View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Maastricht University Research Portal

Maastricht University

Momentary predictors of compliance in studies using the experience sampling method

Citation for published version (APA):

Rintala, A., Wampers, M., Myin-Germeys, I., & Viechtbauer, W. (2020). Momentary predictors of compliance in studies using the experience sampling method. *Psychiatry Research*, *286*, [112896]. https://doi.org/10.1016/j.psychres.2020.112896

Document status and date:
Published: 01/04/2020

DOI:
10.1016/j.psychres.2020.112896

Document Version:
Publisher's PDF, also known as Version of record

Document license:
Taverne

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy
If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Download date: 06 Jan. 2021

Psychiatry Research 286 (2020) 112896

Contents lists available at ScienceDirect

Psychiatry Research

journal homepage: www.elsevier.com/locate/psychres

Momentary predictors of compliance in studies using the experience sampling method

Aki Rintala[a,*], Martien Wampers[a,b], Inez Myin-Germeys[a], Wolfgang Viechtbauer[a,c]

[a] Department of Neurosciences, Center for Contextual Psychiatry, KU Leuven, Kapucijnenvoer 33 bus 7001 (blok h), Leuven 3000, Belgium

[b] Universitair Psychiatrisch Centrum, UZ Leuven, Leuvensesteenweg 517, Kortenberg 3070, Belgium

[c] Department of Psychiatry and Neuropsychology, School for Mental Health and Neuroscience, Maastricht University, Universiteitssingel 40, Maastricht 6229 ER, Netherlands

ARTICLE INFO

Keywords:
Momentary assessment
Compliance
Experience sampling
Diary
Ecological momentary assessment

ABSTRACT

The influence of momentary experiences on compliance has not yet been studied extensively in diary methods such as the experience sampling method (ESM). This study investigated to what extent momentary experiences at the moment of responding (hereafter 'beep') can predict compliance in high frequency ESM protocols. Lagged-analyses were conducted using a pooled dataset of seven studies including 1,318 healthy volunteers and individuals with different mental health conditions. All studies used an ESM design of 10 beeps per day over 4 to 6 days. Overall compliance was 86% (to beeps where a subject was compliant at the previous beep). Results indicated that participants who reported higher positive affect overall were more compliant. Feeling disturbed by a beep, being outside the home, medication use, or longer inter-prompt interval decreased the chances of compliance to the subsequent beep. While participants with depression tended to be more compliant, chances to be compliant decreased in the evenings and over the course of the study days. When more beeps were missed consecutively, the chances to miss the subsequent beep increased. Findings suggest that disturbance of the beep, being outside the home, medication use, and inter-prompt interval might decrease the chances of compliance to the subsequent beep.

1. Introduction

The experience sampling method (ESM), also known as ecological momentary assessment (EMA), is a structured diary method to frequently assess momentary experiences in daily life, including mood, thoughts, symptoms, activity, location, stress, and social context (Ebner-Priemer and Trull, 2009; Trull and Ebner-Priemer, 2009). ESM is believed to be a feasible method to capture daily life experiences and behaviors in mental health research (Myin-Germeys et al., 2009, 2018), with advantages over traditional retrospective questionnaires including less recall bias and improved ecological validity (e.g., Myin-Germeys et al., 2009; Solhan et al., 2009; Trull and Ebner-Priemer, 2009). This has resulted in an exponential growth of ESM studies, especially in mental health research (Aan het Rot et al., 2012; Fahrenberg et al., 2007; Morren et al., 2009; Myin-Germeys et al., 2009; Shiffman et al., 2008).

Despite these advantages, the high frequency of the assessments can be a serious burden for participants (Delespaul, 1995; Palmier-Claus et al., 2011). In addition, compliance with the intended assessment scheme might be negatively impacted by having participants evaluate their experiences without the presence of a researcher or a health-care professional (Myin-Germeys et al., 2009).

Little is still known about factors that influence compliance. Previous methodological ESM studies that have studied predictors of overall compliance have mainly focused on personal, study, and time characteristics (Courvoisier et al., 2012; Hartley et al., 2014; Messiah et al., 2011; Ono et al., 2019; Rintala et al., 2019; Vachon et al., 2019). However, the influence of a subject's mental and physical state and the context surrounding a diary signal (hereafter 'beep') on compliance has been studied far less frequently (e.g., are participants less likely to respond to a beep if they are experiencing high levels of stress at the moment of the beep?). If a participant does not respond to a beep, this is generally assumed to be a random occurrence, that is, independent of the phenomenon one wants to capture, such as mood, symptomatology, or other daily life experiences (McLean et al., 2017). However, if a participant is more likely to be

* Corresponding author.
E-mail addresses: aki.rintala@kuleuven.be (A. Rintala), martien.wampers@kuleuven.be (M. Wampers), inez.germeys@kuleuven.be (I. Myin-Germeys), wolfgang.viechtbauer@maastrichtuniversity.nl (W. Viechtbauer).

https://doi.org/10.1016/j.psychres.2020.112896
Received 31 October 2019; Received in revised form 22 February 2020; Accepted 23 February 2020
Available online 29 February 2020
0165-1781/ © 2020 Elsevier B.V. All rights reserved.

compliant under certain circumstances, for example when he or she is in a positive mood or when alone, then the ecological validity of the ESM could be diminished (Scollon et al., 2003).

To study such associations, we would need information about the momentary state of participants not only surrounding beeps that were responded to, but also for planned assessments that were missed. However, we typically lack self-reported information for beeps that were missed unless passive sensors are used in combination with ESM (e.g., high heart rate (variability) as an indicator of stress or frequent changes in Bluetooth signals as a rough proxy for being outside the home). See Boukhechba et al. (2018) and Sarker et al. (2014) for example studies combining ESM with passive sensor data.

One way to circumvent this conundrum is by studying whether certain momentary experiences captured at a beep that was responded to are predictive of compliance to the subsequent beep. To date, only a few ESM studies have investigated such time-lagged associations (Silvia et al., 2013; Sokolovsky et al., 2014). Silvia et al. (2013) investigated the lagged effects of self-reported emotions (i.e., happiness, relaxation, enthusiasm, sadness, and anxiety) and experiences (i.e., current activity and fatigue) in a sample of 450 young healthy adults using an ESM protocol of 7 days with 8 random beeps per day (Silvia et al., 2013). The authors concluded that emotions and experiences reported at particular beeps are not a major source of non-compliance at the following beeps, and they only found that feeling enthusiasm predicted higher chances of nonresponse at the following beep. The authors hypothesized that participants might have felt more enthusiasm when engaging in certain activities, which might have interfered with responding to the following beep (Silvia et al., 2013). In a similar vein, Sokolovsky et al. (2014) observed that healthy adolescents (n = 461) reporting higher positive affect (PA) and being outside their home at a beep increased the risk of non-compliance at the following beep when using an ESM protocol of 7 days with 5 to 7 random beeps per day. The same authors also found that longer inter-prompt intervals between beeps predicted non-compliance. The association between PA and non-compliance at the following beep might result from decreased awareness of the beep when adolescents are in a high positive emotional state (Sokolovsky et al., 2014).

Although these studies provide some initial evidence that compliance might be associated with certain momentary states, further research into this topic is warranted. Moreover, to the best of our knowledge, previous studies have not yet investigated such associations in clinical populations. Therefore, the objective of this study is to investigate to what extent beep-level factors can predict compliance at the subsequent beep when using high frequency ESM protocols (4 to 6 study days with 10 semi-random assessments per day) in a large sample of participants with different mental health conditions.

2. Methods

2.1. Participants and procedure

Analyses were conducted using a pooled dataset of seven studies comprising a total of 1354 participants and 72,954 ESM observations. An overview of the studies in the pooled dataset is provided in supplementary file 1. Sufficient data for inclusion in the analysis was available for 1318 (97.3%) participants. Twenty-eight participants were excluded due to missing data in one or multiple predictors at each beep,[1] four participants were excluded due to missing age values, three participants were excluded due to missing information on mental health status, and one participant was excluded due to missing data for gender. Removal of the first beep of each day (to avoid lagged associations that span the night) and beeps where the prior beep (and hence all predictor variables) were missing (or where participants left one or more of the predictor variables of interest unanswered) left a total of 36,326 observations with 5665 non-response observations for the actual analyses. The final sample comprised 972 (73.6%) female and 346 (26.3%) male participants with a mean (SD, range) age of 32 (11.0, 16–65) years. Among the participants, 796 (60.4%) were classified as healthy subjects, 203 (15.4%) were at-risk for psychosis (e.g., having a first-degree relative with a psychotic disorder or having a high score on a subclinical psychosis scale), 190 (14.4%) were suffering from some form of psychosis, and 129 (9.8%) had a diagnosis of major depressive disorder or were suffering from residual symptoms. The local ethics committees approved all studies included in the pooled dataset.

All studies in the dataset used an ESM protocol where data were collected through a paper-and-pencil diary and a digital wristwatch for either 4, 5, or 6 consecutive study days (e.g., Collip et al., 2011; Collip et al., 2013a; Collip et al., 2013b; Geschwind et al., 2011; Wigman et al., 2015). Participants received 10 randomized beeps per day within 90-minute intervals between 7.30 a.m. and 10.30 p.m (i.e., 7.30 a.m. to 9.00 a.m., 9.00 a.m. to 10.30 a.m., and so on, with 10-minute time windows at the beginning and end of each interval where no beeps were allowed to occur, to avoid adjacent beeps occurring too close or too far apart in time). The beep times were programmed in a digital wristwatch, and these times were blinded from the participants. Participants reported the response time to every beep into the diary. After the study, a researcher matched the self-reported response times with the randomized schedule triggered by the digital wristwatch.

Most items in the ESM questionnaires were scored on a 7-point Likert scale (e.g., "I feel cheerful" with 1 = 'not at all' to 7 = 'very much'). A few questions were open-ended (e.g., "What am I doing?") or were scored on a bipolar scale (e.g., event-related question "This event was" -3 = 'very unpleasant' to +3 = 'very pleasant') or a binary scale (e.g., "I am alone" with response options 'Yes' or 'No'). A questionnaire example is presented in supplementary file 2.

2.2. Measures

2.2.1. Compliance

Compliance was defined as having a self-reported response time that fell within a time window of 5 min before and 15 min after the beep programmed into the wristwatch. This time window is considered acceptable when using a paper-and-pencil diary and a digital wristwatch, as participants might report a response time from a different device than the digital wristwatch itself (e.g., a kitchen or a cell phone clock) that is not synchronized with the wristwatch (Delespaul, 1995). In addition, participants might need time to stop their current activity to start answering the diary (e.g., driving). Compliance was scored dichotomously (0 = not answered within the time window, 1 = answered within the time window). This resulted in 40 to 60 values for each participant depending on the length of the study (i.e., 4 to 6 days).

2.2.2. Beep-level predictors

Beep-level variables were divided into six categories. An overview of these predictors and the corresponding hypotheses is provided in Table 1.

Emotional characteristics. Participants rated their positive (PA) and negative (NA) affect on a 7-point Likert scale, ranging from 1 (not) to 7 (very). PA was computed based on the mean of the items "*cheerful*", "*satisfied*", and "*relaxed*", while NA was based on the mean of the items "*uncertain*", "*lonely*", "*anxious*", "*guilty*", "*down*", and "*irritated*". The items for these composites were chosen because they were assessed in all of the included studies (the set of emotional characteristic items included in the various studies varied slightly) and because previous ESM studies have used the same or similar items to assess PA and NA

[1] These participants missed different items at each responded beep for unknown reasons. For example, one subject filled in affect-related items on some occasions, but did not fill in stress- or contextual-related items. In other words, these participants lacked consistently either one or several beep-level predictors of interest in their dataset.

Table 1
List of predictors at the beep level and corresponding hypotheses.

Predictor	Type	Compliance tends to be lower following beeps
Emotional characteristics		
Positive affect	Continuous	… where persons report less positive affect
Negative affect	Continuous	… where persons report high negative affect
Stress-related characteristics		
Activity stress	Continuous	… where persons report more activity stress
Social stress	Continuous	… where persons report more social stress
Event pleasantness	Continuous	… where persons report more event pleasantness
Contextual characteristics		
Location	Dichotomous	… where persons are outside their home
Social situation	Dichotomous	… where persons are with others
Disturbed by the beep	Continuous	… where persons are disturbed by the beep
Physical state characteristics		
Being active	Continuous	… where persons are more active
Being tired	Continuous	… where persons are more tired
Being unwell	Continuous	… where persons are more unwell
Being hungry	Continuous	… where persons are more hungry
Substance use		
Tobacco use	Dichotomous	… where persons report tobacco use since the last beep
Alcohol use	Dichotomous	… where persons report alcohol use since the last beep
Medication use	Dichotomous	… where persons report medication use since the last beep
Time characteristics		
Inter-prompt interval	Continuous	… that have a longer inter-prompt interval

(e.g.,Collip et al., 2013a; Habets et al., 2012; Lataster et al., 2013). Reliabilities of these composites at the person and beep levels were 0.90 and 0.76 for PA and 0.91 and 0.68 for NA, respectively.

Stress-related characteristics. Activity and social stress were defined according to previous ESM studies (e.g., Collip et al., 2013b; Myin-Germeys et al., 2001). Activity stress was assessed using the mean of the items "*I am skilled to do this activity*" (reverse scored), "*I prefer doing something else*", and "*This is a challenge*", while social stress was measured using the mean of the items "*I prefer being alone*" and "*I like this company*" (reverse scored). All items were rated on a 7-point Likert scale. Social stress was only measured when participants reported being with others and hence was not filled out at every beep.

For event pleasantness, participants were asked to recall the most important event that happened between the current and the previous beep. This event was then rated on a 7-point bipolar scale (-3 = 'very unpleasant' to +3 = 'very pleasant'). Event pleasantness was reverse coded so that higher scores reflect higher unpleasantness. Previous ESM studies have used this item to indicate "event-related stress" (Collip et al., 2011; Lataster et al., 2013).

Contextual characteristics. We defined a participant's location, social situation, and the level of disturbance due to the beeps as contextual characteristics. Location (0 = at home, 1 = outside home) and social situation (0 = with others, 1 = alone) were coded as dummy variables. The degree of disturbance due to a beep was measured with the item "*This beep disturbed me*" (7-point Likert scale).

Physical state characteristics. Items that reflect a participant's physical state were evaluated with 4 items, namely "*I am active*", "*I am unwell*", "*I am tired*", and "*I am hungry*" (7-point Likert scale).

Substance use characteristics. The questionnaires also included questions for tobacco, alcohol, and medication use. Participants were asked to indicate which substances they had consumed since the previous beep (each coded dichotomously).

Time characteristics. We included one time characteristic in the analyses, namely the inter-prompt interval (IPI). IPI was included as a continuous variable that reflects the time between the prior and the subsequent beep. To clarify, suppose a participant responded to beeps 4 and 5 on a particular day. Then the contextual variables at beep 4 were used to predict compliance at beep 5 (in this case, 'yes') and the amount of time (in minutes) between beep 4 and beep 5 was the IPI. Similarly, the contextual variables at beep 5 were used to predict compliance at beep 6 (in this case, 'no') and the IPI was the time between beeps 5 and 6. For the analyses, we centered the IPI values around the expected average time between beeps (i.e., 90 min) and rescaled these deviations to 15-minute units to avoid an overly small coefficient (i.e., this variable was coded (IPI – 90) / 15).

2.3. Statistical analyses

Analyses were conducted using multilevel mixed-effects logistic regression models with dichotomous and continuous predictors as time-lagged variables. First, we fitted an empty model to investigate the overall compliance in the pooled studies. Second, we fitted individual predictor models with one time-lagged variable at a time using age, gender, clinical population, chronological study days, time of the day, and weekdays as covariates. The coding of these covariates is described in supplementary file 3. We also included a covariate in these models that represents the number of missed beeps prior to the beep from which the lagged predictor values were obtained, as a way to approximate the autocorrelation in the outcome (which cannot be modeled directly by including the lagged outcome as predictor, since the analysis only included beeps where participants responded to the previous beep). To illustrate the coding of this predictor (denoted 'prior missed beeps'), see supplementary file 4.[2] Third, we fitted a multivariable model with all individual predictors and covariates except social stress. We decided to exclude social stress from the multivariable model, because it was the only item that was linked to a branching logic (i.e., only people engaged in social interactions during the lagged beep answered items related to social stress), and therefore many observations would be lost for this analysis. Finally, we fitted a model similar to the one presented by Sokolovsky et al. (2014) to see if we could replicate their findings using similar characteristics. In this replicate model, we only included the healthy participants as the sample studied by Sokolovsky et al. (2014) had no clinical diagnosis. All of these models included random effects for subjects within study, day number within

[2] We conducted an additional analysis where we investigated whether the number of prior missed beeps predicts compliance at the subsequent beep, controlling for age, gender, clinical population, chronological study days, time of the day, and day of the week. The model confirmed our hypothesis that a higher number of prior missed beeps increases the chances of non-compliance at the subsequent beep ($\chi^2 = 36.4$; $df = 1$; $OR = 0.894$; $p < .0001$), as would be expected under positive autocorrelation. Based on this model, we estimate that the odds of non-compliance increases by 12% (95% CI: 8% to 16%) for each additional prior missed beep.

subjects, and beep number within subjects, with the last two random effects entered as crossed random effects. The time-lagged predictors were included in the models as between-subject (BS) variables (by entering the subject-level means) and as within-subject (WS) variables (by entering the deviations from the subject-level means) simultaneously. This allowed us to investigate differences in compliance for example between individuals that report on average different amounts of negative mood throughout the study period (BS) and changes in compliance at the subsequent beep when a person reports different amounts of negative mood at particular moments (WS). Model coefficients were tested with Wald-type chi-square tests, with $\alpha = 0.05$ as the cutoff for significance. Accordingly, all reported confidence intervals (CIs) are 95% CIs. Analyses were conducted using R 3.5.1 (R Development Core Team, 2016) with packages *lme4* (Bates et al., 2015), *car* (Fox and Weisberg, 2011), and *multcomp* (Hothorn et al., 2008).

3. Results

Overall response compliance was 86% (CI: 85% to 86%). This value is higher than the overall compliance (78%) that has been reported for the individual studies included in this pooled dataset (Rintala et al., 2019), because our analyses were restricted to beeps where a subject was compliant at the previous beep. Descriptive information for the predictors and compliance rates for the different mental health conditions are presented in Table 2. The findings from the individual predictor and multivariable models are presented in Table 3, and the model replicating Sokolovsky et al. (2014) is presented in Table 4.

Table 2
Descriptive information for the lagged predictor variables (n = 1381).

Predictor	Number of observations	Mean (SD, range), unless otherwise mentioned
Positive affect	36,326	4.71 (1.29, 1 – 7)
Negative affect	36,326	1.53 (0.83, 1 – 7)
Activity stress	36,326	2.67 (1.11, 1 – 7)
Social stress	10,387	1.98 (1.16, 1 – 7)
Event pleasantness	36,326	2.74 (1.63, 1 – 7)
Location		
Being at home	20,123	55.4%
Being outside home	16,203	44.6%
Social situation		
Being with others	23,438	64.5%
Being alone	12,888	35.5%
Feeling disturbed by ESM beep	36,326	2.89 (1.88, 1 – 7)
Being active	36,326	3.59 (1.95, 1 – 7)
Being physically tired	36,326	2.72 (1.74, 1 – 7)
Being physically unwell	36,326	1.76 (1.34, 1 – 7)
Being hungry	36,326	2.13 (1.67, 1 – 7)
Tobacco use		
No	28,813	79.3%
Yes	7513	20.7%
Alcohol use		
No	34,702	95.5%
Yes	1624	4.5%
Medication use		
No	34,338	94.5%
Yes	1988	5.5%
Inter-prompt interval[a]	36,326	88.08 (31.84, 21 – 159)
Compliance rates[b]		
Healthy participants	85%	
At-risk for psychosis	85%	
Psychosis	85%	
Depression	89%	

SD = Standard deviation;.

[a] In the analyses, the inter-prompt interval (IPI) variable was coded (IPI – 90) / 15;.

[b] Compliance rates are higher than the overall compliance rates that have been reported previously for this pooled dataset (Rintala et al., 2019), because the present analyses were restricted to beeps where a participant was compliant at the previous beep.

3.1. Individual predictors of compliance

Covariates. In all individual predictor models, higher age and fewer missed beeps prior to the lagged response increased the chances to be more compliant at the subsequent beep (supplementary file 5). Gender was not a significant predictor of compliance in any of the models. In all individual predictor models except the one for social stress, participants with depression were more likely to be compliant compared to the healthy participants, the chances to be compliant was lower for every day compared with the first study day, for every beep compared with the second beep of the day, and during the weekday compared to the weekend. Participants who were at-risk for psychosis tended to be less compliant compared to the healthy participants, but only in the model examining social stress.[3]

Emotional characteristics. On the between-subject level, participants who reported higher PA overall were more compliant ($\chi^2 = 24.0$; $df = 1$; $OR = 1.169$; $p < .0001$) and participants who reported higher NA overall were more likely to be non-compliant ($\chi^2 = 17.9$; $df = 1$; $OR = 0.818$; $p < .0001$). On the within-subject level, higher PA increased the chances of compliance at the subsequent beep ($\chi^2 = 3.9$; $df = 1$; $OR = 1.032$; $p = .049$), while NA was not associated with compliance at the subsequent beep ($p = .77$).

Stress-related characteristics. Participants who reported higher activity stress ($\chi^2 = 13.6$; $df = 1$; $OR = 0.844$; $p = .0002$) or higher event unpleasantness ($\chi^2 = 10.2$; $df = 1$; $OR = 0.887$; $p = .001$) overall across the study were more likely to be non-compliant. Social stress was not associated with compliance (BS: $p = .90$; WS: $p = .84$).

Contextual characteristics. Participants feeling more disturbed by the beep had decreased chances to be compliant in both between- and within-subject levels (BS: $\chi^2 = 5.1$; $df = 1$; $OR = 0.949$; $p = .024$; WS: $\chi^2 = 21.3$; $df = 1$; $OR = 0.954$; $p < .0001$). On the within-subject level, participants who reported being outside their home ($\chi^2 = 23.2$; $df = 1$; $OR = 0.846$; $p < .0001$) were more likely to miss the subsequent beep. Being either alone or with others was not associated with compliance (BS: $p = .60$; WS: $p = .16$).

Physical state characteristics. On the between-subject level, participants who reported being physically tired ($\chi^2 = 14.5$; $df = 1$; $OR = 0.912$; $p = .0001$), physically unwell ($\chi^2 = 11.1$; $df = 1$; $OR = 0.903$; $p = .0009$), or hungry ($\chi^2 = 9.2$; $df = 1$; $OR = 0.895$; $p = .002$) overall across the study had increased chances to be non-compliant. However, being active was not associated with compliance (BS: $p = .94$; WS: $p = .80$).

Substance use characteristics. On the between-subject level, participants reporting less alcohol use overall across the study were more compliant ($\chi^2 = 4.8$; $df = 1$; $OR = 0.451$; $p = .028$). On the within-subject level, the use of tobacco increased the chances to be compliant at the subsequent beep ($\chi^2 = 4.0$; $df = 1$; $OR = 1.160$; $p = .047$). Medication use was not associated with compliance (BS: $p = .17$; WS: $p = .07$).

Time characteristics. Longer IPI increased the chances of non-compliance at the subsequent beep ($\chi^2 = 27.1$; $df = 1$; $OR = 0.963$; $p < .0001$).

3.2. Multivariable model

Based on the covariates in the multivariable model, we found that clinical status remained a significant predictor of compliance ($\chi^2 = 20.5$; $df = 3$; $p = .0001$), with higher compliance among

[3] The model investigating social stress differed from the rest of the models in the number of observations (10,387) and the number of subjects ($n = 536$). This difference is due to the fact that responses to the items measuring social stress were only collected in circumstances when the person was engaged in social interactions.

Table 3
Beep-level predictors of compliance.

Predictor	Level	Individual predictor model[†]				Multivariable model[†]			
		b	*Z*	*p*	*OR (95% CI)*	*b*	*Z*	*p*	*OR (95% CI)*
						2.333[‡]	6.494[‡]		
Emotional characteristics									
Positive affect	Intercept	1.264	7.049						
Between subjects		0.156	4.896	**< 0.0001**	**1.169 (1.098 to 1.244)**	0.102	2.373	**.018**	**1.107 (1.018 to 1.204)**
Within subjects		0.031	1.963	**.049**	**1.032 (1.000 to 1.064)**	0.039	1.916	.06	1.040 (0.999 to 1.082)
Negative affect	Intercept	2.290	19.419						
Between subjects		−0.201	−4.234	**< 0.0001**	**0.818 (0.746 to 0.898)**	−0.040	−0.611	.54	0.961 (0.846 to 1.091)
Within subjects		−0.008	−0.295	.77	0.992 (0.940 to 1.047)	0.027	0.801	.42	1.027 (0.962 to 1.097)
Stress-related characteristics									
Activity stress	Intercept	2.476	15.394						
Between subjects		−0.170	−3.688	**.0002**	**0.844 (0.771 to 0.924)**	−0.081	−1.577	.12	0.922 (0.833 to 1.020)
Within subjects		−0.007	−0.443	.66	0.993 (0.962 to 1.024)	0.007	0.414	.68	1.007 (0.974 to 1.042)
Social stress	Intercept	1.919	8.661						
Between subjects		−0.008	−0.131	.90	0.992 (0.880 to 1.118)	n.a.[†]			
Within subjects		0.006	0.198	.84	1.006 (0.945 to 1.072)	n.a.[†]			
Event pleasantness	Intercept	2.341	16.431						
Between subjects		−0.120	−3.189	**.001**	**0.887 (0.824 to 0.955)**	−0.035	−0.817	.41	0.966 (0.889 to 1.050)
Within subjects		−0.011	−1.126	.26	0.989 (0.969 to 1.009)	−0.007	−0.642	.52	0.993 (0.973 to 1.014)
Contextual characteristics									
Location	Being at home[§]	2.093	17.741						
Between subjects	Being outside home	−0.169	−1.402	.16	0.844 (0.666 to 1.070)	−0.217	−1.744	.08	0.805 (0.631 to 1.027)
Within subjects	Being outside home	−0.167	−4.814	**< 0.0001**	**0.846 (0.790 to 0.906)**	−0.160	−4.358	**< 0.0001**	**0.852 (0.793 to 0.916)**
Social situation	Being with others[§]	1.974	18.508						
Between subjects	Being alone	0.072	0.526	.60	1.074 (0.823 to 1.403)	0.104	0.743	.46	1.110 (0.843 to 1.461)
Within subjects	Being alone	0.049	1.407	.16	1.051 (0.981 to 1.126)	0.001	0.025	.98	1.001 (0.931 to 1.076)
Feeling disturbed by the ESM beep	Intercept	2.167	17.960						
Between subjects		−0.052	−2.252	**.024**	**0.949 (0.907 to 0.993)**	−0.009	−0.353	.72	0.991 (0.945 to 1.040)
Within subjects		−0.047	−4.613	**< 0.0001**	**0.954 (0.935 to 0.973)**	−0.040	−3.880	**.0001**	**0.960 (0.941 to 0.980)**
Physical state characteristics									
Being active	Intercept	1.996	14.972						
Between subjects		0.002	0.080	.94	1.002 (0.950 to 1.057)	−0.007	−0.239	.81	0.993 (0.938 to 1.051)
Within subjects		−0.002	−0.259	.80	0.998 (0.980 to 1.016)	0.006	0.623	.53	1.006 (0.987 to 1.026)
Being physically tired	Intercept	2.228	19.627						
Between subjects		−0.092	−3.814	**.0001**	**0.912 (0.870 to 0.956)**	−0.021	−0.690	.49	0.979 (0.922 to 1.040)
Within subjects		0.010	0.830	.41	1.010 (0.987 to 1.033)	0.015	1.205	.23	1.015 (0.991 to 1.040)
Being physically unwell	Intercept	2.147	20.301						
Between subjects		−0.102	−3.326	**.001**	**0.903 (0.851 to 0.959)**	−0.001	−0.028	.98	0.999 (0.922 to 1.082)
Within subjects		0.000	−0.012	.99	1.000 (0.970 to 1.030)	0.004	0.208	.84	1.004 (0.971 to 1.037)
Being hungry	Intercept	2.280	17.177						
Between subjects		−0.110	−3.036	**.002**	**0.895 (0.834 to 0.962)**	−0.072	−1.867	.06	0.930 (0.862 to 1.004)
Within subjects		−0.003	−0.271	.79	0.997 (0.978 to 1.017)	−0.003	−0.261	.79	0.997 (0.978 to 1.018)
Substance use									
Tobacco use	No[§]	2.032	20.824						
Between subjects	Yes	−0.131	−1.667	.10	0.877 (0.752 to 1.023)	−0.102	−1.275	.20	0.903 (0.772 to 1.056)
Within subjects	Yes	0.149	1.988	**.047**	**1.160 (1.002 to 1.343)**	0.137	1.822	.07	1.147 (0.990 to 1.328)
Alcohol use	No[§]	2.035	20.866						
Between subjects	Yes	−0.795	−2.199	**.028**	**0.451 (0.222 to 0.917)**	−0.710	−1.936	.053	0.492 (0.240 to 1.009)
Within subjects	Yes	−0.031	−0.411	.68	0.969 (0.835 to 1.125)	−0.044	−0.576	.56	0.957 (0.823 to 1.112)
Medication use	No[§]	2.021	20.862						
Between subjects	Yes	−0.411	−1.360	.17	0.663 (0.367 to 1.199)	−0.281	−0.915	.36	0.755 (0.414 to 1.378)
Within subjects	Yes	−0.134	−1.820	.07	0.875 (0.758 to 1.010)	−0.164	−2.224	**.026**	**0.849 (0.734 to 0.981)**
Time characteristics									
Inter-prompt interval	Intercept	2.002	20.557						
Between subjects		0.119	1.428	.15	1.127 (0.957 to 1.327)	0.102	1.227	.22	1.107 (0.941 to 1.304)
Within subjects		−0.038	−5.210	**< 0.0001**	**0.963 (0.949 to 0.976)**	−0.037	−5.045	**< 0.0001**	**0.964 (0.950 to 0.978)**
Covariates									
Age									
	Higher age	*Please see supplementary file 5*				0.046	1.550	.12	1.047 (0.998 to 1.110)
Gender									
	Male[§]								
	Female					0.065	0.970	.33	1.067 (0.936 to 1.216)
Clinical population									
	Healthy participants[§]								
	At-risk for psychosis					−0.001	−0.017	.99	0.999 (0.855 to 1.167)
	Psychosis					0.150	1.571	.12	1.162 (0.964 to 1.401)
	Depression					0.465	4.335	**< 0.0001**	**1.592 (1.290 to 1.965)**
Chronological study days									
	1[§]								
	2					−0.131	−2.571	**.010**	**0.877 (0.794 to 0.969)**
	3					−0.302	−5.940	**< 0.0001**	**0.739 (0.669 to 0.817)**
	4					−0.363	−7.053	**< 0.0001**	**0.696 (0.629 to 0.770)**
	5					−0.391	−7.480	**< 0.0001**	**0.677 (0.611 to 0.750)**
	6					−0.402	−6.563	**< 0.0001**	**0.669 (0.594 to 0.755)**

(*continued on next page*)

Table 3 (*continued*)

Predictor	Level	Individual predictor model[†]				Multivariable model[†]			
		b	*Z*	*p*	*OR (95% CI)*	*b*	*Z*	*p*	*OR (95% CI)*
						2.333[‡]	6.494[‡]		
Time of the day									
	9 a.m. – 10.30 a.m.[§]								
	10.30 a.m. – 12 p.m.					0.022	0.285	.78	1.022 (0.880 to 1.187)
	12 p.m. – 1.30 p.m.					0.003	0.041	.97	1.003 (0.865 to 1.163)
	1.30 p.m. – 3 p.m.					−0.173	−2.349	**.019**	**0.842 (0.729 to 0.972)**
	3 p.m. – 4.30 p.m.					−0.151	−2.021	**.043**	**0.859 (0.742 to 0.995)**
	4.30 p.m. – 6 p.m.					−0.151	−2.015	**.044**	**0.859 (0.742 to 0.996)**
	6 p.m. – 7.30 p.m.					−0.154	−2.055	**.040**	**0.857 (0.740 to 0.993)**
	7.30 p.m. – 9 p.m.					−0.229	−3.083	**.002**	**0.796 (0.688 to 0.920)**
	9 p.m. – 10.30 p.m.					−0.467	−6.351	< **0.0001**	**0.627 (0.542 to 0.724)**
Day of the week									
	Weekdays[§]								
	Weekend					0.070	2.021	**.043**	**1.073 (1.002 to 1.148)**
Number of prior missed beeps[¶]									
						−0.113	−6.003	< **0.0001**	**0.893 (0.861 to 0.927)**

Significant results are displayed in bold;.
[†] = Total number of observations across all models was 36,326 with n = 1318 (except for social stress; number of observations was 10,387 with n = 536); b = estimated value of coefficient; Z = Z-value; p = p-value; *OR* = Odds ratio; *95% CI* = 95% confidence interval;.
[‡] = Intercept coefficient and Z-value for the multivariable model; n.a. = not applicable; Between subjects = Subjects' average value for variable across the study; Within subjects = Subjects' momentary fluctuations in variable;.
[§] = Reference category;.
[¶] = Number of missed beeps prior to the beep from which the lagged predictor values were obtained.

participants with depression compared to the healthy participants (*OR* = 1.592; p < .0001). However, age was no longer a significant predictor of compliance compared to the individual predictor models (p = .12). Other main findings remained significant, namely that the chances of responding were lower every day when compared to the first study day (χ^2 = 91.2; *df* = 5; *OR* = 0.877 to 0.669; p < .0001) and also when comparing beeps between the 5th (1.30 p.m. to 3p.m.) and the 10th (9 p.m. to 10.30 p.m.) interval within a day with the second (9 a.m. to 10.30 a.m.) beep of the day (χ^2 = 77.2; *df* = 8; *OR* = 0.842 to 0.627; p < .0001). When more beeps were missed consecutively, the chances to miss the subsequent beep also increased (χ^2 = 36.0; *df* = 1; *OR* = 0.893; p < .0001). Responding over the weekend remained a significant predictor of compliance compared to weekdays (χ^2 = 4.1; *df* = 1; *OR* = 1.073; p = .043).

On the between-subject level, only higher PA (χ^2 = 5.6; *df* = 1; *OR* = 1.107; p = .018) remained a predictor of higher compliance in the multivariable model (Table 3). On the within-subject level, feeling more disturbed by the beep (χ^2 = 15.1; *df* = 1; *OR* = 0.960; p = .0001), being outside the home (χ^2 = 19.0; *df* = 1; *OR* = 0.852; p < .0001), and longer IPI (χ^2 = 25.5; *df* = 1; *OR* = 0.964; p < .0001) remained significant predictors of non-compliance at the subsequent beep. In addition, use of medication was a predictor of non-compliance at the subsequent beep (χ^2 = 4.9; *df* = 1; *OR* = 0.849; p = .026), which was not observed in the individual predictor model.

3.3. Replicate multivariable model

In agreement with Sokolovsky et al. (2014), a longer IPI (χ^2 = 14.8; *df* = 1; *OR* = 0.964; p = .0001) and being outside the home (χ^2 = 10.4; *df* = 1; *OR* = 0.866; p = .001) both decreased the chances of compliance at the subsequent beep (Table 4). In our replicate multivariable model, higher PA increased the chances of compliance on both the between- and within-subject levels (BS: χ^2 = 8.3; *df* = 1; *OR* = 1.162; p = .004; WS: χ^2 = 6.1; *df* = 1; *OR* = 1.066; p = .013), whereas Sokolovsky et al. (2014) found a negative association between PA and compliance at the within-subject level (and no between-subject association). Compliance decreased over the course of the study days in our replicate multivariable model (χ^2 = 50.2; *df* = 1; *OR* = 0.910; p < .0001), but this was not a significant predictor in the Sokolovsky et al. (2014) model.

4. Discussion

The objective of this study was to investigate beep-level predictors of compliance when using a high-frequency ESM protocol (with 10 beeps per day) in a large sample of participants with different mental health conditions. Our results suggest that most momentary characteristics measured at a particular beep, such as negative mood, stress, and contextual, physical state, or substance use characteristics, are not associated with compliance at the subsequent beep, which provides some support for the assumption that missing data could be considered missing at random. However, we cannot rule out systematic missingness based on our findings, as some momentary experiences were associated with compliance.

On the between-subject level, persons with higher PA overall tended to be more compliant in all of our models. Higher PA was also a predictor of compliance at the within-subject level in our individual model and in our replicate model of the one presented by Sokolovsky et al. (2014). However, Sokolovsky et al. (2014) found the opposite, namely that higher PA (at the within-subject level) increased the chances of *non-compliance*. Sokolovsky et al. (2014) noted that adolescents with a highly positive emotional state might be less able to recognize the beep in their natural environment (Sokolovsky et al., 2014). Another study by Silvia et al. (2013) observed an association between "*I feel enthusiastic right now*" and non-compliance, but found no association between compliance and single items related to PA, such as "*I feel happy right now*" and "*I feel relaxed right now*". Silvia et al. (2013) noted that persons feeling enthusiasm were perhaps engaged in activities that interfered with responding to the beep. Contradictory findings between our and previous studies might be explained by differences in the PA items used, but also in the study sample and age. Sokolovsky et al. (2014) and Silvia et al. (2013) included healthy adolescents (n = 461) and young adults between 14 and 19 years (n = 450) while our study included participants with different mental health conditions that were between 16 and 65 years (n = 1318).

Our findings indicate that a higher degree of disturbance by the beep increased the chances of non-compliance at the subsequent beep. To our knowledge, this is the first study to investigate and demonstrate that the degree of disturbance by the beep is a predictor of compliance at the within-subject level. Researchers are encouraged to discuss with their participants how to deal with possible disturbance by the beeps

Table 4
Replicate model of Sokolovsky et al. (2014) on beep-level predictors of compliance using a subset of healthy participants.

Predictor	Level	Multivariable model of compliance				Predictor	Level	Sokolovsky et al. (2014) model of non-compliance			
		b	*Z*	*p*	*OR (95% CI)*			*b*	*t*	*p*	*OR (95% CI)*
	Intercept[†]	1.416	3.900				Intercept[‡]	−2.82	−7.20		
Study days[§]	5 days	−0.095	−7.084	**< 0.0001**	**0.910 (0.886 to 0.934)**	Study days[§]	8 days	0.022	1.46	.14	1.018 (0.999 to 1.037)
Day of the week[¶]	Weekend	0.052	1.125	.26	1.053 (0.962 to 1.152)	Weekday[¶]	Weekend	0.068	0.92	.36	1.067 (0.929 to 1.227)
Social[¶]	Being alone	0.032	0.666	.51	1.033 (0.940 to 1.135)	Social[¶]	With friends	0.028	0.37	.71	1.031 (0.895 to 1.189)
PA (BS)		0.151	2.882	**.004**	**1.162 (1.049 to 1.288)**	PA (BS)		0.019	0.46	.65	1.018 (0.942 to 1.101)
PA (WS)		0.064	2.474	**.013**	**1.066 (1.013 to 1.121)**	PA (WS)		0.063	2.44	**.015**	**1.062 (1.012 to 1.115)**
NA (BS)		−0.107	−1.096	.27	0.898 (0.742 to 1.088)	NA (BS)		0.062	1.75	.08	1.061 (0.993 to 1.134)
NA (WS)		0.095	1.837	.07	1.100 (0.994 to 1.218)	NA (WS)		−0.024	−1.06	.29	0.977 (0.937 to 1.020)
IPI (per 15 min)		−0.036	−3.846	**< 0.0001**	**0.964 (0.947 to 0.982)**	IPI (per minute)		0.003	2.84	**.005**	**1.003 (1.001 to 1.004)**
Hungry		0.003	0.198	.84	1.003 (0.978 to 1.028)	Hunger		0.019	1.85	.06	1.021 (0.993 to 1.050)
Gender[¶]	Female	0.084	0.824	.41	1.088 (0.890 to 1.329)	Gender[¶]	Male	0.206	2.50	**.012**	**1.219 (1.044 to 1.424)**
Tobacco[¶]	Yes	0.048	0.625	.53	1.050 (0.902 to 1.222)	Tobacco	Log 30-day smoking rate	0.027	1.06	.29	1.026 (0.978 to 1.077)
Alcohol[¶]	Yes	−0.162	−1.658	.10	0.850 (0.702 to 1.030)	Alcohol use		0.044	1.58	.12	1.043 (0.990 to 1.099)
Location[¶]	Outside home	−0.144	−3.228	**.001**	**0.866 (0.793 to 0.945)**	Location	School (vs. home)	0.419	4.63	**< 0.001**	**1.498 (1.262 to 1.777)**
							Other (vs. home)	0.351	4.19	**< 0.001**	**1.402 (1.197 to 1.642)**

Significant results are displayed in bold; *b* = estimate value of coefficient; *Z* = *Z*-value; *p* = *p*-value; *OR* = Odds ratio; *95% CI* = 95% confidence interval; *t* = *t*-value;.
[†] = Our full model consisted of 19,797 observations from 796 healthy adults (16 – 65 year old);.
[‡] = Full model from Sokolovsky et al. (2014) study consisted 461 healthy adolescents;.
[§] = Study days predictor was analyzed as continuous;.
[¶] = predictor was treated as a dummy variable in the models; BS = Between subject level (subjects' average affect across the study); PA = Positive affect; WS = Within subject level (subjects' momentary fluctuations in affect); NA = Negative affect; IPI = Inter-prompt interval.

during the study briefing session to minimize negativity towards the method.

For location, our findings confirmed the hypothesis that being outside the home at a beep may increase the risk of not responding to the subsequent beep, similarly to what was observed by Sokolovsky et al. (2014). Presumably, when a subject is outside the home, chances are increased that the participant is still outside the home at the following beep, which in turn tends to coincide with circumstances during which the participant might not hear the beep or the participant is engaged in an activity where the beep cannot be responded to (e.g., during sport activities or while driving a car). As a recommendation, researchers are encouraged to discuss with their participants the importance of keeping the device within reach at all times and to think of safe strategies that would allow them to respond during particular moments if possible (e.g., by safely pulling over while driving a car).

Medication use was a predictor of non-compliance at the subsequent beep. Although ESM has been shown to be suitable for studying medication use and its associations with pain, fatigue, or mood (Smyth and Smyth, 2003; Stone et al., 1997; Wichers et al., 2009), its association with compliance has not yet been investigated. Our findings may suggest that adherence to fill in the beeps may diminish if participants are temporarily ill or medicated for a certain time of the day. However, this finding needs to be interpreted with caution as the dataset does not contain information on the level or the type of reported medication (hence, participants might have taken an antidepressant medication or just a painkiller). More methodological studies are warrented to confirm this association.

Our findings also suggest that a longer IPI between beeps might increase the risk of non-compliance, which is in agreement with the study by Sokolovsky et al. (2014). Although a high-frequency protocol may increase burden, both studies suggests that higher sampling frequencies may not alter compliance. Sokolovsky et al. (2014) noted that too infrequent beeps may lead to disengaging from the device and forgetting to follow the study protocol. As a further explanation for this finding, we hypothesize that when a participant has responded to a beep and the IPI is short for the subsequent beep, chances are higher that the context remains unchanged. Therefore, the probability to respond to the subsequent beep may also be higher. On the other hand, if the IPI is longer for the following beep, chances that the context has changed increase, and hence, the probability to respond to the subsequent beep decreases.

We found that participants with depression tended to be more compliant compared to the healthy participants in circumstances where the prior beep was responded to (with an overall compliance of 89% versus 85%, respectively; see Table 2) and this difference was also found to be significant in our individual and multivariable models. Our previous study of the same dataset indicated no differences in overall compliance rates between depressed (80%) and healthy (83%) participants (Rintala et al., 2019). This discrepancy can be explained by the fact that our current study only included observations that were preceded by a beep that was responded to, instead of investigating the full dataset as in our previous study. Moreover, although statistically significant, the difference in compliance rates in our current dataset was only 4 percentage points in favor of participants with depression, which one may argue is too small to base further conclusions on. More studies are therefore recommended to investigate momentary predictors of compliance in different clinical populations.

Other covariate findings were in line with previous studies (e.g., Ono et al., 2019; Rintala et al., 2019), namely that the chances to be compliant decreased in the evenings and over the course of the study days. For example, the odds to be compliant were the lowest for beeps between 7.30 p.m. and 9 p.m. and for beeps between 9 p.m. and 10.30 p.m.. Similarly, the odds to be compliant were the lowest on the fifth day and on the sixth day. Finally, our 'pseudo autocorrelation' variable indicated that the odds of compliance decreased for each additional beep missed prior to the lagged (i.e., responded to) beep. This confirmed our hypothesis that if a participant is not responding to the first beeps of the day, he or she also has an increased risk of missing the beeps that will follow later during the day. This may be difficult to prevent, but we can imagine future studies giving additional reminders as needed to participants based on the response pattern to the first beeps of the day, which might be able to increase the chances of being engaged in the study later on that same day.

To our knowledge, this is the first methodological study to examine beep-level predictors of compliance when using a high-frequency ESM protocol in individuals with different mental health conditions. At the same time, we need to consider some limitations. Our study focused only on data collected using a paper-and-pencil approach, a method that is nowadays often replaced by a dedicated device or a smartphone application. Using a similar ESM protocol with an electronic data collection method is recommended to replicate and verify our findings. Finally, our analyses can only provide indirect evidence about the association between momentary states and compliance, because we could only examine associations between predictors observed at beeps where participants responded and the subsequent beep. Regardless of these limitations, this study provides unique information on beep-level predictors of compliance using a high-frequency ESM protocol.

Our findings suggest that disturbance by the beep, location, medication use, and longer inter-prompt intervals might increase the chances of non-compliance at the subsequent beep on the within subject-level, but most momentary characteristics measured at a particular beep are not associated with compliance. A high frequency ESM protocol can therefore be considered a viable option for capturing momentary daily life experiences and contexts.

Funding

This study was supported by the Remote Assessment of Disease and Relapse – Central Nervous System (RADAR-CNS) research program from the Innovative Medicines Initiative 2 Joint Undertaking under a grant agreement number 115902. This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation program and the European Federation of Pharmaceutical Industries and Associations (EFPIA). Inez Myin-Germeys was funded by the Fonds voor Wetenschappelijk Onderzoek (FWO) Odysseus grant (GOF8416N).

CRediT authorship contribution statement

Aki Rintala: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing - review & editing, Visualization, Project administration. **Martien Wampers:** Conceptualization, Methodology, Writing - review & editing, Supervision. **Inez Myin-Germeys:** Writing - review & editing, Supervision, Funding acquisition. **Wolfgang Viechtbauer:** Methodology, Software, Writing - review & editing, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank the research group of ESM MERGE for providing the data for this study.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.psychres.2020.112896.

References

Aan het Rot, M., Hogenelst, K., Schoevers, R.A., 2012. Mood disorders in everyday life: a systematic review of experience sampling and ecological momentary assessment studies. Clin. Psychol. Rev. 32, 510–523. https://doi.org/10.1016/j.cpr.2012.05.007.

Bates, D., Mächler, M., Bolker, B., Walker, S., 2015. Fitting linear mixed-effects models using lme4. J. stat. Softw 67, 51. https://doi.org/10.18637/jss.v067.i01.

Boukhechba, M., Cai, L., Chow I., P., Fua, K., Gerber S., M., Teachman A., B., Barnes E., L., 2018. Contextual Analysis to Understand Compliance with Smartphone-based Ecological Momentary Assessment. Proceedings of the 12th EAI International Conference on Pervasive Computing Technologies for Healthcare 232–238. https://doi.org/10.1145/3240925.3240967.

Collip, D., Habets, P., Marcelis, M., Gronenschild, E., Lataster, T., Lardinois, M., Nicolson, N.a, Myin-Germeys, I., 2013a. Hippocampal volume as marker of daily life stress sensitivity in psychosis. Psychol. Med 43, 1377–1387. https://doi.org/10.1017/S003329171200219X.

Collip, D., Nicolson, N.a, Lardinois, M., Lataster, T., van Os, J., Myin-Germeys, I., 2011. Daily cortisol, stress reactivity and psychotic experiences in individuals at above average genetic risk for psychosis. Psychol. Med 41, 2305–2315. https://doi.org/10.1017/S0033291711000602.

Collip, D., Wigman, J.T.W., Myin-Germeys, I., Jacobs, N., Derom, C., Thiery, E., Wichers, M., van Os, J., 2013b. From epidemiology to daily life: linking daily life stress reactivity to persistence of psychotic experiences in a longitudinal general population study. PLoS ONE 8, e62688. https://doi.org/10.1371/journal.pone.0062688.

Courvoisier, D.S., Eid, M., Lischetzke, T., 2012. Compliance to a cell phone-based ecological momentary assessment study: the effect of time and personality characteristics. Psychol Assess 24, 713–720. https://doi.org/10.1037/a0026733.

Delespaul, P., 1995. Assessing Schizophrenia in Daily life: the Experience Sampling Method. Universitaire Pers Maastricht, Maastricht.

Ebner-Priemer, U.W., Trull, T.J., 2009. Ecological momentary assessment of mood disorders and mood dysregulation. Psychol. Assess. 21, 463–475. https://doi.org/10.1037/a0017075.

Fahrenberg, J., Myrtek, M., Pawlik, K., Perrez, M., 2007. Ambulatory assessment -monitoring behavior in daily life settings: a behavioral-scientific challenge for psychology. Eur. J. Psychol. Assess. 23, 206–213. https://doi.org/10.1027/1015-5759.23.4.206.

Fox, J., Weisberg, S., 2011. An r companion to applied regression. Sage Publications.

Geschwind, N., Peeters, F., Drukker, M., Van Os, J., Wichers, M., 2011. Mindfulness training increases momentary positive emotions and reward experience in adults vulnerable to depression: a randomized controlled trial. J. Consult. Clin. Psychol. 79, 618–628. https://doi.org/10.1037/a0024595.

Habets, P., Collip, D., Myin-Germeys, I., Gronenschild, E., van Bronswijk, S., Hofman, P., Lataster, T., Lardinois, M., Nicolson, N.a, van Os, J., Marcelis, M., 2012. Pituitary volume, stress reactivity and genetic risk for psychotic disorder. Psychol. Med 42, 1523–1533. https://doi.org/10.1017/S0033291711002728.

Hartley, S., Varese, F., Vasconcelos D., e S., Udachina, A., Barrowclough, C., Bentall, R.P., Lewis, S.W., Dunn, G., Haddock, G., Palmier-Claus, J., 2014. Compliance in experience sampling methodology: the role of demographic and clinical characteristics. Psychosis 6, 70–73. https://doi.org/10.1080/17522439.2012.752520.

Hothorn, T., Bretz, F., Westfall, P., 2008. Simultaneous inference in general parametric models. Biometrical J 50, 346–363. https://doi.org/10.1002/bimj.200810425.

Lataster, T., Valmaggia, L., Lardinois, M., Van Os, J., Myin-Germeys, I., 2013. Increased stress reactivity: a mechanism specifically associated with the positive symptoms of psychotic disorder. Psychol. Med. 43, 1389–1400. https://doi.org/10.1017/S0033291712002279.

McLean, D.C., Nakamura, J., Csikszentmihalyi, M., 2017. Explaining system missing: missing data and experience sampling method. Soc. Psychol. Personal. Sci 8, 434–441. https://doi.org/10.1177/1948550617708015.

Messiah, A., Grondin, O., Encrenaz, G., 2011. Factors associated with missing data in an experience sampling investigation of substance use determinants. Drug Alcohol Depend 114, 153–158. https://doi.org/10.1016/j.drugalcdep.2010.09.016.

Morren, M., Dulmen, S., Ouwerkerk, J., Bensing, J., 2009. Compliance with momentary pain measurement using electronic diaries: a systematic review. Eur. J. Pain 13, 354–365. https://doi.org/10.1016/j.ejpain.2008.05.010.

Myin-Germeys, I., Kasanova, Z., Vaessen, T., Vachon, H., Kirtley, O., Viechtbauer, W., Reininghaus, U., 2018. Experience sampling methodology in mental health research: new insights and technical developments. World Psychiatry 17, 123–132. https://doi.org/10.1002/wps.20513.

Myin-Germeys, I., Oorschot, M., Collip, D., Lataster, J., Delespaul, P., van Os, J., 2009. Experience sampling research in psychopathology: opening the black box of daily life. Psychol. Med. 39, 1533–1547. https://doi.org/10.1017/S0033291708004947.

Myin-Germeys, I., van Os, J., Schwartz, J.E., Stone, A.A., Delespaul, P.a, 2001. Emotional reactivity to daily life stress in psychosis. Arch. Gen. Psychiatry 58, 1137–1144. https://doi.org/10.1001/archpsyc.58.12.1137.

Ono, M., Schneider, S., Junghaenel, D.U., Stone, A.A., 2019. What affects the completion of ecological momentary assessments in chronic pain research? an individual patient data meta-analysis. J Med Internet Res 2019 21 (2), e11398. https://doi.org/10.2196/11398.

Palmier-Claus, J.E., Myin-Germeys, I., Barkus, E., Bentley, L., Udachina, A., Delespaul, P., Lewis, S.W., Dunn, G., 2011. Experience sampling research in individuals with mental illness: reflections and guidance. Acta Psychiatr. Scand. 123, 12–20. https://doi.org/10.1111/j.1600-0447.2010.01596.x.

R. Development Core Team, 2016. R: a language and environment for statistical computing, R Foundation for Statistical Computing Vienna Austria. 10.1038/sj.hdy.6800737.

Rintala, A., Wampers, M., Myin-Germeys, I., Viechtbauer, W., 2019. Response compliance and predictors thereof in studies using the experience sampling method. Psychol. Assess. 31, 226–235. https://doi.org/10.1037/pas0000662.

Sarker, H., Sharmin, M., Ali A., A., Rahman M., M., Bari, R., Hossain M., S., Kumar, S., 2014. Assessing the Availability of Users to Engage in Just-in-Time Intervention in the Natural Environment. Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing 909–920. https://doi.org/10.1145/2632048.2636082.

Scollon, C.N., Chu, K.-P., Diener, E., 2003. Experience sampling: promises and pitfalls, strengths and weaknesses. J. Happiness Stud. 4, 5–34. https://doi.org/10.1023/A:1023605205115.

Shiffman, S., Stone, A.A., Hufford, M.R., 2008. Ecological momentary assessment. Annu. Rev. Clin. Psychol 4, 1–32. https://doi.org/10.1002/9781118384404.ch20.

Silvia, P.J., Kwapil, T.R., Eddington, K.M., Brown, L.H., 2013. Missed beeps and missing data: dispositional and situational predictors of nonresponse in experience sampling research. Soc. Sci. Comput. Rev. 31, 471–481. https://doi.org/10.1177/0894439313479902.

Smyth, J.M., Smyth, J.M., 2003. Ecological momentary assessment research in behavioral medicine. J. Happiness Stud. 4, 35–52. https://doi.org/10.1023/a:1023657221954.

Sokolovsky, A.W., Mermelstein, R.J., Hedeker, D., 2014. Factors predicting compliance to ecological momentary assessment among adolescent smokers. Nicotine Tob. Res. 16, 351–358. https://doi.org/10.1093/ntr/ntt154.

Solhan, M.B., Trull, T.J., Jahng, S., Wood, P.K., 2009. Clinical assessment of affective instability: compating ema indices, questionnaire reports, and retrospective recall. Psychol Assess 21, 425–436. https://doi.org/10.1037/a0016869.

Stone, A.A., Broderick, J.E., Porter, L.S., Kaell, A.T., 1997. The experience of rheumatoid arthritis pain and fatigue: examining momentary reports and correlates over one week. Arthritis Rheum.10.1002/art.1790100306.

Trull, T.J., Ebner-Priemer, U.W., 2009. Using experience sampling methods/ecological momentary assessment (ESM/EMA) in clinical assessment and clinical research: introduction to the special section. Psychol. Assess. 21, 457–462. https://doi.org/10.1037/a0017653.

Vachon, H., Viechtbauer, W., Rintala, A., Myin-Germeys, I., 2019. Compliance and retention with the experience sampling method over the continuum of severe mental disorders: meta-Analysis and recommendations. J Med Internet Res 21, e11398. https://doi.org/10.2196/14475.

Wichers, M.C., Barge-Schaapveld, D.Q.C.M., Nicolson, N.A., Peeters, F., De Vries, M., Mengelers, R., Van Os, J., 2009. Reduced stress-sensitivity or increased reward experience: the psychological mechanism of response to antidepressant medication. Neuropsychopharmacology 34, 923–931. https://doi.org/10.1038/npp.2008.66.

Wigman, J.T.W., van Os, J., Borsboom, D., Wardenaar, K.J., Epskamp, S., Klippel, A., Viechtbauer, W., Myin-Germeys, I., Wichers, M., 2015. Exploring the underlying structure of mental disorders: cross-diagnostic differences and similarities from a network perspective using both a top-down and a bottom-up approach. Psychol. Med. 45, 2375–2387. https://doi.org/10.1017/S0033291715000331.