Ecological Momentary Assessment: Statistical Issues

J.B. Nezlek[1]

[1]Psychology, College of William & Mary, Williamsburg

The type of data collected in ESM/EMA studies are usually referred to as nested data. For example, if participants provide data every day for two weeks, this would be described as 'days nested within participants" and would be called a two level data set. Similarly, if participants provided multiple observations each day for two weeks this would be described as 'observations nested within days nested within participants" and would be called a three level data set.

Such data cannot be analyzed using OLS multiple regression (e.g., the data provided each day by each participant is a unit of analysis) because such analyses assume that observations are independent, and they are not. Moreover, although some OLS analyses may take such dependencies into account, they fall short in other ways, e.g., they do not model error properly.

Analyzing such data structures requires the use of specific techniques that are referred to as *multilevel modeling*. In such analyses, relationships at all levels of analysis are modeled simultaneously. For example, in a two level model (days nested within persons), we could model within-person relationships between stress and symptom expression, and then model the between-person differences in the strength of these within-person relationships. Who reacts more strongly to stress, who reacts less strongly? Within- and between-person measures can be continuous or categorical, and they can be combined, and temporal dynamics at either level of analysis can be examined to reveal mechanisms underlying symptom expression.