

## EHPS conference 2012

# Opportunities and challenges in real time data capture: Methods in Health Psychology Symposium II

**Martyn Jones\****University of Dundee***Daryl O'Connor***Leeds University***Joseph Schwartz***Stony Brook University***Derek Johnston***University of Aberdeen***David French***University of Manchester***Marijn de Bruin\****University of Amsterdam*

Following the successful first symposium on “Current issues in Randomized Controlled trials” at the EHPS conference in Crete (2011), the decision was made to organize a yearly state-of-the-art and thought-provoking symposium on methods in health psychology (a collaborative initiative by Marie Johnston and Marijn de Bruin). This year’s symposium was on a topic that has the potential to radically alter the way in which we collect our data, to enhance the validity and reliability of the data collected, to advance our statistical approaches and theories, and to allow the

design of individually-tailored interventions in real-time with exciting technological advances: Ecological Momentary Assessment (EMA). Fairly straightforward applications of this approach in other domains, like the electronic monitoring of medication intake behaviours in drug trials, have revolutionized models for understanding complex processes and opened up opportunities for intervention as problems occur in real-time within people’s normal, everyday context (e.g., Blaschke, Osterberg, Vrijens, & Urquhart, 2012; de Bruin, Hospers, van Breukelen, Kok, Koevoets, & Prins, 2010; Haberer, Robbins, Ybarra, et al., 2012).

The many advantages of using real time data capture techniques are best summed up in the words of Affleck and colleagues (1999), who argued that these approaches 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). Moreover, techniques such as EMA can be used not just to record on-going daily processes but also to examine how the co-variation between important behavioural processes (e.g., effects of daily stressors on food intake; see O’Connor et al., 2008) varies as a function of psychological interventions and different personality traits.

This symposium overview offers a summary of what EMA does best in health psychology (by Martyn Jones), and considers the challenges of linking individual difference and trait data which may be captured from different devices or gathered at different intervals (by Joseph Schwartz). This is further illustrated with studies that detail the value of examining health psychology theory in within-person as opposed to between-person designs (by Derek Johnston) and relate laboratory and field measures (by Daryl O’Connor). The symposium ends with a critical overview of the main challenges to conducting well-designed EMA studies, and our take on the future of EMA

Note: All authors contributed equally to this paper.

within health psychology (by David French and Marijn de Bruin, with contributions from the other authors).

### **What does EMA do best?**

This symposium is timely given the focus of health psychology on dynamic processes that underlie behaviours which are often studied using methods and frequencies of data collection that cannot capture such complexity. Retrospective approaches fail to capture such within-person variation using data that are aggregated and collapsed over time (Jones & Johnston, 2011). One might question the relevance and validity of findings using such measures.

Stone & Shiffman (1994) have advocated ecological momentary assessment (EMA) as an alternative. EMA allows collection of longitudinal data from a representative part of the participant's daily experience, in real time and in the participant's natural environment. For example, behavioural diaries can capture data provided repeatedly over time using paper (Takarangi, Garry, & Loftus, 2006), PDA handheld computers, or Smartphones (Johnston, Beedie, & Jones, 2006). Behavioural diaries can capture within-person data on cognition, affect, behaviour and even performance in the social world (Bolger, Davis, & Rafeli, 2003). EMA accounts are gathered more closely in time to the event and are less biased by heuristic, autobiographical memory strategies. Experiential Sampling Methodology (ESM) is a closely related approach, first developed in the Netherlands (e.g. Csikszentmihalyi & Larson, 1987).

Diary-based EMA approaches allow the provision of reminders for diary completion, called signal-contingent recording. Data may be gathered at a set interval or times of day (interval-contingent), or following particular incidents of interest (event-contingent). Diary-

based EMA methods generally have good or excellent compliance (Takarangi, Garry, & Loftus, 2006). Real-time longitudinal data may be combined from a range of devices like self-reports linked with physiological data. EMA allows testing of within-person variation in variables in a way that is difficult, often impossible, to achieve using retrospective measures and between-person (group level) designs.

The key benefit of this approach lies in the examination of events as they occur in their natural setting and allows the time course of the behaviour of interest to be modelled. Data collection can be scheduled to fit the respondent's day to explore the antecedents, correlates and consequences of daily experiences. Repeated, real-time EMA approaches are thought to improve the reliability and validity of data collection and to improve the quality of collected data (Piasecki, Hufford, Solhan, & Trull, 2007; Burton, Weller, & Sharpe, 2007). Data are time-stamped and entered into a spread-sheet automatically, with no error (Bolger, Davis, & Rafeli, 2003). Real time data collection can also be programmed to request information following state or physiological changes in the respondent (Picard & Liu, 2007). This approach has huge potential to provide accurate, real time evidence to assist in the therapeutic decisions of practitioners.

In sum, EMA approaches provide more data, probably better data and certainly different data than previously, allowing the application of more powerful analytic techniques to critical, real life questions than ever before.

### **Can intensive 1-day EMA monitoring be used to assess traits?**

EMA is advocated as a strategy for generating ecologically valid assessments of individuals' emotions, cognitions, behaviors, and physical

states. The extent to which these assessments provide reliable, reproducible and valid measures of individual differences is unknown. Their relationship to traditional trait questionnaires is also not known. In the Masked Hypertension Study, we collected electronic diary assessments of affect (e.g., anxiety, depression) every 30 minutes during waking hours for two 24h periods (i.e., the EMA measures), several months apart, and a variety of traditional questionnaires including the Spielberger Trait Anxiety Inventory (STAI, 1970) from 157 employed individuals. In a multilevel model (PROC Mixed in SAS; Schwartz & Stone 2007) we treated 24h subject-level means of EMA anxiety as a latent variable measure of trait anxiety and estimated 'EMA trait stability' (correlation of Time 1 means with Time 2 means), and the correlations of these EMA means with the STAI scores. The results showed that the 24h average of EMA anxiety is very stable ( $r = 0.91$ ), strongly suggesting that one day of intensive EMA monitoring is adequate for capturing individual differences. EMA mean anxiety, however, correlated only modestly with the STAI assessment of the same construct (correlations ranging from .21-.24). The question for future studies is now whether this implies that the EMA assessment (100-pt VAS rating of a single item, "anxious/tense") fails to capture important aspects of the STAI, or that the STAI suffers (more) compared with EMA from recall bias, the difficulty of mentally aggregating over time, reliance on semantic memory (self-image), and/or social desirability response bias. Hence, using EMA within this context was feasible, provided reliable and reproducible results, and raised interesting questions regarding the validity of a widely used traditional questionnaire measure.

### Testing theories within individuals

One of the powerful features of EMA is that it enables, indeed encourages, the repeated

measurement of the behaviour of individuals. This allows the testing of theory within individuals as well as the between-person tests that are more common in psychology. The importance of testing theory within individuals has been pointed out very vigorously by Molenaar (2004) who argues that most psychology theories should explain the behaviour of individuals, and variations in that behaviour. He has clearly described the danger of accepting the fallacy that a theory that explains differences between individuals will necessarily explain variations within an individual. We tested two of the main theories of work related stress, namely Karasek's (1979) Demand Control (DC) model and Siegrist's (1996) Effort Reward Imbalance (ERI) model, in a large sample of nurses measured every 90 minutes over 3 working shifts. We used PDA-based EMA that we have previously described (Johnston, Beedie & Jones, 2006) and multilevel modelling to conduct within-person tests. We showed that, as predicted, Negative Affect (NA) was greatest when Demand or Effort was high and this was moderated by Control and Reward. This indicates that the same determinants of work related stress operate within people as between. However EMA studies and multilevel modelling enables one to take this further and examine whether one's models apply to the individuals under study. We can show that while the DC model applied to virtually all nurses, the ERI Model appeared to be inappropriate for approximately 30%. This could not be established by traditional between-subject methods and illustrates how EMA studies can increase our understanding in unique ways. EMA studies lead us to ask new questions; in this case the challenge becomes to establish what other factors (environmental or personal) determine this difference between people. This is an important theoretical question which has practical implications since it suggests that some interventions, such as increasing reward,

may not be effective or might even be harmful for some individuals.

### Testing the efficacy of interventions in real time

In the current study, we used real time data capture techniques to explore whether a brief, easy to administer intervention, known as the written emotional disclosure paradigm (Pennebaker & Beall, 1986), was able to buffer against the effects of maladaptive rumination (i.e., brooding) on daily cardiovascular outcomes. Most previous research in this area has traditionally investigated the impact of written emotional disclosure on one-off, single assessment outcome measures (e.g., number of visits to general practitioners, frequency of cold symptoms, antibody concentrations). However, using innovative techniques allowed us to investigate whether the intervention was effective at lowering blood pressure and/or momentary levels of psychological stress on multiple occasions throughout normal working days. To this end, EMA was applied to ambulatory blood pressure assessments taken every 30 minutes for 12 hours on two weekdays following the intervention yielding 1339 observations from 55 participants. In addition, we were able to explore whether the efficacy of the intervention was moderated by important between-participant factors such as personality, whilst controlling for baseline levels of blood pressure and other potential confounders (e.g., physical activity level, body mass index). To our mind, these approaches are important as they generate a large number of observations in real-world contexts, which in turn increase ecological validity and confidence that any observed intervention effects are real and meaningful for health.

### Research priorities and future directions

The studies presented at this conference contrast key opportunities for advancing health

psychology research with a range of methodological challenges and questions. These opportunities and challenges are discussed in terms of the main features of EMA, namely those that relate to Ecological aspects, Momentary aspects, and Assessment aspects of such research.

First, studying measures taken within the context where behaviours, emotions, and cognitions actually occur may seem like such an obviously good thing as to not require stating. However, in literatures relating to common social cognition models, such as the Theory of Planned Behaviour (TPB, Ajzen, 1991), measurement within context is the exception rather than the norm. People are usually asked to complete questionnaires about behaviours such as alcohol consumption, exercise, and screening attendance either at home or in laboratories/classrooms. This lack of context has been shown to be misleading, at least in relation to alcohol consumption. Most TPB studies indicate that normative factors are not important in predicting drinking intentions and behaviour, whereas when people are asked to complete questionnaires about drinking behaviour in the context in which it occurs, i.e. in bars, then normative factors become very strongly predictive (Cooke & French, 2011; Cooke & French, 2012). There is a need for more consideration of context generally, and more examination of where completing measures in an inappropriate context may produce misleading results.

The momentary aspects of EMA may be potentially both a strength and a weakness. First, as already noted, intensive repeated measurement allows more appropriate tests of theory within people, rather than between people. A second potential advantage is that designs using such intensive measurement may make more sense to participants: the focus of

the research is on variation within themselves. This may partly explain the good levels of retention which EMA studies show, despite high respondent burden (Burton, Weller, & Sharpe, 2007). However, a potential downside is that the prospect of high respondent burden may lead to higher levels of selection bias in recruitment, relative to less intensive measurement. This issue warrants closer attention, especially for “convenience” samples not drawn from a clearly defined sampling frame.

The final aspect of EMA relates to Assessment, which is an enormous topic in its own right (Meier, 1986). Repeated measurements may make EMA liable to reactivity of measurement effects (for an overview see French & Sutton, 2010). Mean levels of variables such as reported pain do not alter across repeated measurement (Aaron et al, 2005), suggesting a lack of reactivity. However, it may be worth considering the framework of Golembiewski, Billingsley and Yeager (1976), who propose three kinds of change in measurement, namely alpha, beta and gamma change. Gamma change indicates a reconceptualization of the domain of interest, e.g. a person may initially consider “stress” to be synonymous with anxiety but later understand “stress” to be more composed of excess demand. Beta change indicates a recalibration in scaling, where e.g. a rating of “5” on the first occasion of measurement is not the same as a rating of “5” on the hundredth occasion. Alpha change indicates “true” change, which is what researchers are usually interested in. It is possible that the observed lack of change in mean scores across time may be taken at face value as due to a genuine lack of alpha, beta or gamma change, or it may indicate that such changes are occurring, but are not resulting in mean score changes. It would be worth examining such potential changes in response shift, given that repeated measurement is a core characteristic of EMA studies. It is

certainly the case that questionnaire measurement can involve the creation of new cognitions, as well as their assessment, as indicated by the use of “think aloud” methods (e.g. Darker & French, 2009).

Apart from reactive effects, other measurement challenges include making sure that any measures used possess good sensitivity to change. This psychometric criterion is even more important in within-person designs, and tends to be neglected relative to the criteria of reliability and validity. There is still a need to establish that single item measures are reliable and valid, which may be a challenge given that single item measures are often used to reduce respondent burden. In addition, asking questions within context should result in more valid self-reports. It does not require people to recall their past behaviour nor does it require people to mentally aggregate their experiences to produce an overall summary score, as is the case with more traditional approaches. However, a comparison of where there are differences between more traditional summary measures and EMA measures may shed light on the mental processes that people use to produce such summary scores. More generally, these advantages of EMA should not deflect attention away from the fact that people still need to interpret questions, retrieve appropriate information from memory and format their responses (Jobe et al, 1991). Although EMA may ameliorate some of these issues for self-report measurement, it is still important to examine how people approach the task of completing self-report measures.

## Conclusions

Repeated real-time EMA offers the possibility of providing better, more reliable, more context-specific data that are relatively unaffected by a range of recall and other biases. EMA is a flexible methodology that combines self-report

of behaviours, cognition and emotion with other forms of real-time data, e.g. physiological measurement. It allows the assessment of both trait-like individual difference variables and within-person changes allowing the possibility of providing innovative within-person tests of Health Psychology theory. EMA offers the possibility of testing mechanisms of change within complex interventions set in a real-world context, in a manner not previously possible. An exciting prospect is that it also allows us to intervene with unhealthy behaviours or cognitions as they occur in real-time. These aspects of EMA methodology are ripe for further development. EMA is not without its challenges, however. Further research is needed to detail the precise effects of perceived burden and selection bias, to establish the psychometric properties of short scale EMA measures, and to further elaborate the context specific effects of repeated measurement on reactivity, complexity and entrainment for EMA-derived outcomes. ■

## References

- Aaron, L., Turner, J., Lloyd, M., Brister, H., & Sawchuk, C. (2005). Electronic Diary Assessment of Pain Related Variables: Is Reactivity a Problem? *Journal of Pain*, *6*, 107-115. doi: 10.1016/j.jpain.2004.11.003
- Affleck, 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. doi: 10.1037/0022-006X.67.5.746
- Ajzen, I. (1991). The theory of planned behaviour. *Organizational Behavior and Human Decision Processes*, *50*, 179-211. doi: 10.1016/0749-5978(91)90020-T
- Blaschke, T.F., Osterberg, L., Vrijens, B., & Urquhart, J. (2012). Adherence to medications: insights arising from studies on the unreliable link between prescribed and actual drug dosing histories. *Annual Review of Pharmacology & Toxicology*, *10*, 275-30. doi: 10.1146/annurev-pharmtox-011711-113247
- de Bruin, M., Hospers, H.J., van Breukelen, G.J. P., Kok, G., Koevoets, W.M., & Prins J.M. (2010). Electronic monitoring-based counselling to enhance adherence among HIV-infected patients: A randomized controlled trial. *Health Psychology*, *29*, 421-428. doi: 10.1037/a0020335.
- Bolger, N., Davis, A., & Rafaeli, E. (2003). Diary Methods: Capturing Life as it is Lived. *Annual Review of Psychology*, *54*, 579-616. doi: 10.1146/annurev.psych.54.101601.145030
- Burton, C., Weller, D., & Sharpe, M. (2007). Are Electronic Diaries Useful for Symptoms Research? A Systematic Review. *Journal of Psychosomatic Research*, *62*, 553-61. doi: 10.1016/j.jpsychores.2006.12.022
- Cooke, R., & French, D.P. (2011). The role of context and timeframe in shaping relationships within the theory of planned behaviour. *Psychology and Health*, *26*, 1225-1240. doi: 10.1080/08870446.2011.572260
- Csikszentmihalyi, M., & Larson R. (1987). Validity and Reliability of the Experience Sampling Method. *Journal of Nervous and Mental Disease*, *175*, 526-37. doi: 10.1097/00005053-198709000-00004
- Darker, C.D., & French, D.P. (2009). What sense do people make of a Theory of Planned Behaviour questionnaire? A think-aloud study. *Journal of Health Psychology*, *14*, 861-871. doi: 10.1177/1359105309340983
- French, D.P., & Hevey, D. (2008). What do people think about when answering questionnaires to assess Unrealistic Optimism about skin cancer? A think aloud study. *Psychology, Health and Medicine*, *13*, 63-74. doi: 10.1080/13548500701243959
- French, D.P., & Sutton, S. (2010). Reactivity of measurement in health psychology: How much of a problem is it? What can be done about it? *British Journal of Health Psychology*, *15*, 453-468. doi: 10.1348/135910710X492341
- French, D.P., & Cooke, R. (2012). Using the Theory of Planned Behaviour to understand binge drinking: the importance of beliefs for developing interventions. *British Journal of Health Psychology*, *17*, 1-17. doi: 10.1111/j.2044-8287.2010.02010.x
- Golembiewski, R.T., Billingsley, K., & Yeager, S. (1976). Measuring change and persistence in human affairs: Types of change generated by OD designs. *Journal of Applied Behavioral Science*, *12*, 133-157. doi: 10.1177/002188637601200201

- Haberer, J.E., Robbins, G.K., Ybarra, M., Monk, A., Ragland, K., Weiser, S.D., Johnson, M.O., & Bangsberg, D.R. (2012). Real-time electronic adherence monitoring is feasible, comparable to unannounced pill counts, and acceptable. *AIDS & Behavior, 16*, 375-82. doi: 10.1007/s10461-011-9933-y
- Jobe, J.B., & Mingay, D.J. (1991). Cognition and survey measurement: History and overview. *Applied Cognitive Psychology, 5*, 175-192. doi: 10.1002/acp.2350050303
- Johnston, D., Beedie, A., & Jones, M. (2006). Using computerized ambulatory diaries for the assessment of job characteristics and work-related stress in nurses. *Work and Stress, 20*, 163-72. doi: 10.1080/02678370600902872
- Jones, M.C., & Johnston, D. (2011). Understanding phenomena in the real world: The case for real time data collection. *Journal of Health Services Research & Policy, 16*, 172-176. doi: 10.1258/jhsrp.2010.010016
- Karasek, R.A. (1979). Job demands, job decision latitude, and mental strain: implications for job redesign. *Administrative Science Quarterly, 24*, 285-307. doi: 10.2307/2392498
- Meier, S.T. (1994). *The chronic crisis in psychological measurement and assessment: A historical survey*. San Diego, CA: Academic Press.
- Molenaar, P.C.M. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement, 2*, 201-218. doi: 10.1207/s15366359mea0204\_1
- Piasecki, T., Hufford, M., Solhan, M., & Trull, R. (2007). Assessing clients in their natural environments with electronic diaries: rationale, benefits, limitations and barriers. *Psychological Assessment, 19*, 25-43. doi: 10.1037/1040-3590.19.1.25
- Picard, R., & Liu, K. (2007). Relative subjective count and assessment of interruptive technologies applied to the remote monitoring of stress. *International Journal of Human-Computer Studies, 65*, 361-75. doi: 10.1016/j.ijhcs.2006.11.019
- O'Connor, D.B., Jones, F.A., Conner, M., McMillan, B. & Ferguson, E. (2008). Effects of daily hassles and eating style on eating behaviour. *Health Psychology, 27*, 20-31. doi: 10.1037/0278-6133.27.1.S20
- van Oort, L., Schröder, C., & French, D.P. (2011). What do people think about when they answer the Brief Illness Perception Questionnaire? A 'think aloud' study. *British Journal of Health Psychology, 16*, 231-245. doi: 10.1348/135910710X500819
- Pennebaker, J.W., & Beall, S.K. (1986). Confronting a traumatic event: Toward an understanding of inhibition and disease. *Journal of Abnormal Psychology, 95*, 274-281. doi: 10.1037/0021-843X.95.3.274
- Schwartz, J. E., & Stone, A. A. (2007). The analysis of real-time momentary data: A practical guide. In A. Stone, S. Shiffman, A. A. Atienza & L. Nebeling (Eds.), *The Science of Real-Time Data Capture: Self-Report in Health Research* (pp. 76-113). Oxford: Oxford University Press.
- Siegrist, J. (1996). Adverse health effects of high-effort/low-reward conditions. *Journal Occupational Health Psychology, 1*, 27-41. doi: 10.1037/1076-8998.1.1.27
- Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1970). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA.: Consulting Psychologists Press.
- Stone, A., Shiffman, S. (1994) Ecological Momentary Assessment (EMA) in Behavioral Medicine. *Annals of Behavioral Medicine, 16*, 199-202.
- Takarangi, M., Garry, M., & Loftus, E. (2006) Dear diary, is plastic better than paper? I can't remember: Comment on Green, Rafaeli, Bolger, Shrout and Reis. *Psychological Methods, 11*, 119-22. doi: 10.1037/1082-989X.11.1.119

---

\*Corresponding authors:

**Martyn Jones**

**m.c.jones@dundee.ac.uk**

**Marijn de Bruin**

**m.debruin@uva.nl**