# The Psychologist

# Archive - 2008

## [Volume 21 - Part 12 - (December 2008)](http://www.thepsychologist.org.uk/archive/archive_home.cfm?volumeID=21&editionID=167)

## Methods: Studying the natural history of behaviour

**Daryl B. O’Connor and Eamonn Ferguson on the use of diary methods in psychology**

**Pages:** 1034-1037

Methods that allow scientists to identify the momentary patterns of our behaviour – their ‘natural history’, their hierarchies or responses to interventions or experimental manipulations – are incredibly useful. Medical researchers, epidemiologists, economists and policy researchers can use them in seeking to understand the burden of different diseases, or the health consequences of psychological, social and environmental stressors (see Ferguson, 2005; Kahneman et al., 2004). These techniques, variously referred to as experience sampling methods (ESM) or diary methods (DM), allow for in situ assessments of behaviour – they get at the there and then. This means they can reduce retrospective bias, as long as certain design constraints are achieved (Affleck et al., 1999; Ferguson, 2005). Recently an alternative to DM has been introduced – the day reconstruction method (DRM: Kahneman et al., 2004) – and we shall compare these two methods after describing both below.

Methods that allow scientists to identify the momentary patterns of our behaviour – their ‘natural history’, their hierarchies or responses to interventions or experimental manipulations – are incredibly useful. Medical researchers, epidemiologists, economists and policy researchers can use them in seeking to understand the burden of different diseases, or the health consequences of psychological, social and environmental stressors (see Ferguson, 2005; Kahneman et al., 2004). These techniques, variously referred to as experience sampling methods (ESM) or diary methods (DM), allow for in situ assessments of behaviour – they get at the there and then. This means they can reduce retrospective bias, as long as certain design constraints are achieved (Affleck et al., 1999; Ferguson, 2005). Recently an alternative to DM has been introduced – the day reconstruction method (DRM: Kahneman et al., 2004) – and we shall compare these two methods after describing both below.

**Diary methods**
In psychological research, the dependent variable under investigation is often  a process (e.g. stress, mood, recall, attention), that changes from moment  to moment (where a moment can be minutes, hours, days, weeks, etc.). Knowing the ‘natural history’ of these patterns is extremely useful. In field studies (experimental or otherwise) conventional cross-sectional or longitudinal study designs without
DM do not allow us to examine research questions within the context of fluctuating daily processes.
There are three different DM protocols:

* Interval-contingent: the participant completes the diary at specified intervals (e.g. end of each day). For example, it could be used to ascertain whether psychological stress on one day ‘predicts’ symptom ‘flare-ups’ on the next day in psoriasis patients. Interval-contingent protocols are especially useful for frequent behaviours without a definitive start and end.
* Event-contingent: the participant completes the diary each time a specific event happens. For example, we could investigate whether the act of smoking moderates daily mood (Moghaddam & Ferguson, 2007). Event-contingent protocols are especially useful to estimate event prevalence
* Signal-contingent: the participant completes the diary in responseto random ‘alarms’ or ‘beeps’ from a palmtop computer or similar device.

It can be used, for example, to examine the pattern of mood and its moderation by personality. Signal-contingent protocols are especially useful for recording data on the distribution, frequency and duration of events It should be noted that these DM protocols are not mutually exclusive, and researchers should consider DM combinations (Shiffman et al., 2008). For example, it is possible to examine the relationship between mood and smoking behaviour (event-contingent) within different temporal epochs (interval contingent: e.g. specified periods across
a day) (cf. Moghaddam, 2007).

It is also important to consider how frequently and for how long (hours, to months) participants are expected to complete their diaries (known as the duration-frequency decision; see Ferguson, 2005, for more details). The answer depends on the predicted temporal properties of the variable being studied (i.e. mood would be assessed at many points over a short interval, whereas long-term illness may be monitored at fewer points over a longer period).

**Compliance and causality**
Diary methods allow causality to be inferred when observing systems that cannot be directly manipulated (see West & Hepworth, 1991). That is, where more then one variable is measured at a number of different time points these can be compared and lagged relationships examined (e.g. does variable X measured on a Monday predict variable Y measured on a Tuesday or is the reverse true?). Essential to this, therefore, is ensuring that measures reported on Monday were actually recorded on Monday and measures on Tuesday were actually recorded on Tuesday and so forth. Therefore, compliance is important so that lagged relationships can be examined. In relation to standard (interval-contingent) daily diary studies, for example, it has been recommended that data are returned at the end of each day so that causal lags between days can be examined. This also prevents hoarding or backfilling, whereby participants complete missed or forgotten entries when they make their next entry. With signal-contingent DM, then, it is necessary to record if the participants have responded when the signal occurs and palmtop computers or time-locked paper or test tubes are essential here (Stone et al., 2002). For event-contingent DM compliance is harder to ascertain, as the event itself is the trigger. Other design parameters to try to increase compliance include ‘bogus pipeline’ procedures: convincing the respondent that the researcher has a reliable and valid means, usually in the form of a sham lie detector, of checking the truthfulness of responses.

Recently the major pros and cons of paper-and-pencil versus palmtops have been debated (e.g. Green et al., 2006; Tennen et al., 2006). Stone et al. (2002) published a report highlighting the potential seriousness of not using electronic diary methods. Using a time-based diary design, participants were asked to complete measures at three specified times each study day – 10am, 4pm and 8pm. Compliance rates were compared in two groups: one used an electronic diary and the other used paper diaries equipped, unknown to the participants, with a small chip that recorded the openings and closings of the diary. The results were rather startling. The electronic diary group was found to be compliant on 94 per cent of occasions whereas the paper diary group was compliant on only 11 per cent of occasions. More alarmingly, on only a few occasions did the latter group report non-compliance. However, a number of factors have recently been identified that may account for these apparent large differences between the diary techniques. Green et al. (2006) have argued that  findings of Stone et al. are confounded by differential awareness of being monitored, feedback about actual compliance; and participant motivation between the groups. In a series of studies, the authors demonstrated very close matches between electronic and paper formats and presented convincing evidence that the results could not have been fabricated. In fact, these authors concluded that problems relating to compliance were more likely to be associated with the study design and participant motivation than due to whether the data were collected in an electronic or paper format.

**Data analyses**
The data from DM (as all data in fact) are essentially hierarchical as a consequence of our sampling and experimental procedures. Multilevel modelling (MLM) procedures are therefore used to examine such data. If we wanted to know if daily hassles influence health behaviours and whether this relationship is moderated by personality, we would set up an MLM at two levels. Level 1 would be the daily variation in hassles and health behaviours (within-subject variation) and Level 2 would be the between-subject variation in personality (e.g. Jones et al., 2007; O’Connor et al., 2008). In MLM the outputs at one level (regression coefficients) become the parameters that are modelled/analysed at the next level.

MLM procedures can be used to examine experimental data with repeated assessments (Ferguson et al., 2007) and physiological data (Ferguson, 2008; see also Newman et al., 2007). For a detailed introduction to MLM procedures see Kreft and De Leeuw (2006) or Hox (2002).

**Fallacies**
MLM should be used with hierarchical data sets in order to avoid certain erroneous conclusions based on the atomistic or ecological fallacies. These fallacies are based on the inappropriate assumption that relationships at one level in a hierarchy apply at another.

The ecological fallacy (also known as the ‘Robinson effect’: Robinson, 1950) occurs when inferences are made about associations between variables at an individual level based upon group level data. The problem occurs when the association between two variables at the group level differs from the associations between similar variables measured at the individual level. For example, while it is true that people with strong family ties are less likely to be criminals, it does not follow that all people with strong family ties are law abiding.

Within the context of psychological research, the same problem can arise if we draw within-person inferences based upon aggregated data units from across-person associations. For example, the relationship between stress levels at an organisational level and job control may be negative, indicating that higher job control is associated with lower stress. However, this inference may fail to account for some employees with high control over their jobs who will always experience higher stress than employees with lower control (because they are less well equipped to cope with the demands of their job).

The atomistic fallacy is the converse, with inferences about groups inferred from individual level data (e.g. implying group productivity from knowledge of individual work rates). It is important to note that both fallacies concern inference and not measurement issues.

A related fallacy is the ‘Simpson paradox’, which occurs when data from heterogeneous groups is collapsed and analysed as if it were one homogeneous group. For instance, Sober and Wilson (1998) give an example where the data suggest that at a university level there is a discriminatory bias against women applicants to courses; however, at a departmental level there is no evidence for any sex discrimination. This is due to variable acceptance rates across departments. If a large proportion of men apply to a high acceptance rate department and a high proportion of women to a low acceptance rate department and selection is made proportionally on the acceptance rates within the department, then the within-department ratio remains as fair but over the university there will be relatively fewer women.

**The day reconstruction method**
As a potential alternative to DM, the day reconstruction method (DRM) has been introduced (Kahneman et al., 2004). The DRM is a structured recall technique that enables participants to recall events, mood and well-being indicators over the previous day.

While recognising daily in situ assessment as the gold standard, the DRM is a more efficient and less time-consuming technique. For example, Stone et al. (2006) recently used DRM to examine the diurnal rhythms of emotions in a large-scale investigation and identified a number of specific patterns that had not previously been observed. These authors argue that DM are frequently expensive, are associated with a heavy participant burden and they do not allow for the detection of finer- grained patterns of behaviour or emotion. The DRM technique does not interfere with daily activities and allows for detailed examination of daily events. However, it does not tell us anything about causal relations between study variables. DM also allow for the possibility to assess cognitive functions in situ (e.g. emotional stroops, dot probes) in a way that you cannot with the DRM. With this in mind, the DRM should be developed to include clinical and health-relevant variables and extended to include multiple assessment points over time (e.g. one DRM per week for two months).

**Advantages of diary methods**
What are the advantages of using DM? This is best summed up in the words of Affleck and colleagues (1999) who argue that daily diary studies allow researchers ‘(a) to capture as closely as possible the “real-time” occurrences or moments of change (in study variables); (b) to reduce recall bias; (c) to mitigate some forms of confounding by using participants as their own controls, and (d) to establish temporal precedence to strengthen causal inferences’ (p.747). DM can also be used not just to record ongoing behaviour patterns but also to examine how the co-variation between patterns of behaviour (e.g. stress and diet) varies as a function of an intervention.

In addition, using daily diaries permit researchers to use sophisticated statistical techniques (e.g. multilevel modeling) to examine day-to-day within-person effects together with the impact of between-person factors such as personality or gender. Finally, most of what we study and theorise about in psychology has a temporal element, changing with respect just to time and other variables. To fully appreciate the complexity and dynamics of human social and cognitive behaviour, diary methods are crucial.

* Daryl O’Connor is Senior Lecturer in Health Psychology at the University of LeedsD.B.O'Connor@leeds.ac.uk
* Eamonn Ferguson is Professor of Health Psychology at the University of Nottingham Eamonn.Ferguson@nottingham.ac.uk

References

ffleck, G., Zautra, A., Tennen, H. & Armeli, S. (1999). Multi-level daily process designs for consulting and clinical psychology: A preface for the perplexed. Journal of Consulting and Clinical Psychology, 67, 746–754.

Ferguson, E. (2005). The use of diary methods in clinical and health psychology. In J. Miles & P. Gilbert (Eds.) A handbook of research methods in clinical and health psychology (pp.111–124). Oxford: Oxford University Press.

Ferguson, E. (2008). Health anxiety moderates the daytime cortisol slope. Journal of Psychosomatic Research, 64, 484–494.

Ferguson, E., Moghaddam, N.G. & Bibby, P. (2007). Memory bias in health anxiety is related to the emotional valence of health related words. Journal of Psychosomatic Research, 62, 263–274.

Green, A.S., Rafaeli, E., Bolger, N. et al. (2006). Paper or plastic? Data equivalence in paper and electronic diaries. Psychological Methods, 11, 87–105.

Hox, J. (2002). Multilevel analysis. London: Lawrence Erlbaum.

Jones, F., O’Connor, D.B., Conner, M. et al. (2007). Impact of daily mood, work hours, and iso-strain variables on self-reported health behaviors. Journal of Applied Psychology, 92, 1731–1740.

Kahneman, D., Krueger A.B., Schkade D. et al (2004). A survey method for characterizing daily life experience. Science, 306, 1776–1780.

Kreft, I. & De Leeuw, J. (2006). Introducing multilevel modeling. London: Sage.

Moghaddam N.G. (2007). Modelling smoking motivation. PhD thesis, University of Nottingham.

Moghaddam N.G. & Ferguson, E. (2007). Smoking, mood regulation and personality: An event contingent-sampling exploration of potential models and moderation. Journal of Personality, 75, 451–478.

Newman, E., O’Connor, D.B., & Conner, M. (2007). Daily hassles and eating behaviour. Psychoneuroendocrinology, 32, 125–132.

O’Connor, D.B., Jones, F.A., Conner, M. et al. (2008). Effects of daily hassles and eating style on eating behaviour. Health Psychology, 27, S20–S31.

Robinson, W.S. (1950). Ecological correlations and the behavior of individuals. American Sociological Review, 15, 351–357.

Shiffman, S., Stione, A.A. & Hufford, MR. (2008). Ecological momentary assessment. Annual Review of Clinical Psychology, 4, 1–32.

Sober, E. & Wilson, S.D. (1998). Unto others: The evolution and psychology of unselfish behaviour. London: Harvard University Press.

Stone A.A., Schwartz, J.E., Schwarz, N., et al. (2006). A population approach to the study of emotion. Emotion, 6, 139–149.

Stone, A.A., Shiffman, S., Schwartz, J.E., et al. (2002). Patient non-compliance with paper diaries. British Medical Journal, 324, 1193–1194.

Tennen, H., Affleck, G., Coyne, J.C. et al. (2006). Paper and plastic in daily diary research: Comment on Green, Rafaeli, Bolger, Shrout, and Reis. Psychological Methods, 11, 112–118.

West, S.G. & Hepworth, J.T. (1991). Statistical issues in the study of temporal data: Daily experiences. Journal of Personality, 59, 609–661.