

## Naturalistic Observation of Daily Behaviour in Personality Psychology

MATTHIAS R. MEHL

*Department of Psychology, University of Arizona, Tucson, AZ, USA*

mehl@email.arizona.edu

### *Abstract*

*This comment highlights naturalistic observation as a specific method within Furr's (this issue) cluster direct behavioural observation and discusses the Electronically Activated Recorder (EAR) as a naturalistic observation sampling method that can be used in relatively large, nomothetic studies. Naturalistic observation with a method such as the EAR can inform researchers' understanding of personality in its relationship to daily behaviour in two important ways. It can help calibrate personality effects against act-frequencies of real-world behaviour and provide ecological, behavioural personality criteria that are independent of self-report. Copyright © 2009 John Wiley & Sons, Ltd.*

Furr's target paper (this issue) provides an excellent and sorely needed analysis of the nature, potentials and state-of-the-science of behavioural assessment in personality

psychology. This comment highlights *naturalistic observation* as a specific method within Furr's cluster *direct behavioural observation* and discusses ways in which it can help the field become 'a truly *behavioural science*.' Naturalistic observation is the observation of subjects in their natural habitat. Whereas the method is fairly common in neighbouring disciplines (e.g. anthropology, primatology) and areas (e.g. developmental psychology), it has a thin history in personality psychology (Barker & Wright, 1951; Craik, 2000).

Over the last 10 years, I have co-developed and validated the Electronically Activated Recorder or EAR (Mehl et al., 2001) as a naturalistic (acoustic) observation sampling method that can be used in relatively large, nomothetic studies. The EAR is a pocket-sized audio-recorder that periodically samples snippets of ambient sounds from people's momentary environments (e.g. 30 second every 12.5 minute). Participants carry the device around while going about their normal lives. That way, the EAR produces acoustic logs of their daily behaviours as they naturally occur over the course of a day.

In a series of studies, my colleagues and I have used the method to show that a broad spectrum of acoustically detectable behaviours can be assessed reliably and with low levels of reactivity from the sampled ambient sounds (Mehl & Holleran, 2007), show very large between-persons variability and good temporal stability (Mehl & Pennebaker, 2003) and have good convergent validity with theoretically related trait measures such as the Big Five (Mehl et al., 2006) and subclinical depression (Mehl, 2006).

Naturalistic observation with a method such as the EAR can inform researchers' understanding of personality in its relationship to daily behaviour in at least two ways.

### **NATURALISTIC OBSERVATION CAN HELP CALIBRATE PERSONALITY EFFECTS AGAINST ACT-FREQUENCIES OF REAL-WORLD BEHAVIOUR**

The field has long defended itself against criticisms of the limited magnitude of personality effects. Even though two recent landmark studies went a long way to silence them (Ozer & Benet-Martinez, 2006; Roberts, Kuncel, Shiner, Caspi, & Goldberg, 2007), the field still needs more effective ways to communicate the implications of its findings. In psychology, the vast majority of measures use arbitrary metrics. Calibrating effects based on arbitrary metrics against inherently meaningful, real-world referents is one way to come to a 'better understanding of the measures by which the phenomena with which we concern ourselves are gauged' (Sechrest, McKnight, & McKnight, 1996, p. 1068).

One advantage of the EAR is that its sound-file based behavioural codings can be readily converted into a metric that is non-arbitrary, intuitively meaningful and inherently real-world relevant. If the EAR captures a person talking in 40 out of 120 recordings, one can estimate that the person spent about a third of her time awake (or about 5 hours) talking. By linking EAR-derived act-frequencies of daily behaviour to individual differences, a better understanding of effect sizes can be obtained. For example, in one study conscientiousness correlated  $r = -.29$  with EAR-assessed swearing and  $r = .42$  with EAR-coded time in class (Mehl et al., 2006). Converted into a more meaningful metric, this suggests that individuals who marked a '4' on the 5-point conscientiousness scale, compared to those who marked a '2,' used profanity at less than half the rate (0.5 vs. 1.2%) and spent about three times as much time in class (11.9 vs. 4.1%).

Similarly, testing the myth that women are by a factor more verbose than men, we estimated based on six EAR studies that both men and women use about 16 000 words per day (Mehl et al., 2007). A sex difference of 546 words compared to a range of over 46 000 words between the least and most talkative individual (695 vs. 47 016) rendered significance testing close to meaningless—and speaks powerfully to the magnitude of individual differences. Thus, in facilitating an absolute metric in the measurement of daily behaviour, naturalistic observation can help ‘benchmark’ personality effects.

### **NATURALISTIC OBSERVATION CAN PROVIDE ECOLOGICAL, BEHAVIOURAL PERSONALITY CRITERIA THAT ARE INDEPENDENT OF SELF-REPORT**

The ‘criterion problem’ is a vexing issue in the field. Generally, behavioural criteria are deemed preferable to those based on self- or informant reports. For assessing personality in the field, experience sampling has emerged as the best available proxy to behavioural observation (Spain, Eaton, & Funder, 2000). In cases where it is necessary to measure real-world, behavioural personality criteria independent of self-report, the EAR can help accomplish this.

For example, we tested the accuracy of self- and other-reports by comparing the predictive validity of participants’ self-ratings of how much they engage in different daily behaviours to similar ratings obtained from people who knew the participants well. The frequency with which the EAR captured participants actually engaging in these behaviours served as ‘impartial’ accuracy criterion. Self- and other-ratings showed identical validity but also uniquely predicted certain behaviours (Vazire & Mehl, 2008). Importantly, to avoid giving one perspective an undue advantage, it was critical to minimize shared method variance with the two. The EAR-derived behaviour counts maximally accomplished this while at the same time preserving the study’s ecological focus.

Similarly, responding to Terracciano et al.’s (2005) influential finding that national stereotypes have zero validity, Heine, Buchtel, and Norenzayan (2008) argued that ‘comparing means on subjective Likert self-report scales is the most commonly used method for investigating cross-cultural differences, yet there are many methodological challenges associated with this approach’ (p. 309). Following their solution to concentrate on behavioural trait markers, we compared Americans’ and Mexicans’ sociability in a binational EAR study (Ramírez-Esparza, Mehl, Álvarez Bermúdez, & Pennebaker, 2009). We found that although American participants reported being more sociable than their Mexican counterparts, they spent less time with others and had fewer social (i.e. non-instrumental) conversations. Intriguingly, whereas Americans rated themselves significantly *higher* than Mexicans on the item ‘I see myself as a person who is talkative,’ they spent in fact almost 10% *less* time talking (34.3 vs. 43.2%). Again, in providing ecological, behavioural personality criteria, naturalistic observation can help resolve important debates within the field.

To summarize, naturalistic observation clearly occupies a methodological niche; it is not for everyone and everything. It is highly labour-intensive and thus requires careful deliberation as to when it should be used instead of more economic methods. However, in providing arguably the strongest form of behavioural data, it can yield valuable findings that cannot be obtained otherwise and that way contribute to the field becoming a leading behavioural science.