ELSEVIER

International Journal of Psychophysiology 39 (2000) 67–77

INTERNATIONAL JOURNAL OF PSYCHOPHYSIOLOGY

www.elsevier.nl/locate/ijpsycho

The effect of data aggregation on temporal stability of cardiovascular reactivity

Andreas Hinz*, Heinz Schaffernicht, Doris Kuchenbecker, Romy Schetschorke, Doreen Lachnitt

Institute of Occupational and Social Medicine, University of Leipzig, Riemannstrasse 32, D-04107 Leipzig, Germany

Received 19 January 2000; received in revised form 11 April 2000; accepted 2 May 2000

Abstract

Temporal stability of behaviorally evoked cardiovascular responses is important for theoretical (concept of activation) and practical (risk for cardiovascular diseases) reasons. As in test psychology, reliability of physiological responsivity depends on the degree of data aggregation across several measurements. This paper describes a statistical approach based on intra-class correlations. This approach is suited to define certain stability measures based on variance components representing different levels of data aggregation. An empirical investigation is presented comprised of 58 subjects, three physiological parameters (heart rate, systolic and diastolic blood pressure), two mental tasks, two sequences of the tasks within one session, and 2 days with an interval of 4 weeks between them. In addition to the finding that data aggregation can generally increase stability, the different sources of aggregation (across phases within a task, across tasks, and across task sequences) and their combinations are systematically compared with regard to their contribution to this enhancement. Finally, it will be shown how the approach can be utilized to explain aggregation effects for other psychophysiological research questions such as covariation, consistency, and ambulatory assessment of cardiovascular functioning. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: Reproducibility; Temporal stability; Cardiovascular reactivity; Data aggregation; Mental challenge; Intra-class correlation

* Corresponding author. Tel.: +49-341-9715-404; fax: +49-341-9715-409.
E-mail address: hinz@medizin.uni_leipzig.de (A. Hinz).

PII: S0167-8760(00)00117-3

1. Introduction

Behavioral stimuli often evoke responses of the autonomic system, and the magnitude of such responses varies significantly among subjects. It has been hypothesized that exaggerated physiologic responses to psychological challenge may be implicated in the etiology of cardiovascular diseases, including coronary heart disease and essential hypertension (Fredrikson and Matthews, 1990; Pickering and Gerin, 1990). One requirement that needs to be met in order to use cardiovascular reactivity as a risk factor is a certain degree of temporal stability and trans-situational consistency of the cardiovascular responsiveness. In recent years, the issue of temporal stability has been increasingly investigated. While Manuck et al. (1989) summarized the results of 12 published studies which were performed to test stability of cardiovascular responses, Swain and Suls (1996) presented a meta-analysis of 31 papers with 95 coefficients for heart rate (HR) and 73 coefficients for systolic blood pressure (SBP) and diastolic blood pressure (DBP). Mean stability coefficients in the Swain and Suls study were: HR, $r = 0.55$; SBP, $r = 0.41$; and DBP, $r = 0.35$. The studies incorporated in this meta-analysis differed with regards to several criteria: size and composition of the subject sample, type and duration of the mental tasks, definition of baseline values and change scores, and interval between the replications. Some of these aspects are explicitly addressed by Swain and Suls (1996). One of their results was that the stability of blood pressure reactivity depends on the degree of data aggregation. When the blood pressure assessments comprised only a few readings, stability coefficients were lower than in cases with more readings. A similar result was obtained by Kamarck et al. (1992) and Manuck et al. (1993). They found that the reliability of reactivity assessments can be enhanced by aggregating across individual's responses over several different tasks. From a psychometric point of view, this finding is not surprising. Epstein (1979, 1983, 1986) showed how temporal reliability can become higher when the data are aggregated.

The generalizability theory (Cronbach et al., 1972; Brennan, 1992) is a very useful tool to describe how single measurements can generalize across several conditions. The statistical basis of this theory is an ANOVA model with random factors. For simple (single-facet) designs the generalizability coefficient corresponds with the intra-class correlation. One possible application of generalizability coefficients is in the assessment of temporal stability, understood as generalizability over time. Llabre et al. (1993) use generalizability coefficients for this purpose as an alternative to the usually adopted Pearson correlations. In a multi-facet design it is possible to assess the generalizability across multiple sources of variation. For example, Llabre et al. (1988) demonstrate the generalizability of cardiovascular responses across replications, across settings (laboratory, home and work), and across instruments. Gerin et al. (1996) also show how the assessment of generalizability or reliability of cardiovascular responses can be qualified when several sources of variance (replications under identical conditions and changes in the setting) are differentiated. One application of this theory concerns the number of measurements that are necessary in order to assess the 'true' score accurately. For example, Llabre et al. (1988) conclude from generalizability coefficients that six readings of systolic blood pressure are needed at home or at work to assess systolic blood pressure with a sufficient accuracy.

A problem that occurs sometimes in the application of the generalizability theory is that of negative estimates of variance components. These negative estimates are usually replaced with zero (Llabre et al., 1988, Table 3; Van Doornen et al., 1994, Table 6; Gerin et al., 1996, Table 4; Marwitz and Stemmler, 1998, Table 3a). These estimates indicate that the model assumptions have not been perfectly met and that there are inconsistencies in the estimated variance components.

Based on the technique of covariance partitioning (Stemmler and Fahrenberg, 1989; Stemmler, 1992), which was developed to specify several aspects of covariation between physiological variables, Hinz et al. (1994) constructed several measures for temporal stability which also take into account the degree of data aggregation. Stability