative Affect Schedule (PANAS). European Journal of Psychological Assessment **19**(2), 131–141 (2003). doi:10.1027//1015-5759.19.2.131
16. Toepoel, V., Lugtig, P.: Online surveys are mixed-device surveys. Issues associated with the use of different (mobile) devices in web surveys. Methods, Data, Analyses **9**(2), 155–162 (2015). doi:10.12758/mda.2015.009
17. Tyrväinen, L., Ojala, A., Korpela, K., Lanki, T., Tsunetsugu, Y., Kagawa, T.: The influence of urban green environments on stress relief measures: A field experiment. Journal of Environmental Psychology **38**, 1–9 (2014)
18. Ulrich, R.S.: Aesthetic and affective response to natural environment. In: Altman, I., Wohlwill, J.F. (eds.) Behavior and the Natural Environment, pp. 85–125. Plenum, New York (1983)
19. Watson, D., Clark, L.A.: On traits and temperament - General and specific factors of emotional experience and their relation to the 5-factor model. Journal of Personality **60**, 441–476 (1992)
20. Watson, D., Clark, L.A., Tellegen, A.: Development and validation of brief measures of positive and negative affect: The PANAS scales. Journal of Personality and Social Psychology **54**, 1063–1070 (1988)
21. Watson, D., Wiese, D., Vaidya, J., Tellegen, A.: The two general activation systems of affect: Structural findings, evolutionary considerations, and psychobiological evidence. Journal of Personality and Social Psychology **76**, 820–838 (1999)
22. Wong, C.K.H., Lam, C.L.K., Mulhern, B., Law, W.-L., Poon, J.T.C., Kwong, D.L.W., Tsang, J.: Measurement invariance of the Functional Assessment of Cancer Therapy—Colorectal quality-of-life instrument among modes of administration. Quality of Life Research **22**, 1415–1426 (2013). doi:10.1007/s11136-012-0272-x