Journal of Marketing Analytics
https://doi.org/10.1057/s41270-024-00347-6

ORIGINAL ARTICLE

On the use of self-reports in marketing research: insights about initial response biases from daily diary data

Shelley A. Blozis[1] · **Ricardo Villarreal**[1]

Revised: 3 June 2024 / Accepted: 20 August 2024

Abstract
Self-report data are regularly used in marketing research when consumer perceptions are central to understanding consumer responses to marketing efforts. Self-report data are convenient and cost-effective. A widely known response bias that is inherent to self-report data and illuminated by daily diary data is a tendency of the first report by study participants to be more extreme relative to responses at subsequent points in time but no such effects are expected. A critical issue is that statistical data summaries can be impacted and generate misleading conclusions about perceptions. This article demonstrates the impact of initial-report effects by analyzing self-report daily diary media use (TV) data from an observational study. Based on a large and representative sample of adults in the U.S., there was a greater tendency for respondents to report watching TV, and given that TV was watched, to report more time spent watching TV on the first interview day relative to subsequent days. Initial-report effects were also evident in tests of the effects of daily and daily averages of positive and negative affect on the likeliness to watch TV and reported time spent, further indicating the importance of accounting for first-report effects in studies of media use. The need to collect repeated measures of self-report data in consumer research is also highlighted by this evidence of these response patterns that would otherwise be undetectable.

Keywords TV · Consumer behavior · Media use · Mixed-effects models · Mixed-effects location scale models · Repeated measures · Longitudinal data · Multilevel regression

Introduction

Consumer behavior research is reliant on self-report data, especially for measuring psychological variables that are foundational to marketing and consumer behavior research (Dhiman and Kumar 2023; Foxall 2003) and for measuring relatively objective behaviors, including brand and media use (Eastman et al. 2021; Blozis et al. 2019; Rich et al. 2015). Self-report data are generally inexpensive and easy to obtain through surveys, and that helps to increase the feasibility of obtaining large and representative samples. Self-report data are known, however, to exhibit a response pattern in which the first assessment tends to be more extreme than subsequent assessments (Knowles et al. 1996; Robins 1985; Sharpe and Gilbert 1998; Windle 1954). This response pattern is independent of external influences, such as an experimental manipulation. Historically, this pattern was thought to reflect an attenuation in responses following the first assessment. Recent experimental research in social psychology, however, provides evidence that strongly suggests this pattern likely reflects an initial elevation bias (Shrout et al. 2018). Termed the 'initial elevation' (IE) effect, this response pattern holds significant implications for repeated measures and longitudinal inquiries in consumer behavior studies, including experimental and observational studies. Importantly, it holds implications for studies that rely on a single assessment because its impact on such data is uncertain.

To gain a clearer understanding of the source of an elevation in the initial mean response in repeated measures obtained by self-report surveys, Shrout et al. (2018) carried out four separate experiments using intensive data collections (e.g. daily survey) in which participants were randomly assigned to different start dates of a self-report survey that was repeated daily or bimonthly. In one study, for instance, all participants had plans to take the same stressful exam, but their entry into the survey study occurred at different

Shelley A. Blozis
sablozis@ucdavis.edu

[1] Psychology Department, University of California, One Shields Ave, Davis, CA 95616, USA

Published online: 19 September 2024

points in time relative to the date of the common exam. All four experiments surveyed participants about psychological state variables (e.g. daily affect, daily anxiety) and some included measures of more objective behaviors (e.g. daily reported time spent studying). Mean responses were consistently elevated on the first survey day no matter when the survey began for a given group. In one study, for instance, an elevation in the mean response was evident for participants who started the survey far in advance of the exam, close to the time of the exam, as well as after the exam had been administered. In other words, it did not matter when a group started the survey study about daily affect and behavior measures, the IE response pattern was evident. From these results, the evidence strongly suggests a response bias due to the first exposure to a survey. These studies also showed that relatively subjective psychological variables (e.g. affect), as opposed to the more objective variables (e.g. reported time spent studying), showed stronger IE effect sizes.

These findings have implications for any research study that uses self-report data, and marketing research is one such domain. Indeed, marketing and consumer behavior research rely on self-reports to understand consumer experiences (Norton et al. 2015). Measuring the magnitude of an IE effect, however, requires repeated measures data so that the first assessments can be evaluated against those that follow. Current analytic trends in marketing research make it possible to test for IE effects in self-report data. That is, market research is increasingly stressing the importance of evaluating consumer data over time to better understand the individual-level consumer experience, in addition to differences between consumers (Beuk et al. 2014; Blozis et al. 2020, 2019; Chaney and Martin 2017; Katakam et al. 2021; Kumar et al. 2017; Payan et al. 2010). Thus, given the inherent bias of self-report data documented in psychology, consumer behavior research that relies on self-report data could benefit from a close look at possible IE effects.

The current study

This study uses publicly available daily diary data from a large, random sample of adults selected to be representative of the U.S. to test for and evaluate the magnitude of an IE response in self-reported media-use data. Unlike cross-sectional data for which responses are recorded at a single occasion, daily diary and similar intensive data collections allow for a comparison of responses at the first assessment to those that follow soon after. Intensive data collections also differ from longitudinal data that make assessments at times that are relatively far apart, making it difficult to assess the magnitude of an IE effect. As reported in Shrout et al. (2018), relatively intensive data collections yield stronger estimates of an IE effect.

The current study specifically examines IE effects in TV use reports, though the methods applied may be more generally considered to other types of media use reports. Self-reported TV use is also studied in relation to positive and negative affect, two psychological variables deemed important in marketing research (Thürridl et al. 2020). Multiple studies primarily in the field of psychology have documented an IE response pattern in affect reports, such that self-reported affect levels tend to be more extreme at the first assessment relative to subsequent assessments, and so that work is not duplicated here. Instead, measures of positive and negative affect are included as daily covariates of TV use. Importantly, we test for the moderating effects of IE on the relationships between affect and TV use. The remainder of this paper is organized as follows: Research on the relationships between affect and media use, including TV use, is reviewed. We then describe data from the daily diary study that are used to study IE effects. Statistical models selected for the analyses are then described and applied to the data. Findings and implications from the analyses follow.

Media use and affect

Holbrook and Hirschman (1982) were the first to stress the inherent importance of the experiential aspects of consumer behavior. Of the seven issues they addressed, one in particular alluded to an individual's affect as "feelings arising from consumption" (p. 139). Among the many reports building on this foundational work, Chen and Pham (2019) emphasized the necessity of understanding the link between affect regulation and consumption-based affect regulation. Indeed, media engagement has been widely studied in relation to an individual's mood (e.g. Bowman and Tamborini 2015; Cohen et al. 2008; Greenwood and Long 2009; Hoffner and Lee 2015; Schimmack and Crites 2005; Wolfers and Scheider 2021). An individual's affective state can be influenced by consumer behavior (Chen and Pham 2019), including shopping (Rick et al. 2014), eating (Gibson 2006), and media consumption (Nabi et al. 2017; Nabi and Krcmar 2004; Wolfers and Scheider 2021). Both positive and negative affect states have been associated with media use, and TV use in particular (Festinger and Katz 1953), where selective TV programming tends to align with an individual's psychological needs (Knobloch-Westerwick and Alter 2006). This selective exposure is thought to help individuals manage their emotions effectively (Nabi et al. 2017; Wolfers and Scheider 2021). Ultimately, consumers can either maintain or improve their positive emotional state through their consumption behaviors, with media use being one of the most common methods for doing so.

Example using daily diary data

Data for this study were sourced from the Daily Diary Project, a subproject of the Midlife in the United States (MIDUS) Refresher 1 (Ryff and Almeida, 2012–2014). From that sample, 782 adults were selected to participate in a telephone interview for 8 consecutive days. A select set of questions were asked repeatedly across the 8 days and asked with regard to the previous 24 h. The day of the week for the first interview varied between participants. About 30 people were surveyed in blocks across the study years. We analyze daily reports of time spent watching TV and measures reflective of positive and negative affect. The sample was 55.6% female. Females were 47.6 (*SD* 12.9, min = 25, max = 75) years old on average; males were 48.2 (*SD* 12.5, min = 25, max = 75) years on average. The data analyzed for this report are publicly available from the Inter-university Consortium for Political and Social Research (ICPSR).

Measures

We converted daily responses about the time spent watching TV in minutes to hours. Positive affect was measured by averaging responses to 13 questions that had a common stem: 'How much of the time today did you feel …' All items used a 5-point response scale (0: none of the time; 1: a little of the time; 2: some of the time; 3: most of the time; 4: all of the time). Examples of survey items are "like you belong?" and "close to others?" Higher scores reflect greater positive affect levels. For the sample, reliability measured by Cronbach's alpha was .95. Negative affect was measured by averaging responses to 14 questions that had the same stem and response scale used to measure positive affect. Examples of survey items are "hopeless?" and "worthless?" Higher scores reflect greater negative affect levels. For the sample, reliability measured by Cronbach's alpha was .87.

Covariates

Analyses accounted for the day of the week of the interview day, as well as the age and biological sex of participants. Biological sex was coded as $female = 1$ if female and $female = 0$ if male. Day of the week was represented by a set of indicator variables for each day, day_k for $k \in \{1,\ldots,6\}$, with Sunday serving as the reference day. For example, $day_1 = 1$ if the interview was conducted on a Monday and was equal to 0 otherwise.

Analytic strategy

Daily measures of reported time spent watching TV are semi-continuous, meaning that values can be discrete or continuous. For time spent watching TV, zero is a discrete value that indicates that the respondent did not watch TV; the remaining values represent the time spent when the respondent reported to have watched TV. To account for these two aspects of TV use, a two-part mixed-effects model is applied to daily measures of reported TV use (Blozis et al. 2019). This joint model simultaneously models whether or not an individual engaged in a measured behavior and the extent of that engagement on occasions when the individual was engaged. Technically, the model combines a logistic mixed-effects model for repeated measures of the binary indicator of whether or not an individual engaged in the behavior and a generalized linear mixed-effects model for repeated measures of the level of engagement conditional that there was a positive amount of the behavior. Given that time spent watching TV is known to vary from day to day for some individuals, a version of the model, known as a two-part mixed-effects location scale (MELS) model (Blozis et al. 2019, 2020) is used that includes a random scale effect for the continuous model part to permit individual differences in the day-to-day variability in time spent. In sum, a two-part MELS model is used to estimate the daily likeliness to watch TV, the daily mean time spent when TV was watched, and the day-to-day variability in time spent on days when TV was watched. Maximum likelihood estimation was carried out using SAS version 9.4 with PROC NLMIXED and Gaussian quadrature. We used parameter estimates from relatively less complex models as the starting values for more complex models.

The analysis of TV use reports proceeded by fitting a model that excluded covariates and assumed a lognormal distribution for the continuous model part to address the positive skew in values (comparisons of fit between models that made different assumptions about the response distribution of scores were carried out in Blozis et al. 2019). Next, an indicator of the first interview day entered the model to test for differences in the expected likeliness to watch TV and the expected time spent between the first interview versus subsequent daily interviews. This was done to test if TV use reports showed evidence of IE effects. Next, daily positive and negative affect, with the other covariates discussed previously, entered the model to evaluate the impact of an IE effect. In particular, we tested if the relationships between positive and negative affect on TV use measures differed between the first interview versus the interviews that followed. We end by fitting a model that excluded all IE effects to evaluate how ignoring IE effects could impact statistical inference of the associations between positive and negative affect and TV use measures. It is worth noting that from any model in which aspects of TV use are regressed on affect measures, daily measures of affect and TV use reflect reports for the same

24-h period, and so inferences about the effects of affect measures on TV use measures are correlational.

Missing data

Estimation of mixed-effects models does not require that each respondent has complete response data for all planned occasions. For individuals with incomplete data, statistical inference is considered valid if the data are missing at random, meaning that if an individual has missing data, the source of the missing data is independent of the missing values (Laird and Ware 1982). We assume data were missing at random.

Two-part MELS model for TV use reports

Let y_{ti} be the reported time spent watching TV on day t for individual i, where $i = 1, \ldots, 782$ and $t = 1, \ldots, \mathrm{n_i}$, with $\mathrm{n_i}$ being the number of daily responses for the individual. To fit a two-part model, two variables were created from the original response y_{ti}. The first, u_{ti}, was set equal to 1 if the individual reported watching TV (i.e., $y_{ti} > 0$) and was set equal to 0 if they reported not watching TV (i.e., $y_{ti} = 0$). The second variable, m_{ti}, was set equal to the reported time spent watching TV if any amount was reported (i.e., $m_{ti} = y_{ti}$ if $y_{ti} > 0$) and was coded as missing if the individual reported no time spent.

The binary indicator of whether or not TV was watched was modeled using a mixed-effects logistic regression model. Let η_{ti} denote the logit of the probability that individual i reported any amount of TV use on day t (i.e., $\mathrm{P}(u_{ti} = 1)$):

$$\eta_{ti} = \log\left[\mathrm{P}(u_{ti} = 1)/\left(1 - \mathrm{P}(u_{ti} = 1)\right)\right].$$

The logit η_{ti} was assumed to follow a two-level model:

$$\eta_{ti} = \alpha_0 + a_i, \tag{1}$$

where α_0 is the population log odds of watching TV. The coefficient a_i is a random subject effect that indicates how much an individual's log odds differs from the population log odds. The random effect a_i is assumed to be independently and identically distributed (i.i.d.) between subjects as normal with a mean of 0 and variance ϕ_{a}^2 that quantifies the degree to which individuals varied in their log odds of TV use.

A positive report of time spent watching TV was modeled by a generalized linear mixed-effects model with scores assumed to follow a lognormal distribution:

$$m_{ti} = \beta_0 + b_i + e_{ti}, \tag{2}$$

where β_0 is the mean log time spent across days and individuals. The coefficient b_i is a random subject effect that indicates how much an individual's mean log time spent differs from the population value. The random effect b_i is assumed to be i.i.d. between subjects as normal with a mean of 0 and variance ϕ_{b}^2 that quantifies individual differences in mean log time spent on days when TV was watched. The residual e_{ti} is assumed to be i.i.d. between subjects and days as lognormal with a mean of 0 and variance σ_{e}^2.

The within-subject residual variance of the continuous model characterizes the degree to which observed scores deviate about an individual's fitted mean response across days. In modeling this variance, a random scale effect was used to allow for between-subject heterogeneity of the within-subject variance. That is, the day-to-day variability in reported time spent watching TV on days when TV was watched could differ between respondents. The within-subject residual variance was expressed using an exponential function (cf: Hedeker et al. 2012):

$$\sigma_{\mathrm{e}}^2 = \exp\left(\tau_0 + c_i\right), \tag{3}$$

where the exponentiated value of τ_0 is the residual variance when the random scale effect c_i is equal to zero. The random scale c_i is assumed to be i.i.d. normal with a mean of 0 and variance ϕ_c^2 that quantifies the variance of the log normal perturbations of the within-subject residual variance. At the subject level, the models are joined by the covariances between the random intercepts of Eqs. (1 and 2) and the random scale of Eq. (3). The covariance matrix of the random intercepts and the random scale is given by $\mathbf{\Phi}$:

$$\mathbf{\Phi} = \begin{bmatrix} \phi_{\mathrm{a}}^2 & \phi_{\mathrm{ab}} & \phi_{\mathrm{ac}} \\ \phi_{\mathrm{ba}} & \phi_{\mathrm{b}}^2 & \phi_{\mathrm{bc}} \\ \phi_{\mathrm{ca}} & \phi_{\mathrm{cb}} & \phi_{\mathrm{c}}^2 \end{bmatrix},$$

where the variances and covariances are the diagonal and off-diagonal elements, respectively.

Testing for IE effects

To test for IE effects in reported TV use, an indicator of the first interview day was created: $\mathrm{IE_{ti}} = 1$ if the tth report by person i took place on their first interview day, and $\mathrm{IE_{ti}} = 0$ otherwise. The indicator was added to each model part:

$$\eta_{ti} = \alpha_0 + \alpha_{\mathrm{IE}}\mathrm{IE_{ti}} + a_i,$$

$$m_{ti} = \beta_0 + \beta_{\mathrm{IE}}\mathrm{IE_{ti}} + b_i + e_{ti},$$

where α_0 and β_0 are the population log odds of watching TV and the expected mean log time spent when TV was watched, respectively, on days other than the first interview day and for a subject whose respective random effects (a_i and b_i) are equal to 0. α_{IE} and β_{IE} are the expected differences in the population log odds of watching TV and mean log time spent watching TV when watched, respectively, between the first interview day and the remaining days. If α_{IE} and β_{IE} are greater than 0, then this is taken to reflect an IE effect on the likeliness to report watching TV and the (log) time spent when watched, respectively.

Adding positive and negative affect measures and other covariates

Daily measures of positive and negative affect, indicators of the days of the week, participant's age and biological sex were added to the model. Indicators of the days of the week were centered within-subject to interpret their effects as within-subject effects (Raudenbush and Bryk 2002).[1] For an indicator of a particular day of the week, the individual's mean across days represents the proportion of interview days that corresponded to that day of the week. Consequently, the within-subject effect is interpreted as the typical (across days and subjects) effect of a given day of the week on the outcome variable. Daily measures of positive and negative affect were centered about their respective person-level means to test for within-subject effects of each measure. Within-subject effects of positive and negative affect measures are interpreted as the typical (across days and subjects) effect of an affect measure on a given day. The person-level mean of each affect score (calculated by averaging values across days) was included to test for between-subject effects of each affect measure. These person-level means were centered about the respective mean of the individual-level means. Between-subject effects of positive and negative affect measures are interpreted as the typical (across subjects) effect of the daily average of an affect measure across days. The within- and between-subject effects are mathematically independent (Raudenbush and Bryk 2002). Age was centered about the sample mean age of 48 years.

The logit was specified as

$$\begin{aligned}\eta_{ti} = {} & \alpha_0 + \alpha_{IE}\mathrm{IE}_{ti} + \sum_{1}^{k}\alpha_{1k}\mathrm{weekday}_{tik} + \alpha_2\mathrm{age}_i \\ & + \alpha_3\mathrm{female}_i + \alpha_{4a}\mathrm{PA}_{ti} + \alpha_{4b}\overline{\mathrm{PA}}_i + \alpha_{5a}\mathrm{NA}_{ti} \\ & + \alpha_{5b}\overline{\mathrm{NA}}_i + \alpha_{6a}\mathrm{IE}_{ti} * \mathrm{PA}_{ti} + \alpha_{6b}\mathrm{IE}_{ti}\overline{\mathrm{PA}}_i \\ & + \alpha_{7a}\mathrm{IE}_{ti} * \mathrm{NA}_{ti} + \alpha_{7b}\mathrm{IE}_{ti}\overline{\mathrm{NA}}_i + a_i,\end{aligned}$$

where IE_{ti} is an indicator of whether the logit was measured on the first interview day or not, $\mathrm{weekday}_{tik}$ is indicator k corresponding to a particular day of the week when the logit was measured, age_i is the individual's age centered to 48 years, female_i is an indicator of whether the individual is female, PA_{ti} is the within-subject centered value of positive affect on day t for individual i, $\overline{\mathrm{PA}}_i$ is the daily mean of positive affect for individual i, NA_{ti} is the within-subject centered value of negative affect on day t for individual i, $\overline{\mathrm{NA}}_i$ is the daily mean of negative affect for individual i, $\mathrm{IE}_{ti} * \mathrm{PA}_{ti}$ is the interaction between the indicator of the first interview day and the within-subject centered value of positive affect at day t for individual i, $\mathrm{IE}_{ti} * \overline{\mathrm{PA}}_i$ is the interaction between the indicator of the first interview day and the daily mean of positive affect for individual i,$\mathrm{IE}_{ti} * \mathrm{NA}_{ti}$ is the interaction between the indicator of the first interview day and the within-subject centered value of negative affect at day t for individual i, and $\mathrm{IE}_{ti} * \overline{\mathrm{NA}}_i$ is the interaction between the indicator of the first interview day and the daily mean of negative affect for individual i.

Interpretation of the model coefficients is as follows: α_0 is the population log odds of watching TV on a day other than the first interview day for a male who was 48 years old, surveyed on a Sunday, whose daily-average affect scores were equal to the respective sample means, and whose random effect α_i was equal to 0. Unlike the preceding model in which the effect of the first interview day was the only covariate included in the model, this model included the moderating effects of the first interview day on the relationships between affect measures and TV use (in addition to other covariates). Consequently, α_{IE} is the expected difference between the first interview day and the remaining days in the log odds of watching TV for a subject whose daily means of both positive and negative affect scores were equal to the respective sample mean[2]; α_{1k} is the within-subject effect of the kth day of the week relative to Sunday, α_2 is the between-subject effect of age, and α_3 is the mean difference in the logit between males and females. The coefficients α_{5a} and α_{5b} are the within- and between-subject effects of positive affect, respectively. The coefficients α_{6a} and α_{6b} are the within- and between-subject effects of negative affect, respectively. The coefficients α_{7a} and α_{7b} are the interaction effects between the first interview day and the within- and between-measures of positive affect, respectively, and α_{8a} and α_{8b} are the interaction effects between the first interview day and the within- and between-subject measures of negative affect, respectively. Unless otherwise specified, covariate effects are adjusted for the effects of other covariates. Finally, a_i is a random subject effect that indicates how much an individual's log odds, conditional on the covariates, differs from the

[1] Within-subject centering of a time-varying covariate requires that the mean of the variable across occasions is calculated separately for each individual and the result subtracted from the individual's daily measures.

[2] A within-subject centered covariate yields a model intercept that is not adjusted for the covariate. By including the subject-level daily mean of the covariate, the model's intercept is adjusted for the effect of the covariate.

population log odds. The random effect a_i is assumed to be i.i.d normal with a mean of 0 and variance ϕ_a^2 that quantifies the degree to which individuals vary in their log odds of watching TV conditional on covariates.

The conditional measure of time spent watching TV was modeled as

$$\begin{aligned} m_{ti} =& \beta_0 + \beta_{IE}IE_{ti} + \sum_{1}^{k} \beta_{1k}weekday_{tik} + \beta_2 age_i \\ &+ \beta_3 female_i + \beta_{4a}PA_{ti} + \beta_{4b}\overline{PA}_i + \beta_{5a}NA_{ti} \\ &+ \beta_{5b}\overline{NA}_i + \beta_{6a}IE_{ti} * PA_{ti} + \beta_{6b}IE_{ij}\overline{PA}_i \\ &+ \beta_{7a}IE_{ti} * NA_{ij} + \beta_{7b}IE_{ti}\overline{NA}_i + e_{ti}, \end{aligned}$$

where β_0 is the population log mean time spent for a male who was 48 years old and surveyed on a Sunday and whose random effect b_i was equal to 0, where b_i indicates how much an individual's mean log time use, conditional on the covariates, differed from the population mean log time. Similar to the logit model part, this model part includes the moderating effects of the first interview day on the relationships between the affect measures and log time spent watching TV. Consequently, the first interview day effect is interpreted as the typical (across days and subjects) effect for a subject whose daily means of positive and negative affect scores were at the respective sample means. The coefficient β_{IE} is the expected difference between the first interview day and the remaining days in the mean log time spent watching TV (conditional that TV was watched) for a subject whose daily mean positive and negative affect scores are equal to the respective sample means; β_{1k} is the within-subject effect of the kth day of the week relative to Sunday (e.g. β_{11} is the within-subject difference in the reported mean log time spent watching TV on a Monday versus a Sunday), β_2 is the between-subject effect of age, and β_3 is the between-subject effect of biological sex. The coefficients β_{4a} and β_{4b} are the within- and between-subject effects of positive affect, respectively. The coefficients β_{5a} and β_{5b} are the within- and between-subject effects of negative affect, respectively. The coefficients β_{6a} and β_{6b} are the interaction effects between the first interview day and the within- and between-measures of positive affect, respectively, and β_{7a} and β_{7b} are the interaction effects between the first interview day and the within- and between-subject measures of negative affect, respectively. Unless otherwise specified, the interpretation of covariate effects are adjusted for the effects of other covariates. The random effect b_i is assumed to be i.i.d. normal with a mean of 0 and variance ϕ_b^2 that quantifies the degree to which individuals vary in their reported mean log time spent, conditional on the covariates. Finally, the residual e_{ti} is assumed to be independent between and within individuals with mean = 0 and variance σ_e^2, where σ_e^2 was assumed to be a function of a random scale effect to permit heterogeneity of the day-to-day variability in reported conditional measures of the log time spent on days when TV was watched.

Approach to model comparisons

For each model, we report the − 2*loglikelihood (− 2lnL) and relative fit indices, namely the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). When deemed warranted, pairwise model comparisons are carried out using an approximation to the Bayes Factor (BF) (Kass and Wasserman 1995; Nagin 1999):

$$\widehat{BF}_{A,B} = \exp\left(SIC_A - SIC_B\right),$$

where SIC_A and SIC_B are based on two selected models with SIC = − 0.5BIC. The value of $\widehat{BF}_{A,B}$ is the ratio of the estimated probability of Model A being the correct model to Model B being correct. A $\widehat{BF}_{A,B} = 10$, for example, is taken as evidence that Model A is 10 times more likely than Model B to be the correct model.

Results

Table 1 includes descriptive statistics of daily measures, including the sample sizes according to biological sex and the proportions of missing data by interview day and biological sex. Figure 1 displays daily reported time spent watching TV.

The first model, Model 1, assumed scores of the second model part followed a lognormal distribution with a loglikelihood function defined by (cf: Blozis et al. 2019)

$$L_i \propto \frac{1}{\sqrt{2\pi\sigma_e^2\left(y_{ti}\right)}}\exp\left(-\frac{1}{2}\left(\frac{\left(\ln\left(y_{ti}\right)-\mu_i\right)^2}{\sigma_e^2}\right)\right).$$

Although the natural log transformed scores reduced skew, scores remained slightly positively skewed, and a second model (Model 2) was applied that assumed scores in the second model part followed a 2-class mixture of two lognormal distributions with unknown and mutually exclusive class membership for subject i, with a likelihood function for subject i in class j = 1,2 defined as

$$L_{ij} \propto \frac{1}{\sqrt{2\pi\sigma_{e_j}^2\left(y_{tij}\right)}}\exp\left(-\frac{1}{2}\left(\frac{\left(\ln\left(y_{tij}\right)-\mu_{ij}\right)^2}{\sigma_{e_j}^2}\right)\right),$$

where subject i's expected value and residual variance were class specific: $\mu_{ij} = \beta_{0j} + b_{0ij}$ and $\sigma_{e_j}^2$, respectively. The logis-

Table 1 Daily sample descriptive statistics (n = 782)

	Interview Day	1	2	3	4	5	6	7	8
Females	n(% Observed)	346(44.4)	340(44.9)	336(44.9)	336(45.1)	329(44.8)	313(44.7)	315(45.1)	313(45.8)
Males	n(% Observed)	434(55.6)	417(55.1)	413(55.1)	409(54.9)	406(55.2)	388(55.3)	383(54.9)	371(54.2)
total	n (% Missing)	780(0.3)	757(3.2)	749(4.2)	745(4.7)	735(6.0)	701(10.4)	698(10.7)	684(12.5)
TV use (yes/no)									
Females	No. users (% users)	382(88.0)	329(80.1)	317(80.1)	316(80.4)	315(81.2)	319(82.2)	312(81.5)	286(77.3)
Males	No. users (% users)	305(88.2)	285(84.8)	282(86.2)	267(81.4)	264(83.5)	255(81.5)	266(84.7)	264(84.4)
Time spent watching TV									
Females	mean/median	2.7/2	2.2/2	2.2/2	2.2/2	2.1/2	2.1/2	2.2/2	2.7/2
	SD	2.4	1.9	2.0	2.0	1.8	1.6	1.7	1.8
	Min/max	0.08/16	0.03/10	0.08/16	0.17/12	0.08/18	0.17/8	0.08/14	0.17/14
Males	Mean/median	2.6/2	2.1/1.5	2.2/1.5	2.2/2	2.2/2	2.3/2	2.2/2	2.2/2
	SD	2.4	2.2	2.2	2.2	2.0	2.0	1.8	2.2
	Min/max	0.17/20	0.08/20	0.17/23	0.17/20	0.08/18	0.08/16	0.17/13	0.17/22
Positive Affect									
Females	Mean/median	2.5/2.5	2.6/2.7	2.5/2.7	2.5/2.6	2.5/2.7	2.5/2.6	2.5/2.6	2.5/2.5
	SD	0.81	0.82	0.89	0.88	0.89	0.87	0.88	0.89
	Min/max	0/4	0.08/4	0.23/4	0/4	0/4	0/4	0/4	0/4
Males	Mean/median	2.5/2.6	2.6/2.5	2.5/2.6	2.5/2.6	2.5/2.6	2.5/2.6	2.5/2.6	2.5/2.5
	SD	0.79	0.81	0.80	0.85	0.85	0.81	0.83	0.83
	Min/max	0.15/4	0/4	0.08/4	0.08/4	0.31/4	0.31/4	0.23/4	0.31/4
Negative affect									
Females	Mean/median	0.35/0.21	0.26/0.14	0.23/0.14	0.22/0.14	0.22/0.07	0.20/0.07	0.18/0.07	0.20/0.07
	SD	0.44	0.37	0.35	0.35	0.36	0.31	0.27	0.31
	Min/max	0/2.9	0/3.0	0/2.4	0/2.9	0/2.6	0/2.6	0/1.6	0/3.1
Males	Mean/median	0.32/0.21	0.24/0.14	0.20/0.07	0.20/0.07	0.20/0.07	0.19/0.07	0.17/0.07	0.16/0
	SD	0.40	0.36	0.34	0.33	0.35	0.35	0.29	0.28
	Min/max	0/3.4	0/2.9	0/2.8	0/2.5	0/2.6	0/3.3	0/2	0/1.9

tic model part was unchanged. The logliklihood was the log sum of the two class likelihoods weighted by their respective class proportions:

$$\ln L_{ij} \propto \ln \left(L_{i1}\pi_1 + L_{i2}\left(1 - \pi_1\right) \right),$$

where π_1 and $\pi_2 = 1 - \pi_1$ are the population proportions in class 1 and 2, respectively. The aim in fitting this alternative model was only to account for unexplained heterogeneity of scores and not to assign individuals to classes. Comparisons between models that involve a mixture of response distributions are often based on a comparison of BIC values. Here, the difference in BIC values between Models 1 and 2 was 69, indicating a mixture of lognormal distributions for the second model part provided a better fit to the data. The approximated Bayes Factor value[3] indicated that Model 2 was 9.6E14 times more likely than Model 1 to be the correct model. Model 2 was retained for subsequent analyses. ML estimates of Models 1 and 2 are provided in the first and second columns of estimates of Tables 2 and 3,[4] respectively. Across days and individuals, the estimated probability that an individual watched TV was Prob $= \frac{\exp(2.6)}{1+\exp(2.6)} = .93$. For the continuous model part that was based on a 2-class mixture of lognormal distributions, the estimated proportions of individuals in the two classes were 0.06 and 0.94, respectively, and the estimated mean difference between the two distributions was 1.05 (95% CI 0.86, 1.23). Mixing across the two classes of distributions, the estimated time spent watching TV when respondents reported to watch was $\exp = (-0.61 * .06 + (-0.61 + 1.0) * .94) = 1.39$ hours.

[3] $\widehat{\mathrm{BF}}_{2,1} = \exp(-9101 + 9135.5) = 9.6\mathrm{E}14$.

[4] Although estimates of the two model parts appear in separate tables, the two model parts were estimated simultaneously. For the sake of brevity in this report, we provide estimates of the fixed effects and make estimates of the variances and covariances of all models that were fit available upon request.

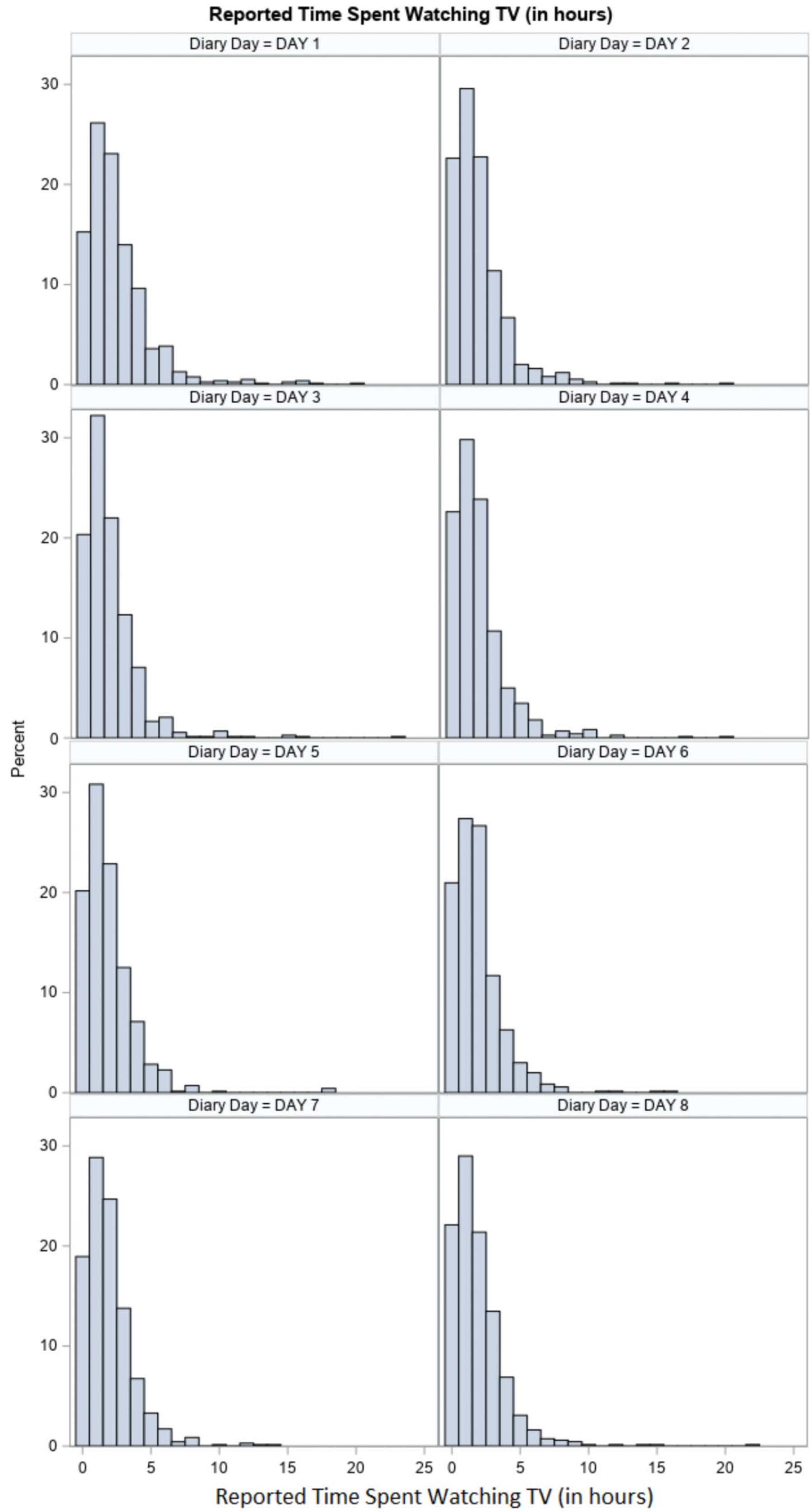

Fig. 1 Daily reported time spent watching TV (n = 782)

Table 2 ML estimates (95% CI) of the logistic model part for TV use (yes/no) (n = 782)

Log odds of use	Model 1 Est(95% CI)	Model 2 Est(95% CI)	Model 3 Est(95% CI)	Model 4 Est(95% CI)	Model 5 Est(95% CI)
α_0	2.6(2.4, 2.8)	2.6(2.4, 2.8)	2.5(2.3, 2.7)	2.9(2.5, 3.2)	2.9(2.6, 3.3)
fd_{ij},α_{IE}			0.79(0.50, 1.1)	− 0.76(− 2.1, 0.60)	
Mon_{ij},α_{1a}				− 0.27(− 0.60, 0.06)	− 0.14–0.47, 0.19
Tue_{ij},α_{1b}				− 0.17(− 0.50, 0.17)	− 0.09(− 0.42, 0.24
Wed_{ij},α_{1c}				− 0.22(− 0.55, 0.12)	− 0.15(− 0.48, 0.19
Thu_{ij},α_{1d}				− 0.08(− 0.42, 0.27)	− 0.05(− 0.39, 0.29)
Fri_{ij},α_{1e}				− 0.07(− 0.42, 0.27)	− 0.05(− 0.39, 0.29)
Sat_{ij},α_{1f}				− 0.35(− 0.69, − 0.01)	− 0.34(− 0.68, − 0.00)
Age_i,α_2				0.02(0.01, 0.04)	0.02(0.01, 0.04)
$Female_i,\alpha_3$				− 0.40(− 0.75, − 0.05)	− 0.38(− 0.72, − 0.04)
PA_{ij}, α_{4a}				− 0.14(− 0.42, 0.14)	− 0.02(− 0.27, 0.23)
$\overline{PA}_i, \alpha_{4b}$				− 0.37(− 0.65, − 0.09)	− 0.30(− 0.56, − 0.03)
NA_{ij}, α_{5a}				− 0.71(− 1.2, − 0.24)	− 0.21(− 0.63, 0.21
$\overline{NA}_i, \alpha_{5b}$				− 0.40(− 1.1, 0.35)	− 0.24(− 0.93, 0.45)
$PA_{ij} * fd_{ij}, \alpha_{6a}$				0.31(− 0.52, 1.1)	
$\overline{PA}_i, * fd_{ij}, \alpha_{6a}$				0.54(0.08, 1.0)	
$NA_{ij} * fd_{ij}, \alpha_{7a}$				2.1(0.64, 3.6)	
$\overline{NA}_i, * fd_{ij}, \alpha_{7a}$				0.38(− 0.84, 1.6)	
− 2lnL	18,211	18,122	18,021	17,798	17,930
AIC	18,229	18,146	18,049	17,890	18,002
BIC	18,271	18,202	18,114	18,104	18,170

Model 3 included the first interview day effects in the logistic and generalized linear models. Estimates of this model are in the third columns of estimates of Tables 2 and 3. The difference in BIC values between Models 2 and 3 was 88, indicating that overall, including the effect of the first interview day improved model fit. Further, the approximated Bayes Factor value[5] indicated that Model 3 was 1.9E19 times more likely than Model 2 to be the correct model. For the logistic model part (see the third column of estimates in Table 2), the log odds that an individual watched TV on their first interview day was estimated to be higher (est[6] = 0.79, 95% CI 0.50, 1.1) relative to the expected log odds across all subsequent interview days. Converting log odds to probabilities, the probability (across days and individuals) that an individual reported watching TV on the first interview day 0.96, whereas the probability was 0.92 across the following days.

For the continuous model part (see the third column of estimates in Table 3), the reported mean log time spent watching TV when TV was watched on the first interview day was estimated to be greater (0.16, 95% CI 0.12, 0.20) relative to the mean log reported time spent across all subsequent days. Converting the mean log time spent to hours, the estimated mean time spent watching TV on the first interview day if TV was watched was about 2 h $(\exp = ((-0.67 + 0.16) * .06 + (-0.67 + 0.16 + 1.1) * .94) = 1.96)$. Conversely, time spent on the days that followed was, on average, estimated to be about a half hour less $(\exp = (-0.67 * .06 + 0.43 * .94) = 1.44$ hours). Thus, both aspects of reported TV use were elevated for the first interview day relative to the following interview days.

Model 4 included the first interview day effect in addition to measures of positive and negative affect and covariates, as described previously. Our focus was on the interpretations of the within- and between-subject effects of positive and negative affect and the moderating effects of the first interview day on those relationships. These relationships are illustrated in Fig. 2.[7] As shown in Table 2 (see the 4th columns of estimates), the first interview day did not clearly moderate the within-subject effect of positive affect

[5] $\widehat{BF}_{3,2} = \exp(-9057 + 9101) = 1.9E19$.

[6] est = estimate.

[7] To simplify the illustration of Model 4, Fig. 2 does not show the effects of age, biological sex and the indicators of the days of the week, although these variables were included in the model that was fitted to the data.

Table 3 ML estimates (95% CI) of the generalized linear model part for time spent watching TV (n = 782)

Log mean time	Model 1 Est(95% CI)	Model 2 Est(95% CI)	Model 3 Est(95% CI)	Model 4 Est(95% CI)	Model 5 Est(95% CI)
β_0	0.38(0.33, 0.43)	− 0.61(− 0.81, − 0.40)	− 0.67(− 0.85, − 0.48)	− 0.58(− 0.75, − 0.41)	− 0.55(− 0.73, − 0.38)
Diff		1.0(0.86, 1.2)	1.1(0.91, 1.2)	1.1(0.90, 1.2)	1.1(0.89,1.2)
π_1		.06(.03, .09)	0.06(0.03, 0.08)	0.06(0.04, .09)	0.06(0.04, 0.09)
fd_{ti},α_{IE}			0.16(0.12, 0.20)	0.04(− 0.14, 0.21)	
Mon_{ij},α_{1a}				− 0.14(− 0.19, − 0.09)	− 0.11(− 0.16, − 0.07)
Tue_{ij},α_{1b}				− 0.19(− 0.24, − 0.14)	− 0.17(− 0.22, − 0.12)
Wed_{ij},α_{1c}				− 0.25(− 0.30, − 0.20)	− 0.24(− 0.30, − 0.20)
Thu_{ij},α_{1d}				− 0.25(− 0.30, − 0.20)	− 0.24(− 0.29, − 0.19)
Fri_{ij},α_{1e}				− 0.21(− 0.26, − 0.16)	− 0.20(− 0.26, − 0.16)
Sat_{ij},α_{1f}				− 0.09(− 0.15, − 0.04)	− 0.10(− 0.15, − 0.05)
Age_i,α_2				0.01(0.01, 0.01)	0.01(0.01, 0.01)
$Female_i$,α_3				− 0.01(− 0.07, 0.05)	− 0.00(− 0.06, 0.06)
PA_{ij}, α_{4a}				0.00(− 0.04, 0.04)	0.01(− 0.02, 0.05)
$\overline{PA_i}$, α_{4b}				− 0.06(− 0.10, − 0.02)	− 0.05(− 0.09, 0.00)
NA_{ti}, α_{5a}				− 0.03(− 0.11, 0.05)	0.04(− 0.02, 0.11)
$\overline{NA_i}$, α_{5b}				0.17(0.08, 0.26)	0.19(0.11, 0.27)
$PA_{ij} * fd_{ij}$, α_{6a}				0.05(− 0.05, 0.14)	
$\overline{PA_i}, * fd_{ij}$, α_{6a}				0.05(− 0.01, 0.11)	
$NA_{ti} * fd_{ti}$, α_{7a}				0.09(− 0.12, 0.30)	
$\overline{NA_i}, * fd_{ij}$, α_{7a}				0.02(− 0.12, 0.17)	

'Diff' is the difference in means between the two lognormal response distributions for the two-class mixture. $\pi 1$ is the proportion of subjects in the first of the two classes for the two-class mixture. Model fit indices are reported in Table 2. Model 1 assumes a lognormal response distribution. Models 2–5 assume a 2-class mixture of lognormal distributions

(est = 0.31, 95% CI −0.52, 1.1), although it did moderate the between-subject effect of positive affect on the log odds of watching TV (est = 0.54, 95% CI 0.08, 1.0). These results indicate that on the first interview day, the direction of the association between the daily average of positive affect on the likeliness to watch TV was not clear (est = − 0.59, 95% CI −1.6, 0.43), whereas on days following the first interview day, individuals tended to have an increased likeliness to report TV use if they had a relatively low positive affect level across those days (est = − 0.37, 95% CI −0.65, −0.09). Regarding the within-subject effect of positive affect, there was no clear relationship between affect and the likeliness to watch TV (est = − 0.14, 95% CI − 0.42, 0.14).

The effect of the first interview day clearly moderated the within-subject effect (est = 2.1, 95% CI 0.64, 3.6) but not the between-subject effect (est = 0.38, 95% CI −0.84, 1.6) of negative affect on the log odds of reported TV use. Specifically, the direction of the within-subject effect of negative affect on the first interview day was not clear (est = 0.65, 95% CI −0.99, 2.3), whereas the within-subject effect for subsequent interview days was negative (est = − 0.71, 95% CI −1.2, −0.24). From this, there was no clear impact of negative affect on the likeliness to watch TV on the first interview day, but on days following the first interview, there was a clear increase in the likeliness to watch TV if one also had a relatively high negative affect level on the same day (est = 2.10, 95% CI 0.80, 3.4). Regarding the between-subject effect of negative affect, there was no clear relationship between affect and the likeliness to watch TV (est = − 0.40, 95% CI − 1.1, 0.35).

For time spent watching TV on days when TV was watched, the first interview day did not clearly moderate any of the within- or between-subject effects of positive or negative affect. Positive and negative affect did, however, relate to the time spent watching TV. The daily average of positive affect was negatively related to an individual's average time spent watching TV on days when TV was watched (est = − 0.06, 95% CI −0.10, −0.02), and the daily average of negative affect was positively related to the individual's average time spent watching TV (est = 0.17, 95% CI 0.08, 0.26). Neither of the within-subject measures of affect were clearly related to time spent watching TV.

Lastly, we fit a model based on Model 4 that excluded all IE effects to evaluate how ignoring the effects of the first interview day on parameter estimates. From this model, Model 5, we evaluated the sensitivity of parameter estimates of the within- and between-subject effects of positive

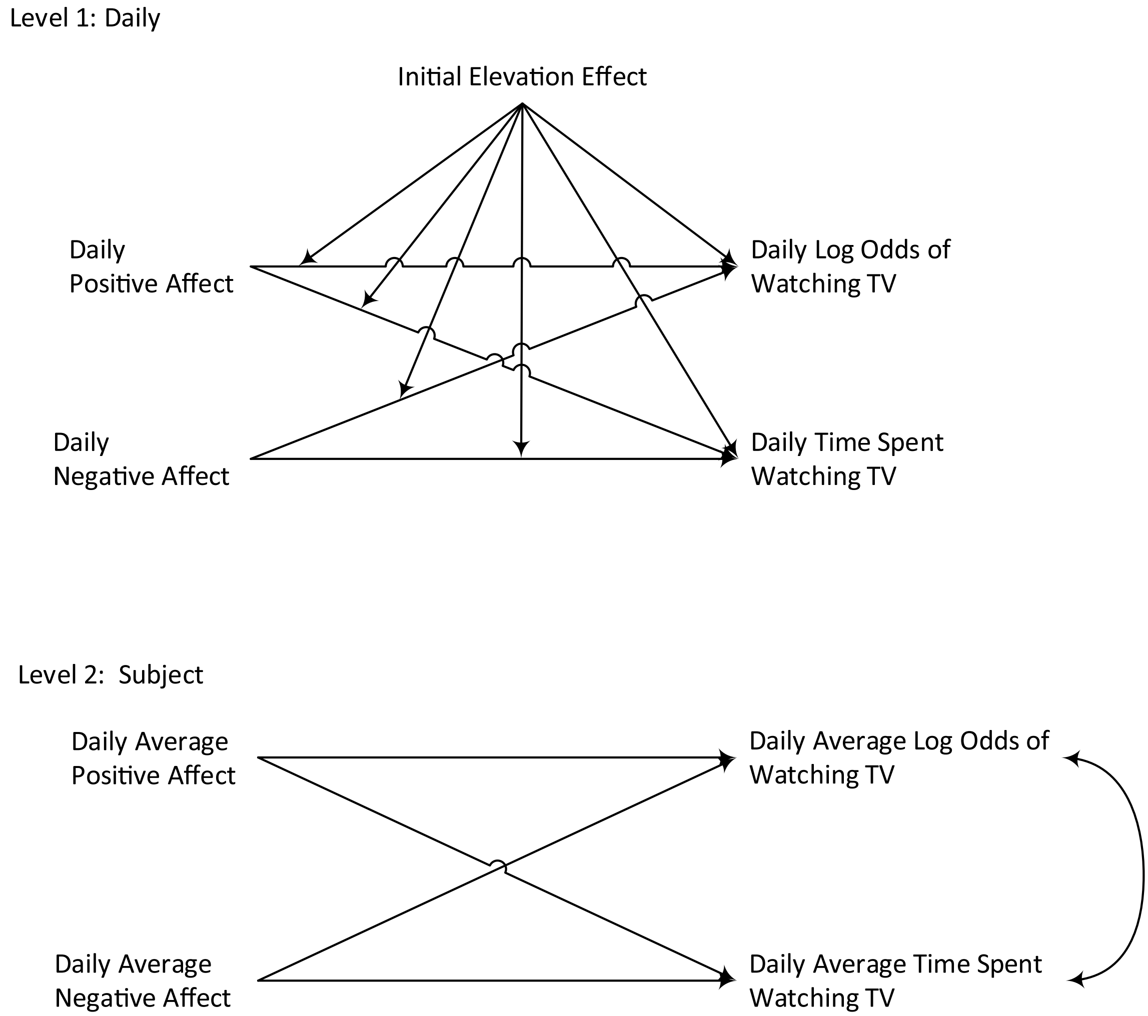

Fig. 2 An illustration of a two-part mixed-effects model for daily TV use reports

and negative affect on the likeliness to watch TV and the reported time spent when TV was watched. Estimates of Model 5 are given in the last columns of Tables 2 and 3. The difference in BIC values between Models 4 and 5 was 66, indicating an overall improvement in model fit by including the first interview day effects in both model parts. The approximated Bayes Factor value[8] indicated that Model 4 was 2.1E14 times more likely than Model 5 to be the correct model, supporting the importance of accounting for the first interview day effects.

From the last column of estimates in Table 2, the estimated between-subject effect of negative affect differs from that obtained under Model 4, but the conclusion is generally the same in that the direction of the effect is not clear as both estimated 95% CIs span from a negative to a positive value. The estimated within-subject effect of negative affect under Model 5 suggests an unclear direction in the effect. This is in contrast to Model 4 that indicates a clear direction in the effect but specifically for the days that follow the first interview and not the first interview day. The estimated between-subject effect of positive affect differs from that obtained under Model 4, but both models indicate a negative effect on the likeliness to watch TV. The estimated effect, however, is reduced under Model 5. Finally, the estimated within-subject effect of positive affect under Model 5 suggests an unclear direction in the effect, and this is consistent with Model 4.

Discussion

Research in psychology using self-report data from daily diary studies has illuminated the potential for an elevated mean response at the first interview where no such patterns are expected. The effect has been shown to impact statistical inference about a variety of self-report measures,

[8] $\widehat{\mathrm{BF}}_{4,5} = \exp(-9052 + 9085) = 2.1\mathrm{E}14.$

including psychological state measures, health symptoms and objective behavioral reports. Marketing research, especially research involving consumer perceptions and measures of relatively objective behaviors, relies on self-report data, suggesting a need to investigate IE response patterns to understand the impact on statistical inference and subsequent marketing strategies that might be developed from such investigations.

We analyzed daily self-report measures of time spent watching TV using data collected from an observational study where participants were surveyed for 8 consecutive days. Using a two-part MELS model, we documented that on the first survey day, participants were more likely to report that they watched TV and that they tended to report more time spent watching TV (if TV was watched) compared to the days that followed. For reports of time spent watching TV on days when TV was watched, the estimated mean time spent on the first interview day was about two hours, whereas the estimated mean time spent across the interview days that followed was about 1.4 h. These analyses were based on an observational study. Further, participants started their daily diary survey on different days of the week and at different times of a calendar year, and so there was no reason to anticipate that responses would differ between the first interview day and those that followed. The findings present potential implications for self-reported daily diary data used in marketing research, but they also pose implications for cross-sectional research where it is not possible to estimate or account for the magnitude of an IE effect if it is present.

Market research has benefitted from including affective measures in studies to better understand consumer behavior, and we explored the role of affect in TV use. Specifically, we studied the effects of IE on the relationships between affect and TV use. Using a two-part MELS model, we tested for moderating effects of the first-interview day on the within- and between-subject effects of positive and negative affect on the likeliness to watch TV and the reported amount of time spent on days when TV was watched (Model 4). A within-subject effect refers to the typical (across days and subjects) relationship between an outcome and a covariate at a given occasion. A between-subject effect refers to the typical (across subjects) relationship between an outcome's tendency across days and the tendency of a covariate across days.

With regard to the likeliness to watch TV, we found that the first interview day did not impact the daily relationship between positive affect and the likeliness to watch TV. It did, however, impact the subject-level relationship. That is, a relatively low daily average for positive affect corresponded to a greater likeliness to watch TV but only on days following the first interview. In other words, the daily average of positive affect had no detectable effect on the likeliness to watch TV on the first interview day but it did on the following days. When the moderating effect of the first interview day was ignored (Model 5), this relationship was replicated, but the magnitude of the effect of positive affect on the likeliness to watch TV was reduced.

We also found an increased likeliness for participants to report watching TV if they also had a relatively high level of negative affect during that same day, but this was the case only for the days that followed the first interview day and not the first interview day itself (Model 4). This is a clear example of the importance of accounting for the effect of the first interview day because the effect of negative affect was only detectable for the set of days that followed the first interview. When the moderating effect of the first interview day was ignored (Model 5), there was no clear direction for the relationship between negative affect and the likeliness to watch TV. This further stresses the importance of accounting for the first interview day in understanding the relationship between daily negative affect and daily TV use because the relationship would have otherwise been missed.

The second part of a two-part model for a semi-continuous variable permits the study of the relationships between covariates and the level of the response when the response is positive. In this study, the second model part was for the reported amount of time spent watching TV on days when TV was watched. We tested the strength of the relationships between within- and between-subject measures of positive and negative affect and the reported amount of time spent, while attending to the effects of the first interview day. Unlike the first model part that was used to model the likeliness to watch TV, the effect of the first interview day did not clearly moderate the relationships between the affect measures and time spent watching TV. The clearest and strongest relationship was that between the daily average of negative affect and time spent watching TV. That is, individuals whose daily average negative affect was high tended to spend more time watching TV. Although the daily average of positive affect was clearly related to time spent watching TV under the model that took the first interview day into account (Model 4), the lower bound of the 95% CI was close to 0. This relationship was not detected under the model that ignored the effect of the first interview day (Model 5).

Managerial implications

The time consumers spend with media, whether the engagement is through TV or other screened devices, is significant for media planners. Obtaining self-reported usage through surveys is a common research tool that media planners use. Using daily diary survey data, this study reported a difference of about 30 min more per day, on average, in the time spent watching TV based on data from the first interview day relative to the daily average based on data from the interview days that followed. Although a difference of 30 min may seem trivial, the value of this result lies in aggregating daily estimates across a week. That is, the estimated time spent based on data from the first interview day would suggest consumers spend about 14 h per week watching TV, whereas the estimated time spent based on data from the 2nd to the final interview day would suggest that consumers spend just under 10 h per week. This difference is especially important for studies that rely on a single interview to estimate weekly totals, as the result implies that a single interview can result in overestimated values of the reported behavior. For planners who use usage reports to help develop media strategies, time spent engaged with media provides insight into how reach and frequency strategies should be implemented. That is, when users spend more time with media, a frequency strategy is attractive because greater time spent with media provides more opportunities to expose viewers to a brand's message through a smaller number of media vehicles. Conversely, when users spend less time engaged with media, a reach strategy may be more attractive as a larger number of unique viewers have the opportunity to be exposed to a brand's message. This strategy is met by placing a brand message on more media vehicles. Based on TV use reports presented here, the results suggest that a frequency strategy could be over-emphasized and the need for a reach strategy undervalued.

The reported difference in time spent watching TV between the first interview and the interview days that followed have important implications for media planners. First, if data are collected at a single occasion, time estimates may be overestimated, as was found for the data set studied here. Second, without collecting repeated measures, it is not possible to test for the possibility of an elevated report in the initial assessment. Consequently, media strategies, such as reach and frequency, may not be well informed if the first report is biased upward. With evidence that responses on the first interview day differed from subsequent interview days, it can be important for media planners to consider data collection across multiple days. Not only does a study design that uses repeated measures allow for testing of the presence of a first-interview effect, such a design offers opportunities to better understand the behaviors of individual consumers, especially regarding behaviors that are known to vary at the individual level across days.

As is well known, consumer affect is an integral part of understanding media use. In this study, reports based on data from the first interview suggested that neither positive nor negative affect was related to an individual's likeliness to watch TV. Positive and negative affect were, however, related to the likeliness to watch TV based on data from the days following the first interview. The implication of this finding is important. That is, a study based on a single assessment could result in the conclusion that neither positive nor negative affect impacts a consumer's likeliness to watch TV. Consequently, media planning going forward might mismatch an advertisement message frame with the media vehicle.

This study also found that individuals whose negative affect levels tended to be relatively high across days tended to spend more time watching TV, and that on a given day, individuals whose positive affect levels were relatively high tended to spend less time watching TV on the same day (though the latter of these two effects is smaller). Taken together, these findings suggest that viewers who tend to be higher on negative affect exceed the number of viewers who tend to be higher on positive affect. This has implications for reach and frequency strategies, ad message framing, and media vehicle selection. Considering negative affect, for example, brands will likely select a frequency media strategy given that adults tend to spend more time watching TV. Further, in their reach strategy, brands may want to frame their message in a manner that would be meaningful for viewers higher on negative affect. Finally, brands would then want to select an appropriate vehicle in which to place their message.

Concluding remarks

An important aim in collecting daily diary data is to document the behavioral tendencies of individuals and the day-to-day variation in those behaviors. Gaining this insight about consumer behavior from repeated-measures data is increasingly recognized as an important component of some types of marketing research. From this report and previous studies, it is clear that data collection using repeated measures can illuminate a greater understanding of the development and variability in individual behaviors. When data are based on self-reports, however, unexpected differences in responses between the first interview and those that follow can have consequences for statistical inference. Using observational TV use data from a large and representative sample of adults in the U.S., the effects of the first interview day were evident, as there was a greater tendency for respondents to report having watched TV on the first interview day relative to the days that followed. When TV was watched, respondents tended to report more time spent watching TV on the first interview day relative to the days that followed. The effects of the first interview day were also evident in tests of the relationships between self-report measures of positive and negative affect and aspects of TV use. These findings are relevant for media-use studies that rely on self-report measures at a single point in time (and naturally are not set up to test for the effects of the first interview), as well as studies that use repeated measures and do not account for the potential effects of the first interview.

This study supports the use of study designs that collect repeated measures and the need to account for the effects of the first interview. Here, we accounted for the first interview effect by including an indicator of the first interview day in the statistical model. Another approach may be to disregard data recorded for the initial interview and the analysis performed on the data observed for the remaining days. For an experimental study design, researchers might consider assessing study participants twice to assess possible effects of the first interview and to gain realistic measures of baseline values. As a greater number of repeated measures would naturally decrease the impact of an IE effect, studies may consider increasing the number of repeated measures as the difference between the mean response of the first day would eventually be outweighed by the means on the days that follow.

Future directions stemming from the current work are plentiful. First, applications of the methods describe here could be applied to other types of media use reports to test for the effects of the first interview effect. Second, as this report relied on a large and representative sample of adults in the U.S. population, future work could explore the needed sample size to reliably detect effects of the first interview on summary measures of media use reports. Third, this report relied on measures of psychological state variables, namely positive and negative affect, that were based on the averages of responses to survey item sets developed to reflect the respective underlying latent variables. Future work could expand on the models used here to include the capacity to account for measurement error in the observed responses to the survey data that represent exogenous latent variables. Fourth, we selected a two-part mixed-effects model with a random scale to analyze daily measures of TV use because we thought that it was important to distinguish between TV use and time spent watching TV when it was watched. A mixed-effects model accounts for within-subject dependencies that are common in repeated measures. By also including a random scale, the model accounted for between-subject heterogeneity of the day-to-day variance in reports of time spent watching TV. Given that TV use and time spent are treated as unique features of media use behavior, this model permitted covariates to have unique relationships with these two aspects of TV use. This made it possible to study predictors of whether or not a person watched TV, and conditional on having watched TV, predictors of the reported time spent watching. Naturally, there are other choices in modeling frameworks, and the sensitivity of the IE effects might have been different had a different statistical model been selected. Fifth, future work might consider expanding a survey study to also include objective measures of media use to better understand media use data obtained through these different methods. Though potentially costlier to obtain than self-report data, objective measures should eliminate concerns over self-report bias. Finally, although we are not aware of research that has examined IE effects in qualitative reports, consumer research that rely on mixed methods might consider the potential impact of the first interview on qualitative reports, in addition to those based on quantitative measures.

Appendix A

SAS PROC NLMIXED syntax for estimation of the two-part location scale mixed-effects (MELS) model applied to daily time spent watching TV. Covariates include within-person effects of the day of the week (Sunday is the reference day), an individual's age and biological sex, daily and daily-average measures of positive and negative affect, an overall initial elevation (IE) effect and moderating effects of IE on the effects of daily and daily-average measures of positive and negative affect.

```
proc nlmixed data=data gconv=0 noad qpoints=30;

* starting values;
parms
a0     3.8 a1a -0.39 a1b -0.71 a2 -0.76
a3a    0.38 a3b 2.12
a4a    -0.27 a4b -0.16 a4c -0.21 a4d -0.07 a4e -0.07 a4f -0.35
a5     -0.39 a6 0.02
b0     -0.46
b1a    0.16 b1b -0.02 b2 0.03
b3a    0.02 b3b 0.08
b4a    -0.14 b4b -0.18 b4c -0.25 b4d -0.25 b4e -0.21 b4f -0.09
b5     -0.01 b6 0.01
alpa0 1.56 alpb0 -0.90
xrhoba 0.83 xrhova -0.74 xrhovb -0.71
tau0a -1.17 tau0b -1.51
sdv    0.80
p1     0.06 BFmixture 1.0
a7a    -0.36 a7b -0.13 a8a 0.54 a8b 0.30
b7a    -0.06 b7b 0.00 b8a 0.05 b8b 0.04;

pi = arcos(-1);

u = tv_u; * daily binary indicator of TV watched versus not watched;
m = tv_m; * daily continuous measure of time spent when TV was
watched;

* exponential functions express the standard deviations of the random
intercepts of each model part;
sda = sqrt(exp(alpa0)); sdb = sqrt(exp(alpb0));

* estimation uses Fisher's transformation of the correlation
coefficients and the correlations that define the random-effects
covariance matrix are expressed by the inverse transformation;
rhoba = tanh(Xrhoba); rhova = tanh(Xrhova); rhovb = tanh(Xrhovb);

* set the bounds of the proportion of the first response distribution
to be between 0 and 1;
bounds 0 <= p1 <=1;
* define the second proportion to be 1 – p1;
p2 = 1 - p1;
* set the bounds of the standard deviation of the random scale to be
at least equal to 0;
bounds sdv >=0;

* logistic model part for u;
ueta = a0 + a1a*mean_RA2DNEGAV + a1b*RA2DNEGAV_pc +
a2*IE +
a3a*mean_RA2DNEGAV*IE + a3b*RA2DNEGAV_pc*IE +
a7a*mean_RA2DPOSAV + a7b*RA2DPOSAV_pc +
a8a*mean_RA2DPOSAV*IE + a8b*RA2DPOSAV_pc*IE +
a4a*mon_pc + a4b*tue_pc + a4c*wed_pc + a4d*thu_pc + a4e*fri_pc +
a4f*sat_pc +
a5*female + a6*agec +
ai*sda;
```

```
expeta = exp(ueta);  p=expeta/(1+expeta);

* define the loglikelihood for the logistic model part;
LL1 = log((1-p)**(1-u)) + log(p**(u));

* generalized linear model part for m;
if tvhours > 0 then do;
* exponential function for class 1 within-subject variance with random
scale vi;
s2e1 = exp(tau0a + vi*sdv);
* class 1 predicted mean response;
mu1 = b0 + b1a*mean_RA2DNEGAV + b1b*RA2DNEGAV_pc +
b2*IE +
b3a*mean_RA2DNEGAV*IE + b3b*RA2DNEGAV_pc*IE +
b7a*mean_RA2DPOSAV + b7b*RA2DPOSAV_pc +
b8a*mean_RA2DPOSAV*IE + b8b*RA2DPOSAV_pc*IE +
b4a*mon_pc + b4b*tue_pc + b4c*wed_pc + b4d*thu_pc + b4e*fri_pc +
b4f*sat_pc +
b5*female + b6*agec +
bi*sdb;
* lognormal likelihood for class 1;
L1 = (1/(sqrt(2*pi*s2e1)*m))*exp(-(((log(m)-mu1)**2)/(2*s2e1)));

* exponential function for class 2 within-subject variance with random
scale vi;
s2e2 = exp(tau0b + vi*sdv);
* class 2 predicted mean response;
mu2 = b0 + b1a*mean_RA2DNEGAV + b1b*RA2DNEGAV_pc +
b2*IE +
b3a*mean_RA2DNEGAV*IE + b3b*RA2DNEGAV_pc*IE +
BFmixture +
b7a*mean_RA2DPOSAV + b7b*RA2DPOSAV_pc +
b8a*mean_RA2DPOSAV*IE + b8b*RA2DPOSAV_pc*IE +
b4a*mon_pc + b4b*tue_pc + b4c*wed_pc + b4d*thu_pc + b4e*fri_pc +
b4f*sat_pc +
b5*female + b6*agec +
bi*sdb;
* lognormal likelihood for class 2;
L2=(1/(sqrt(2*pi*s2e2)*m))*exp(-(((log(m)-mu2)**2)/(2*s2e2)));

* loglikelihood is a weighted sum of L1 and L2;
LL2 = log(L1*p1+L2*p2);
end;

if      u=0 then Loglik=LL1;
else if u=1 then Loglik=LL1+LL2;
model tvhours ~ general(Loglik);
* correlation matrix of the random effects;
random ai bi vi ~ normal([0,0,0],[1,
                                  rhoba,1,
                                  rhova,rhovb,1]) subject=MRID;
estimate "day 1 within subject negative affect effect on likeliness to
watch TV" a1b + a3b + a2;
estimate "day 1 daily-average positive affect effect on likeliness to
watch TV" a7a + a8a + a2;
```

Declarations

Conflict of interest The corresponding author states that there are no conflicts of interest.

References

Beuk, F., A.J. Malter, J. Spanjol, and J. Cocco. 2014. Financial incentives and salesperson time orientation in new product launch: A longitudinal study. *Journal of Product Innovation Management* 31 (4): 647–663.

Blozis, S.A., R. Villarreal, S. Thota, and N. Imparato. 2019. Using a two-part mixed-effects model for understanding daily, individual-level media behavior. *Journal of Marketing Analytics* 7 (4): 234–250.

Blozis, S.A., M. McTernan, J. Harring, and Q. Zheng. 2020. Two-part mixed-effects location scale models. *Behavior Research Methods* 52: 1836–1847.

Bowman, N.D., and R. Tamborini. 2015. In the mood to game: Selective exposure and mood management processes in computer game play. *New Media & Society* 17 (3): 375–393.

Chaney, D., and D. Martin. 2017. The role of shared values in understanding loyalty over time: A longitudinal study on music festivals. *Journal of Travel Research* 56 (4): 507–520.

Chen, C.Y., and M.T. Pham. 2019. Affect regulation and consumer behavior. *Consumer Psychology Review* 2 (1): 114–144.

Cohen, J.B., M.T. Pham, and E.B. Andrade. 2008. The nature and role of affect in consumer behavior. In *Handbook of consumer psychology,* ed. C.P. Haugtvedt, P.M. Herr, F.R. Kardes, 297–348. Taylor & Francis Group/Lawrence Erlbaum Associates.

Dhiman, N., and A. Kumar. 2023. What we know and don't know about consumer happiness: Three-decade review, synthesis, and research propositions. *Journal of Interactive Marketing* 58 (2–3): 115–135.

Eastman, J.K., R. Iyer, K.L. Eastman, S. Cordon-Wilson, and P. Modi. 2021. Reaching the price conscious consumer: The impact of personality, generational cohort and social media use. *Journal of Consumer Behavior* 20: 898–912.

Festinger, L., and D. Katz (Eds.). 1953. *Research Methods in the Behavioral Sciences*. Dryden Press.

Foxall, G.R. 2003. The behavior analysis of consumer choice: An introduction to the special issue. *Journal of Economic Psychology* 24 (5): 581–588.

Gibson, E.L. 2006. Emotional influences on food choice: sensory, physiological and psychological pathways. *Physiology & Behavior* 89 (1): 53–61.

Greenwood, D.N., and C.R. Long. 2009. Mood specific media use and emotion regulation: Patterns and individual differences. *Personality and Individual Differences* 46 (5–6): 616–621.

Hedeker, D., R.J. Mermelstein, and H. Demirtas. 2012. Modeling between-subject and within-subject variances in ecological momentary assessment data using mixed-effects location scale models. *Statistics in Medicine* 31 (27): 3328–3336.

Hoffner, C.A., and S. Lee. 2015. Mobile phone use, emotion regulation, and well-being. *Cyberpsychology, Behavior and Social Networking* 18 (7): 411–416.

Holbrook, M.B., and E.C. Hirschman. 1982. The experiential aspects of consumption: consumer fantasies, feelings, and fun. *Journal of Consumer Research* 9 (2): 132–140.

Kass, R.E., and L. Wasserman. 1995. A reference Bayesian test for nested hypotheses and its relationship to the Schwarz criterion. *Journal of the American Statistical Association* 90: 928–934.

Knobloch-Westerwick, S., and S. Alter. 2006. Mood adjustment to social situations through mass media use: how men ruminate and women dissipate angrymoods. *Human Communication Research* 32 (1): 58–73.

Katakam, B.S., R. Bhukya, R.S. Bellamkonda, and N. Samala. 2021. Longitudinal analysis versus cross-sectional analysis in assessing the factors influencing shoppers' impulse purchase behaviour—Do the store ambience and salesperson interactions really matter? *Journal of Retailing and Consumer Services* 61: 102586.

Knowles, E.S., M.C. Coker, R.A. Scott, D.A. Cook, and J.W. Neville. 1996. Measurement-induced improvement in anxiety: Mean shifts with repeated assessment. *Journal of Personality and Social Psychology* 71 (2): 352–363.

Kumar, V., J.B. Choi, and M. Greene. 2017. Synergistic effects of social media and traditional marketing on brand sales: Capturing the time-varying effects. *Journal of the Academy of Marketing Science* 45 (2): 268–288.

Laird, N.M., and J.H. Ware. 1982. Random-effects models for longitudinal data. *Biometrics* 38 (4): 963–974.

Nabi, R.L., and M. Krcmar. 2004. Conceptualizing media enjoyment as attitude: Implications for mass media effects research. *Communication Theory* 14 (4): 288–310.

Nabi, R.L., D. Pérez Torres, and A. Prestin. 2017. Guilty pleasure no more: The relative importance of media use for coping with stress. *Journal of Media Psychology* 29: 126–136.

Nagin, D.S. 1999. Analyzing developmental trajectories: A semiparametric, group-based approach. *Psychological Methods* 4 (2): 139–157.

Norton, T.A., S.L. Parker, H. Zacher, and N.M. Ashkanasy. 2015. Employee green behavior: A theoretical framework, multilevel review, and future research agenda. *Organization & Environment* 28: 103–125.

Payan, J.M., C. Obadia, J. Reardon, and I. Vida. 2010. Survival and dissolution of exporter relationships with importers: A longitudinal analysis. *Industrial Marketing Management* 39 (7): 1198–1206.

Raudenbush, S.W., and A.S. Bryk. 2002. *Hierarchical linear models: Applications and data analysis methods*, 2nd ed. Thousand Oaks: Sage Publications.

Rich, M., D.S. Bickham, and L.A. Shrier. 2015. Measuring youth media exposure: A multimodal method for investigating the influence of media on digital natives. *American Behavioral Scientist* 59 (14): 1236–1754.

Robins, L.N. 1985. Epidemiology: Reflections on testing the validity of psychiatric interviews. *Archives of General Psychiatry* 42 (9): 918–924.

Ryff, C.D. and Almeida, D. (2012–2014). Midlife in the United States (MIDUS Refresher): Daily diary project. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2018–06–06.

Schimmack, U., and S. L. Crites, Jr. 2005. The structure of affect. In *The handbook of attitudes,* ed. D. Albarracin, B.T. Johnson, M.P. Zanna, 397–435. Lawrence Erlbaum Associates Publishers.

Sharpe, J.P., and D.G. Gilbert. 1998. Effects of repeated administration of the beck depression inventory and other measures of

negative mood states. *Personality and Individual Differences* 24 (4): 457–463.
Shrout, P.E., G. Stadler, S.P. Lane, M.J. McClure, G.L. Jackson, F.D. Clavél, M. Iida, M.E.J. Gleason, J.H. Xu, and N. Bolger. 2018. Initial elevation bias in subjective reports. *Proceedings of the National Academy of Sciences of the United States of America* 115 (1): E15–E23.
Thürridl, C., B. Kamleitner, R. Ruzeviciute, S. Süssenbach, and S. Dickert. 2020. From happy consumption to possessive bonds: When positive affect increases psychological ownership for brands. *Journal of Business Research* 107: 89–103.
Windle, C. 1954. Test-retest effect on personality questionnaires. *Educational and Psychological Measurement* 14 (4): 617–633.
Wolfers, L.N., and F.M. Schneider. 2021. Using media for coping: A scoping review. *Communication Research* 48 (8): 1210–1234.

Shelley A. Blozis is Professor of Psychology in the Department of Psychology at the University of California, Davis. She has a Ph.D. in Psychology from the University of Minnesota, Twin Cities.

Ricardo Villarreal is Professor of Marketing in the School of Management at the University of San Francisco. He has a Ph.D. in Advertising from The University of Texas at Austin.