Nicotine & Tobacco Research, Volume 16, Number 3 (March 2014) 351–358

ORIGINAL INVESTIGATION

Factors Predicting Compliance to Ecological Momentary Assessment Among Adolescent Smokers

Alexander W. Sokolovsky MA[1,2], **Robin J. Mermelstein** PhD[1,2], **Donald Hedeker** PhD[1,3]

[1]*Institute for Health Research and Policy, University of Illinois at Chicago, Chicago, IL;* [2]*Department of Psychology, University of Illinois at Chicago, Chicago, IL;* [3]*Division of Epidemiology and Biostatistics, School of Public Health, University of Illinois at Chicago, Chicago, IL*

Corresponding Author: Alexander W. Sokolovsky, MA, Department of Psychology, University of Illinois at Chicago, 1007 W. Harrison Street, Chicago, IL 60607, USA. Telephone: 913-653-2741; Fax: 312-996-2703; E-mail: alexsoko@uic.edu

Received January 30, 2013; accepted August 26, 2013

ABSTRACT

Introduction: Ecological momentary assessments (EMAs) are increasingly used in smoking research to understand contextual and individual differences related to smoking and changes in smoking. To date, there has been little detailed research into the predictors of EMA compliance. However, patterns or predictors of compliance may affect key relationships under investigation and introduce sources of bias in results. The purpose of this study was to investigate predictors of compliance to random prompts among a sample of adolescents who had ever smoked.

Methods: Data for this study were drawn from a sample of 461 adolescents (9th and 10th graders at baseline) participating in a longitudinal study of smoking escalation. We examined 2 outcomes: subject-level EMA compliance (overall rate of compliance over a week-long EMA wave), and in-the-moment prompt-level compliance to the most proximal random prompt. We investigated several covariates including gender, race, smoking rate, alcohol use, psychological symptomatology, home composition, mood, social context, time in study, inter-prompt interval, and location.

Results: At the overall subject level, higher mean negative affect, smoking rate, alcohol use, and male gender predicted lower compliance with random EMA prompts. At the prompt level, after controlling for significant subject-level predictors of compliance, increased positive affect, being outside of the home, and longer inter-prompt interval predicted lower momentary compliance.

Conclusions: This study identifies several factors associated with overall and momentary EMA compliance among a sample of adolescents participating in a longitudinal study of smoking. We also propose a conceptual framework for investigating the contextual and momentary predictors of compliance within EMA studies.

INTRODUCTION

Ecological momentary assessment (EMA) has become an increasingly favored methodology over traditional "paper and pencil" diary methods. EMA maximizes ecological validity, minimizes recall bias, and allows for the examination of micro-contexts that influence behavior (Shiffman, Stone, & Hufford, 2008). However, EMA compliance remains an important methodological concern. For instance, individuals may not respond to all prompts or cues to report experiences, or may otherwise systematically avoid reporting; both instances may introduce important biases into data collection. This study sought to examine factors related to EMA compliance among adolescents from a longitudinal study of smoking patterns.

Compliance issues with traditional paper self-report methods have often centered around participants' failure to report behavior and surrounding events at the time of their occurrence, and then retroactively reconstructing these events at the time of data collection (Stone & Shiffman, 2002). Stone, Shiffman, Schwartz, Broderick, and Hufford (2002) have reported that up to 90% of events may be retrospectively backfilled by participants to give the appearance of good compliance. Concern about the falsification of paper diary entries has prompted the development of better diary tools, specifically the development of electronic diaries for EMA. Although implementation of EMA methods vary, they all involve the repeated measurement of a subject's behavior and experiences in real time—often using a programmed, time-sensitive signaling device, and an associated data collection modality such as interactive voice response (IVR), paper diary, or electronic diary.

Unlike paper diaries, electronic diaries and IVR enable the use of timestamping, a key feature in improving and assessing compliance. Making participants aware that their electronic diary assessments are electronically timestamped and, by extension, resistant to back-filling should improve momentary compliance with EMA protocols (Hufford & Shiffman,

doi:10.1093/ntr/ntt154
Advance Access publication October 4, 2013

2003). It is reassuring then that many studies employing electronic diaries for EMA report compliance rates above 90% (Cain, Depp, & Jeste, 2009; Hufford & Shiffman, 2003; Stone et al., 2003). Such compliance rates, however, are not universal with some electronic diary studies showing compliance rates below 75% (Jamison et al., 2001). To determine whether such data missingness may bias estimates of aggregated scores, it is important to establish first whether specific characteristics of measurement occasions are systematically associated with EMA noncompliance.

Data missingness is inherent to EMA protocols, and methods for addressing such concerns necessitate an expanded awareness of potential predictors of missingness. Current approaches to managing missing data, such as multiple imputation (MI), typically assume that the data are either missing at random (MAR) or missing completely at random (MCAR; Schafer & Graham, 2002). That is to say, missing data points are either a random sample from the full data set (MCAR) or a nonrandom sample of data where the predictors of missingness are known and included in the model (MAR). When data missingness cannot be predicted by other variables in the data set or are only predicted by the missing variable of interest, it is considered not missing at random (NMAR). Hedeker, Mermelstein, and Demirtas (2007) demonstrated a set of methods for imputing under NMAR by allowing for a relaxed relationship between missingness and binary outcomes. However, although MI under NMAR is possible, most currently established MI methods and software do assume MAR. As the assumption of MCAR is highly restrictive, it is important for EMA researchers to evaluate potential predictors of missingness to better maintain the MAR assumption. Knowledge of these predictors is crucial for MI under the MAR assumption.

This study addresses the question of which factors may be associated with compliance to random prompts in an EMA protocol. Although extant literature on predictors of compliance with EMA remains sparse, several studies have examined potentially important empirical predictors of compliance. These studies have focused primarily on background, subject-level factors that might be associated with poor response rates. Adolescents with learning difficulties in school were found to be less compliant with EMA than their peers (Salamon, Johnson, Grondin, & Swendsen, 2009). Although Palermo, Valenzuela, and Stork, 2004 found significant gender differences in compliance, others have not identified such an effect (Hacker & Ferrans, 2007). Additionally, there is some evidence that substance use, particularly polysubstance use predicts EMA noncompliance (Serre et al., 2012). Although these studies provide some general insights about compliance overall in particular populations, they do not provide much guidance within a given population of interest about what might predict compliance.

In addition to subject-level characteristics that may be associated with overall compliance, factors predicting compliance to random prompts at the prompt level may be of equal if not greater empirical interest. EMA studies often examine the relationships between momentary contextual variables and a behavior of interest. To the extent that the contextual variables influence an individual's probability of responding to EMA signaled prompts, conclusions about the relationships between these variables may be biased. Courvoisier, Eid, and Lischetzke (2012) found that although compliance varied throughout a day and across a study week, individual personality characteristics were not predictive of prompt-level compliance. Additionally, although they did not examine prompt-level compliance directly, Stone et al. (2003) found that the number of daily prompted diary entries correlated significantly with increased perceived burdensomeness with electronic diaries (Stone et al., 2003).

Factors Affecting Compliance: A Conceptual Model

We were guided by a multilevel ecological and social-cognitive framework to view compliance as a function of (a) background characteristics of an individual, (b) characteristics of the behavior itself, (c) demands of monitoring protocols, and (d) in-the-moment contextual influences. As evidenced previously, most investigations of EMA compliance to date have focused on evaluating the effects that background individual differences may have in predicting compliance. These investigations primarily aimed to establish the representativeness of their data between putative sample subgroups. However, incorporating additional information from the behavioral and contextual level may be useful in identifying systematic biases in compliance.

Problems with EMA compliance at the behavior level may be conceptualized as consequences of the specific research question being addressed and the population being targeted. In their study of treated alcoholics, Litt, Cooney, and Morse (1998) found that drinking alcohol was associated with a failure to respond to subsequent EMA signaling. Additionally, multiple discrepancies were discovered between EMA records and detailed timeline followback in nearly half of study participants. These results showed a remarkably clear deficit in EMA compliance associated with the very behavior being assessed.

Contextual problems with EMA compliance can result from collecting data in real time in naturalistic contexts. Such problems are specific to the different situations in which the EMA is used, independent of the research question. For example, compliance may decrease during specific times of the day, during working hours, or during social situations when recording might be socially inappropriate or embarrassing to participants. Unfortunately, such problems may limit the generalizability of study findings.

Problems with EMA compliance at the contextual level can lead to erroneous and potentially spurious research findings. For example, Kudielka, Broderick, and Kirschbaum (2003) tested how accurately subjects would comply with instructions to collect saliva samples at six specific times during one day. Sample collection times were recorded by participants in a timetable, but were also collected covertly and objectively through electronic timestamping in the sample vials. Subjects were divided into two groups: one group informed about the covert collection of data and other group uninformed about the covert collection of data. As expected, uninformed subjects significantly overreported their subjective compliance compared to informed subjects. Additionally, uninformed subjects significantly overreported their compliance compared to their own covertly recorded compliance levels. Noncompliance among uninformed subjects was limited to assessments early in the morning after waking—which is an empirically and diagnostically interesting period during which cortisol levels spike. Uninformed subjects consistently provided samples after the required timeframe, when cortisol levels had dropped. The resulting profiles of salivary cortisol in uninformed subjects thus revealed a significantly muted morning cortisol response, a misleading finding when compared to the profiles of informed subjects. The systematic noncompliance to the data collection procedures threatens the validity of the inferences drawn from the unrepresentative cortisol profiles and highlights the dangers of interpreting data derived from studies where potential context-level problems with compliance are not fully explored.

Any of a variety of exogenous factors (i.e., social desirability, task burdensomeness, response scheduling, social context, and so forth) may cause context-specific difficulties for compliance. Although the use of electronic diaries in EMA research has drastically improved compliance, the presence of systematic predictors of noncompliance to electronic diaries may still bias study results.

AIMS

This study aimed to identify predictors associated with EMA compliance among adolescents in a natural history study of smoking. We examined (a) the effect of individual difference and behavioral factors in predicting overall level of responding to random prompts over a seven-day monitoring period and (b) the effect of contextual factors in predicting compliance at the subsequent random prompt assessment point. We selected potential predictors based on a broad consideration of both background individual differences and momentary contextual factors that might affect responding. At the overall subject level of compliance, we examined the role that gender, race, age, academic performance (grade point average [GPA]), psychological symptoms (depression, anxiety, and antisocial behavior), home composition (number of parents in household and number of siblings), overall daily mood, smoking rate, and alcohol use may play in predicting compliance. Likewise, at the prompt level, we examined the effects of the most proximal random record of mood (positive and negative affect), hunger, practice/experience (study day), day of week, inter-prompt interval, social context (being with friends), and location. In general, we expected that the following factors would be associated with lower levels of overall compliance: male gender, lower GPA, higher levels of psychological symptoms, more siblings in the household (leading to more distractions), lower overall mood, higher levels of smoking, and greater alcohol use. We did not have specific directional hypotheses for race, age, or number of parents in the household. At the prompt level, we expected that being with friends, being outside home locations, experiencing increased negative affect, and shorter inter-prompt interval (increased participant burden) would be associated with missing the next prompt. We did not have specific directional hypotheses for positive affect or hunger states.

METHODS

Participants

Data for this study come from a longitudinal study examining the social and emotional contexts of adolescent smoking behavior. Participants were recruited from 16 schools in the Chicago metropolitan area. All 9th and 10th grade students in these schools (N = 12,970) completed a screening survey about smoking behaviors, and invitations to participate were mailed to eligible students and their parents. Of the 3,654 recruited to participate, 1,344 (36.8%) initially agreed and 1,263 (94.0%) completed baseline data collection. All students agreed to participate in several aspects of the longitudinal program project that included paper and pencil questionnaires, in-person interviews, and for a subset of participants, family interviews, psychophysiological assessment, and a week-long EMA event sampling via handheld computers (electronic diaries). We report here on data from the baseline data collection period.

Of those 1,263 students who completed the baseline assessment, a subsample composed only of adolescents who reported smoking at least once during the past 12 months were selected to participate in the electronic diary study; 461 adolescents (55.1% female; 53.2% in 10th grade; 56.8% White, 15.8% Black, 20.0% Hispanic, 2.8% Asian/Pacific Islander, and 4.6% were Other) completed the baseline electronic diary portion of the study. A majority (57.6%) had smoked at least one cigarette in the past month at baseline. Parental consent and student assent were obtained prior to enrollment in the study.

All participants in the electronic diary portion of the study received training on how to use the electronic diaries. The electronic diary was programmed with random prompt interviews and smoking-related interviews (event-recorded; not prompted). Data from this study come from the random prompt interviews. Following training, participants carried the electronic diary for 7 days, with the devices randomly prompting them on average 5–7 times per day. Each random prompt was date- and time-stamped and recorded whether the prompt was completed, missed, delayed, or abandoned. Devices included both suspend and prompt delay features to facilitate compliance. Participants completed a total of 14,105 random prompts (mean 30; range 7–71).

Non-EMA Self-Report Measures

Background Variables

Demographics included self-reported gender, age, and race/ethnicity.

Academic achievement was based on participants' self-reports of their average grade in the current academic year (GPA).

Home composition included the number of siblings and biologic and stepparents living in their household.

Psychological symptomatology included anxiety, depression, and antisocial behavior symptoms.

Adolescent *anxiety symptoms* were assessed with 12 items from the Mood and Anxiety Symptom Questionnaire (MASQ; Watson & Clark, 1991; Watson et al., 1995). The abbreviated version used in this study contains items from the anxious arousal subscale. Adolescents rated the extent to which they felt specific symptoms in the prior week on a 5-point Likert scale, ranging from 1 (*not at all*) to 5 (*extremely*). Individual item scores were summed to create a full scale score that was highly internally reliable (coefficient α = .81 in the current sample). Prior research supports the validity of the MASQ in both adult and adolescent samples (Reidy & Keogh, 1997; Richey, Lonigan, & Phillips, 2002).

Adolescent *depressive symptomatology* was assessed using the Center for Epidemiological Studies Depression Scale (CES-D; Radloff, 1977). The CES-D is a 20-item measure assessing the weekly frequency of depressive symptoms, ranging from 0 (*rarely*) to 3 (*most or all of the time*). Item responses were summed to create a scale score with high internal reliability (coefficient α = .89 in the current sample). Prior research supports the validity of the CES-D in an adolescent sample (Radloff, 1991).

Adolescent *aggressive and antisocial behavior* was assessed using a 22-item survey derived from a reduced set of the 46-item Antisocial Behavior Checklist (Noll, Zucker, Fitzgerald, & Curtis, 1992) and several items from a longitudinal study of adolescent problem behavior (Windle, 1992). The 22 current items tested core aspects of *DSM-IV* Conduct Disorder aligned across six domains of behavior: aggression, deceit, police contact, rule violation, theft, and vandalism. Participants responded to all items with the following choices: *never*, *rarely* (once or twice), *sometimes* (3–9 times), and *often* (more than 10 times). Sum scores were calculated for the full scale (coefficient α = .88) and all subscales.

Behavioral Variables

Smoking Rate: We created a measure of participant monthly smoking rate by multiplying participant self-reported number of days smoked in prior 30 days by the average number of cigarettes smoked on those days. Smoking rate at the baseline assessment wave ranged between 0 and 450 (representing 15 cigarettes/day on all 30 days). This measure was natural log-transformed to lessen the potential influence of large values on this variable in estimating of its effects.

Alcohol Use: Adolescents' alcohol use was assessed using a 5-item scale asking about alcohol use recency, quantity, lifetime maximum consumption, frequency, and problems. Participants' responses to each item, ranging from 1 to 8, were scaled based on the values of the respective responses. Item scores were averaged to form a full alcohol use scale score that showed high internal consistency (coefficient α = .86).

EMA Collected Measures

Affect: We created measures of participants' positive and negative affect means and variability by aggregating EMA event-level ratings of affect to the subject level. Participants rated their mood "Before the signal" by evaluating how strongly they felt about a set of 10 adjectives on a 10-point scale using a visual ladder (i.e., "Before the signal I felt … happy"). Factor analysis identified two distinct factors for positive affect (happy, relaxed, cheerful, confident, and accepted) and negative affect (angry, frustrated, irritable, sad, and stressed). Event-level ratings were used to predict compliance in-the-moment, while aggregated scores were used to predict compliance at the overall subject level.

Social Context: Participants reported whether they were or were not currently "with friends." All other social contexts, including being alone or being with nonfriend others, were coded as *not* being with friends.

Location: Participants reported their location. We created a categorical variable with levels corresponding to home, school, and an aggregate of all other locations.

Hunger: Participants reported their level of hunger on a 10-point scale using a visual ladder.

Automatically Collected Variables: Day of the week and number of days with the EMA computer (ranging from 1 to 8) were automatically collected by the electronic diaries. Day of the week was recoded into a categorical variable with levels corresponding to weekdays and weekend days. Inter-prompt interval was calculated from timing data collected by the EMA device.

Outcome Measures

Overall Compliance: Overall compliance was measured as the proportion of participants' completed random EMA prompts to their own total number of prompted EMA events. Delayed prompts were not counted against the participant's compliance rate and were only counted as one prompt.

Prompt-Level Compliance: In addition to predicting the overall compliance with random prompts, we were interested in examining whether the most proximal prior contextual variables predicted whether the next random prompt (within 150 min) was completed or missed. We identified pairs of random prompts that occurred within 150 min, with the criteria that the first prompt of the pair was completed. The outcome was whether the subsequent random prompt was *completed* or *missed*. We identified 7,745 total random prompt pairs (456 participants; M = 17.0 prompts per participant, SD = 5.64, range: 1–48) that met these criteria. Of these pairs, 6,332 were ones in which the second prompt was competed (complete–complete; 456 participants; M = 13.9, SD = 5.7; range: 1–43), and 1,413 were ones in which the second prompt was not completed (complete–incomplete; 427 participants; M = 3.1, SD = 1.9, range: 0–10). Twenty-nine participants had zero complete–incomplete pairs. Two participants were identified as potential outliers; excluding these participants did not significantly affect study findings, and they are included in the final results.

Analysis Plan

We analyzed our data in two separate steps: (a) a backwards selection of background and behavioral factors predicting overall compliance (level of completed prompts over the week), and (b) a logistic model of contextual, and significant background and behavioral factors predicting prompt-level compliance (whether the second part of two linked pairs of prompts was missed). First, a general linear model using the backwards method was estimated using the GLMSELECT procedure in SAS 9.2. The GLMSELECT procedure extends the forward, backward, and stepwise methods available in traditional regression analyses to general linear models (Cohen, 2006).

Second, factors predicting compliance in-the-moment were entered together into a logistic mixed model using the GLIMMIX procedure in SAS 9.2 (Schabenberger, 2005). Those individual differences that were found to predict compliance at the overall level were also included at the prompt level. Adolescents' responses/scores on these factors were included at each assessment, using the same score across all EMA events. This strategy both controls for the effects of subject-level predictors of compliance and determines whether these factors affect compliance at the prompt level. We included both the subject-level measures of mean negative and positive affect, as well as the prompt-level deviation of a subject's affect from their own mean in this model. This approach enabled us to estimate both the between-subject (BS; mean) and within-subject (WS; deviation) effects of mood on compliance. A random subject effect was included in the model to account for the clustering of the prompt-level observations WS (Hedeker & Gibbons, 2006). The model was used to estimate those factors that predict noncompliance with random prompts. Population-averaged odds ratios (*OR*s) for factors predicting noncompliance were calculated from the estimated fixed effects.

RESULTS

Predicting Overall Level of Compliance

Overall mean compliance rates were 68.1% (*SD* = 16.9%). The mean time between prompt pairs was 81.6 min (*SD* = 34.7) for complete–complete pairs and 85.0 min (*SD* = 34.4) for complete–incomplete pairs. Descriptive statistics for demographic, behavioral, mood and psychological symptomatology variables are shown in Table 1. Participants reported smoking 21.1 cigarettes/month on average and reported moderate levels ($M = 4.05$) of alcohol use (1–8 scale; higher scores represent more problems). On average, participants completed 30.6 random prompts ($SD = 8.2$) and missed 15.6 ($SD = 12.6$). Bivariate correlations revealed that lower levels of smoking ($r = -.19$, $p < .001$), alcohol use ($r = -.1$, $p < .001$), and antisocial behavior ($r = -.15$, $p < .001$) were associated with higher compliance. Additionally, being male ($r = -.15$, $p < .001$) was associated with lower compliance.

To examine overall level of compliance, all background and behavioral variables were entered simultaneously into a general linear model using the backwards selection method. Significance level to stay in the model was defined at $p < .10$. Parameter estimates for the final selected model are displayed in Table 2. Variables excluded from the final model include race, age, mean positive affect, positive affect variability, negative affect variability, number of siblings in household, number of parents in household, GPA, depression, anxiety, and antisocial behavior. Those effects selected to remain in the model suggest that being female and having lower mean negative affect, less smoking, and less alcohol use predicted higher subject-level compliance with EMA.

Predicting Compliance at the Prompt Level

The stability of prompt-level predictors was examined by calculating bivariate correlations for predictors between consecutive completed prompts. Predictors were found to be moderately stable across random prompts, ranging from $r = .50$, $p < .001$ for location to $r = .65$, $p < .001$ for positive affect, with the exception of hunger ($r = .38$, $p < .001$). Results from the logistic mixed model examining whether the most proximate prompt within a 150-min windows predicted next-prompt compliance are presented in Table 3. The fixed effect parameter estimates in the table predict prompt-level noncompliance. Those factors found to predict compliance at the overall level of analysis (gender, smoking rate, and alcohol use) were included as covariates in this model. Additionally, both BS (mean negative and positive affect) and WS (subject's in-the-moment deviation from their own mean) measures of affect were included in the model. These results show that several factors (positive affect WS, location, inter-prompt interval, and gender) are significant determinants of prompt-level noncompliance—that is, failure to respond to the next prompt that occurs within the 150-min window. One of the four covariates included from the subject-level analysis, gender, significantly predicted prompt-level noncompliance. Although alcohol use and smoking rate predicted lower compliance at the overall, subject level, they were not predictive of prompt-level noncompliance. Population-averaged *OR*s and 95% confidence intervals for fixed effects are presented in Table 3.

DISCUSSION

This study identified several factors associated with overall and prompt-level EMA compliance among a sample of adolescents participating in a longitudinal study of smoking. Although EMA methods offer many advantages over traditional retrospective diary or assessment methods, compliance with EMA methods may still present biases in data collection. Electronic data collection methods, however, permit tracking of missing data occurrences in real time, providing potential opportunities to examine momentary predictors of noncompliance.

We considered compliance within a multilevel ecological framework, viewing an adolescent's responsiveness to random prompts as a function of momentary, proximal factors, such as the social context, mood states, and location, as well

Table 1. Descriptive Statistics for Demographic, Behavioral, Mood, and Psychological Symptomatology Variables at Baseline

Measures	*n*	Range	*M*	*SD*
Age	461	13.9–17.3	15.7	0.613
Grade point average	460	2.00–5.00	3.39	1.81
30-day smoking rate	460	0–450	21.2	57.0
Alcohol use	461	1.00–7.10	4.05	1.59
Negative mean affect	456	1.02–8.26	3.44	1.45
Negative affect variability	456	−5.39 to 1.62	−0.11	1.11
Positive mean affect	456	2.90–10.0	6.82	1.20
Positive affect variability	456	−3.80 to 1.57	−0.09	0.73
Anxiety	461	12–55	28.9	8.16
Depression	460	0–52	17.4	10.2
Antisocial behavior	461	23–74	36.3	8.49
Siblings in household	461	0–8	1.66	1.25

Notes. Depression was measured with the Center for Epidemiological Studies Depression Scale on a 0–60 scale; anxiety was measured with the Mood and Anxiety Symptom Questionnaire on a 12–60 scale; antisocial behavior was measured with the Antisocial Behavior Checklist on a 22–80 scale (higher scores on all scales indicate more disordered symptomatology). Negative and positive moods were measured by aggregating prompt-level ecological momentary assessment (EMA) measures. Affect variabilities represent intraindividual standard deviations measured through EMA. Thirty-day smoking rate represents total number of cigarettes smoked in prior 30 days.

Table 2. Parameter Estimates for Final Selected Model Predicting Subject-Level Compliance

Parameter	*df*	Estimate	*SE*	*t* value	*p* value
Intercept	1	0.804	0.0281	28.61	<.001
Gender	1	−0.0503	0.0160	−3.15	.0018
Negative mean affect	1	−0.0113	0.00551	−2.05	.0409
Log 30-day smoking rate	1	−0.0124	0.00493	−2.51	.0124
Alcohol use	1	−0.0110	0.00535	−2.05	.0410

Note. Model estimated using GLMSELECT procedure in SAS 9.2. Gender variable coding: 0 = female; 1 = male.

Table 3. Parameter Estimates for Fixed Effects in Logistic Mixed Model Predicting Prompt-Level Noncompliance

							95% CI	
Parameter	*df*	Estimate	*SE*	*t* value	*p* value	*OR*	Low	High
Intercept	450	−2.82	0.392	−7.20	<.001*	–	–	–
Study day	7280	0.0217	0.0149	1.46	0.144	1.018	0.999	1.037
Weekday	7280	0.0677	0.0736	0.92	0.358	1.067	0.929	1.227
With friends	7280	0.0281	0.0752	0.37	0.709	1.031	0.895	1.189
Positive affect (WS)	7280	0.0625	0.0257	2.44	0.0149*	1.062	1.012	1.115
Negative affect (WS)	7280	−0.0239	0.0226	−1.06	0.289	0.977	0.937	1.020
Positive affect (BS)	7280	0.0189	0.0413	0.46	0.647	1.018	0.942	1.101
Negative affect (BS)	7280	0.0615	0.0351	1.75	0.0793	1.061	0.993	1.134
Inter-prompt interval	7280	0.00255	0.00090	2.84	0.0045*	1.0025	1.0008	1.0042
Hunger	7280	0.0185	0.00997	1.85	0.0642	1.021	0.993	1.050
Gender	7280	0.206	0.0822	2.50	0.0123*	1.219	1.044	1.424
Log 30-day smoking rate	7280	0.0270	0.0255	1.06	0.291	1.026	0.978	1.077
Alcohol use	7280	0.0436	0.0276	1.58	0.115	1.043	0.990	1.099
Location								
School (vs. home)	7280	0.419	0.0905	4.63	<.001*	1.498	1.262	1.777
Other (vs. home)	7280	0.351	0.0837	4.19	<.001*	1.402	1.197	1.642

Notes. *OR* = odds ratio; CI = confidence interval. Significant positive estimates indicate greater noncompliance. Weekday coding: 0 = weekday; 1 = weekend. With friends coding: 0 = alone; 1 = with friends. Model estimated using GLIMMIX procedure in SAS 9.2. *ORs* of continuous variables (positive affect, negative affect, hunger, inter-prompt interval, log smoking rate, and alcohol use) are associated with one scale unit offsets from the mean. **p* < .05. WS = within subject (subject's momentary deviation in affect from their own mean). BS = between subject (subject's mean affect).

as background contextual influences, including family context, overall psychological and academic functioning, and demographics. Unique to this study was our consideration of these multiple levels of influence on compliance within one model. Until now, researchers have focused primarily on evaluating only those individual differences that distinguish EMA compliers from noncompliers. Although this approach has been useful in identifying potential bias and subgroup differences in compliance, it does not facilitate addressing factors that may contribute to noncompliance in-the-moment. Our data also provide the first substantive evaluation of random prompt compliance, using the adolescent's own momentary self-reported states to prospectively predict subsequent prompt compliance.

Our results illustrate the importance that participants' location and mood each have in determining random prompt compliance. Examination of mood data shows that both increased positive affect at the prompt level and higher overall negative affect had deleterious effects on compliance. We believe that adolescents who have higher overall negative affect may be less motivated to respond to EMA prompts across measurement occasions. Additionally, those adolescents who find themselves in relatively more positive emotional states (compared to their average mood) may be emotionally too stimulated and not have the cognitive resources available to fully attend to the EMA protocol. Participants completing random prompts outside the home had significantly greater odds of noncompliance. Many theoretically important events and/or emotional states that may trigger adolescent smoking events occur outside the home and thus may not be fully reported with this protocol. Counter to our expectations, longer inter-prompt intervals predicted greater noncompliance with random prompts.

Although EMA researchers are frequently concerned with the effects of participant burden on compliance, this finding reflects on the balance that must be drawn between over- and under-utilization of the EMA device. Although having too-frequent prompts may be perceived as burdensome in an adolescent population, too-few prompts may lead to participants disengaging from the device and forgetting to attend to the EMA protocol. Finally, it is important to mention that a subject-level factor predicting overall compliance (gender) was also found to predict compliance at the prompt level. These results suggest that additional training or prompt signaling may be indicated for those individuals identified to be at risk for noncompliance.

The results of this study have meaningful implications for dealing with missing data when employing EMA methodology. Identifying significant predictors of data missingness at the overall and prompt level helps to maintain the MAR assumption. The possibility that missing data in EMA was only predicted by the actual missing value and thus NMAR is impossible to assess with the observed data. Thus, it is crucial that investigators routinely do analyses to find variables related to missingness. Variables that predict data missingness can then be included as covariates in EMA analyses to better satisfy the MAR assumption. Additionally, researchers may consider including integrated measures predictive of compliance (e.g., global positioning system for device location during missing prompts). Further investigation may identify other predictors of data missingness that may facilitate variable selection for EMA protocols.

This study's findings are limited in part by examining compliance at only the baseline assessment wave. Because adolescents may develop improved executive functioning skills and coping strategies with age, those effects found to predict compliance in 9th and 10th graders may not be as predictive of compliance in older samples. Indeed, we have continued to follow this sample of adolescents over multiple waves and have found that compliance increases over time, perhaps both as a function of experience with the protocol and increasing maturity. However, compliance in older samples should not be taken for granted. An adequate investigation of factors predicting noncompliance in those samples is needed to identify bias and reveal potential targets for improving responding. We are currently applying the modeling approach used in this study to later data collection waves of this project.

A notable limitation of our findings about prompt-level compliance is that we assessed momentary factors not immediately in the missed prompt situation (which would not be possible by definition), but in close temporal proximity to the missed event (within prior 150 min, 85-min mean interval). Much can change within that timeframe—clearly the context may well change. However, we demonstrated an approach to modeling that may help inform investigations of associations between potential momentary factors and outcomes. In addition, although we included a relatively large number of predictor variables, other key factors associated with noncompliance were not systematically assessed. As this investigation focused exclusively on random prompts, we did not assess whether smoking behavior reported at the prompt level was associated with reduced compliance. Consistent with the deleterious effects of alcohol on compliance, such influences remain a possibility and should be explored with further studies. Additionally, at the end of each data collection wave, we debriefed the participants about their missed prompts. By far, the most common explanation for the missed prompts was failure to hear the device's signal. Indeed, when participants were in noisy environments or when the device was deep in a participant's filled backpack, it was challenging to hear the device. Thus, nonparticipant factors also contributed to level of compliance. We also found that compliance can be improved. At subsequent data collection waves, we offered monetary incentives for achieving benchmark levels of compliance with random prompts, and compliance levels increased significantly.

In sum, this study highlights an approach to investigating the contextual and momentary predictors of compliance within EMA studies and also emphasizes the importance of considering potential biases that exist with missing data. More attention to missing data considerations with EMA data is needed to help further the field.

FUNDING

This work was supported by a grant from the National Cancer Institute (P01 CA098262).

DECLARATION OF INTERESTS

None declared.

ACKNOWLEDGMENTS

The authors thank Siu Chi Wong for assisting with data analysis.

REFERENCES

Cain, A. E., Depp, C. A., & Jeste, D. V. (2009). Ecological momentary assessment in aging research: A critical review. *Journal of Psychiatric Research*, *43*, 987–996. Retrieved from http://dx.doi.org/10.1016/j.jpsychires.2009.01.014

Cohen, R. A. (2006). Introducing the GLMSELECT procedure for model selection. Paper presented at the Proceedings of the Thirty-first Annual SAS Users Group International Conference, Cary, NC: SAS Institute Inc.

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. *Psychological Assessment*, *24*, 713–720. Retrieved from http://dx.doi.org/10.1037/a0026733

Hacker, E. D., & Ferrans, C. E. (2007). Ecological momentary assessment of fatigue in patients receiving intensive cancer therapy. *Journal of Pain and Symptom Management*, *33*, 267–275. Retrieved from http://dx.doi.org/10.1016/j.jpainsymman.2006.08.007

Hedeker, D., & Gibbons, R. D. (2006). *Longitudinal data analysis* (Vol. *451*). Hoboken, NJ: Wiley-Interscience. doi:10.1002/0470036486

Hedeker, D., Mermelstein, R. J., & Demirtas, H. (2007). Analysis of binary outcomes with missing data: Missing = smoking, last observation carried forward, and a little multiple imputation. *Addiction*, *102*, 1564–1573. doi:10.1111/j.1360-0443.2007.01946.x

Hufford, M., & Shiffman, S. (2003). Assessment methods for patient-reported outcomes. *Disease Management & Health Outcomes*, *11*, 77–86. doi:http://dx.doi.org/10.2165%2F00115677-200311020-00002

Jamison, R. N., Raymond, S. A., Levine, J. G., Slawsby, E. A., Nedeljkovic, S. S., & Katz, N. P. (2001). Electronic diaries for monitoring chronic pain: 1-year validation study. *Pain*, *91*, 277–285. doi:10.1016/s0304-3959(00)00450-4

Kudielka, B. M., Broderick, J. E., & Kirschbaum, C. (2003). Compliance with saliva sampling protocols: Electronic monitoring reveals invalid cortisol daytime profiles in non-compliant subjects. *Psychosomatic Medicine*, *65*, 313–319. doi:10.1097/01.psy.0000058374.50240.bf

Litt, M. D., Cooney, N. L., & Morse, P. (1998). Ecological momentary assessment (EMA) with treated alcoholics: Methodological problems and potential solutions. *Health Psychology*, *17*, 48–52. doi:10.1037/0278-6133.17.1.48

Noll, R. B., Zucker, R. A., Fitzgerald, H. E., & Curtis, W. J. (1992). Cognitive and motoric functioning of sons of alcoholic fathers and controls: The early childhood years. *Developmental Psychology*, *28*, 665. doi:http://dx.doi.org/10.1037%2F%2F0012-1649.28.4.665

Palermo, T. M., Valenzuela, D., & Stork, P. P. (2004). A randomized trial of electronic versus paper pain diaries in children: Impact on compliance, accuracy, and acceptability. *Pain*, *107*, 213–219. doi:10.1016/j.pain.2003.10.005

Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, *1*, 385–401. doi:http://dx.doi.org/10.1177%2F014662167700100306

Radloff, L. S. (1991). The use of the Center for Epidemiologic Studies Depression Scale in adolescents and young adults. *Journal of Youth and Adolescence*, *20*, 149–166. doi:10.1007/BF01537606

Reidy, J., & Keogh, E. (1997). Testing the discriminant and convergent validity of the Mood and Anxiety Symptoms Questionnaire using a British sample. *Personality and Individual Differences*, *23*, 337–344. doi:http://dx.doi.org/10.1016%2FS0191-8869%2897%2900048-2

Richey, J. A., Lonigan, C. J., & Phillips, B. M. (2002). Exploratory factor analysis of the Mood and Anxiety Symptom Questionnaire (MASQ) in a youth sample. Poster Presentation at the Annual Convention of the Association for the Advancement of Behavior Therapy (AABT), Reno, NV.

Salamon, R., Johnson, E.-I, Grondin, O., & Swendsen, J. (2009). Étude des problèmes comportementaux et émotionnels chez l'adolescent: Faisabilité et validité de l'approche ESM. *Journal de Thérapie Comportementale et Cognitive*, *19*, 41–46. Retrieved from http://dx.doi.org/10.1016/j.jtcc.2009.04.004

Schabenberger, O. (2005). Introducing the GLIMMIX procedure for generalized linear mixed models. *SUGI 30 Proceedings*, 196–130.

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, *7*, 147–177. doi:http://dx.doi.org/10.1037%2F%2F1082-989X.7.2.147

Serre, F., Fatseas, M., Debrabant, R., Alexandre, J.-M., Auriacombe, M., & Swendsen, J. (2012). Ecological momentary assessment in alcohol, tobacco, cannabis and opiate dependence: A comparison of feasibility and validity. *Drug and Alcohol Dependence*, *126*, 118–123. Retrieved from http://dx.doi.org/10.1016/j.drugalcdep.2012.04.025

Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment. *Annual Review of Clinical Psychology*, *4*, 1–32. doi:http://dx.doi.org/10.1146%2Fannurev.clinpsy.3.022806.091415

Stone, A. A., Broderick, J. E., Schwartz, J. E., Shiffman, S., Litcher-Kelly, L., & Calvanese, P. (2003). Intensive momentary reporting of pain with an electronic diary: Reactivity, compliance, and patient satisfaction. *Pain*, *104*, 343–351. doi:http://dx.doi.org/10.1016%2FS0304-3959%2803%2900040-X

Stone, A. A., & Shiffman, S. (2002). Capturing momentary, self-report data: A proposal for reporting guidelines. *Annals of Behavioral Medicine*, *24*, 236–243. doi:10.1207/s15324796abm2403_09

Stone, A. A., Shiffman, S., Schwartz, J. E., Broderick, J. E., & Hufford, M. R. (2002). Patient non-compliance with paper diaries. *BMJ*, *324*, 1193–1194. doi:http://dx.doi.org/10.1136%2Fbmj.324.7347.1193

Watson, D., & Clark, L. (1991). The Mood and Anxiety Symptom Questionnaire. *Unpublished manuscript,* Department of Psychology, University of Iowa, Iowa City.

Watson, D., Weber, K., Assenheimer, J. S., Clark, L. A., Strauss, M. E., & McCormick, R. A. (1995). Testing a tripartite model: I. Evaluating the convergent and discriminant validity of anxiety and depression symptom scales. *Journal of Abnormal Psychology*, *104*, 3–14. doi:10.1037/0021-843X.104.1.3

Windle, M. (1992). A longitudinal study of stress buffering for adolescent problem behaviors. *Developmental Psychology*, *28*, 522. doi:http://dx.doi.org/10.1037%2F%2F0012-1649.28.3.522