Research Article

Psychological Science
2019, Vol. 30(3) 405–414

DOI: 10.1177/0956797618818476
www.psychologicalscience.org/PS

Do People Know What They're Like in the Moment?

Jessie Sun and Simine Vazire
Department of Psychology, University of California, Davis

Abstract

Knowing yourself requires knowing not only what you are like in general (trait self-knowledge) but also how your personality fluctuates from moment to moment (state self-knowledge). We examined this latter form of self-knowledge. Participants (248 people; 2,938 observations) wore the Electronically Activated Recorder (EAR), an unobtrusive audio recorder, and completed experience-sampling self-reports of their personality states four times each day for 1 week. We estimated state self-knowledge by comparing self-reported personality states with consensual observer ratings of personality states coded from the EAR files, which formed the criterion for what participants were "actually" like in the moment. People had self-insight into their momentary extraversion, conscientiousness, and likely neuroticism, suggesting that people can accurately detect fluctuations in some aspects of their personality. However, the evidence for self-insight was weaker for agreeableness. This apparent self-ignorance may be partly responsible for interpersonal problems and for blind spots in trait self-knowledge.

Keywords

self-knowledge, personality states, social behavior, experience-sampling method, naturalistic observation, open data, open materials

Received 7/2/18; Revision accepted 10/8/18

Imagine that one frazzled moment before your morning coffee, you contemptuously order your romantic partner to get out of your face when he or she attempts to talk to you while you are rushing to leave for work. What happens next might depend on whether you are aware of how you just acted. If you realize in the moment that you just acted like a jerk, you can apologize and try to make amends right away. Alternatively, if you continue with your day, oblivious to how you acted and failing to do anything about it, this might cause resentment to brew over time. Here, we tackle a critical piece of the self-knowledge puzzle: Do people know how they are acting in the moment?

Self-knowledge is defined as the degree to which a person's self-views reflect what they are really like (Vazire & Carlson, in press; T. D. Wilson, 2009). Most previous self-knowledge research has focused on how well people know what they are typically like (*trait self-knowledge*), showing that there are bright spots and blind spots in people's trait self-knowledge (Vazire, 2010; Vazire & Mehl, 2008). A related form of self-knowledge—and the focus of this article—is whether people know what they are like from one moment to the next (*state self-knowledge*). In other words, do people have insight into their personality states (i.e., thoughts, feelings, and behaviors over shorter periods of time; Fleeson, 2001)?

Although state self-knowledge has been a largely neglected phenomenon (cf. Gosling, John, Craik, & Robins, 1998), there are several reasons why it is important to examine whether people know what they are like from moment to moment. First, state self-knowledge may help people understand what they are like in general (trait self-knowledge). If the first step to understanding your general pattern of behaviors is the ability to recognize the instances that form the pattern, trait-level blind spots (e.g., not knowing that you are a jerk) may arise in part from a lack of awareness of one's behavior in the moment (e.g., not realizing when you

Corresponding Author:
Jessie Sun, University of California, Davis, Department of Psychology, 200 East Quad, Davis, CA 95616
E-mail: jesun@ucdavis.edu

are being rude). Thus, identifying blind spots in state self-knowledge can help us understand how to improve both state and trait self-knowledge.

Second, state self-knowledge may pave the way for more comprehensive and contextualized forms of self-knowledge. Knowing only a person's traits has been described as the "psychology of the stranger" (McAdams, 1995, p. 380), with deeper knowledge of a person coming from an understanding of the dynamic, contextual influences on fluctuations in the person's personality states (McAdams, 1995; Vazire & Carlson, in press; R. E. Wilson & Vazire, 2015). When the person you seek to know is yourself, this implies that trait self-knowledge is a relatively rudimentary form of self-knowledge. To really know yourself, it is also important to understand the idiosyncratic influences (e.g., goals, social roles, biological states) that cause you to think, feel, and act differently from one moment to the next. State self-knowledge—the ability to accurately detect these personality fluctuations in the first place—is likely an important part of the process of forming and revising your ideas about what these influences might be.

State self-knowledge might also have immediate practical consequences. Although some evidence suggests that trait self-knowledge may predict better social relationships (Tenney, Vazire, & Mehl, 2013), being aware of and able to do something about your disagreeableness *in the moment* might be more useful than knowing that you are *generally* a disagreeable person.

Finally, assessing the accuracy of state self-reports can help researchers decide when to trust these measures. Many studies now use the experience-sampling method (ESM), which asks people to report on their momentary experiences and behaviors (Mehl & Conner, 2012). Because the ESM minimizes retrospective reporting biases, it is often touted as the gold standard for assessing in-the-moment experiences (Schwarz, Kahneman, & Xu, 2009). However, as with any self-report measure, ESM reports are valid only to the extent that people have self-insight into their momentary states. Thus, the validity of ESM reports for any given behavior or experience should be tested, rather than assumed.

The aim of the present research was to examine whether people (specifically, college students in North America) have self-knowledge about their personality states in their everyday lives. The biggest challenge in self-knowledge research is the criterion problem—how do we establish what people are *actually* like (i.e., the "ground truth")? To assess whether people have accurate self-views, we must compare self-reports with an independent measure (i.e., an accuracy criterion). The ideal criterion would repeatedly capture a broad range of people's behaviors in their natural environments using consensual observer ratings. We used the only method we know of that can do this in an unobtrusive way: the Electronically Activated Recorder (EAR; Mehl, 2017), a wearable device that repeatedly captures short audio recordings of people's observable behaviors and environments. These recordings are coded by multiple observers, providing a reliable outside perspective on people's real-world behaviors and allowing us to compare repeated self-perceptions of personality states with observer ratings in the same moments.

Method

We used data from the first wave of the longitudinal Personality and Interpersonal Roles Study (PAIRS). Other manuscripts have used the ESM personality-state variables (Beck & Jackson, 2018; Breil et al., in press; Finnigan & Vazire, 2018; R. E. Wilson, Harris, & Vazire, 2015; R. E. Wilson, Thompson, & Vazire, 2016) and other variables from this data set (Colman, Vineyard, & Letzring, 2018; Edwards & Holtzman, 2017; Solomon & Vazire, 2016; Weidman et al., in press), but this is the first article that examines within-person associations between self-reported and EAR-coded behavior.

Participants

The study involved 434 students at Washington University in St. Louis, who were recruited in 2012 and 2013 via flyers and classroom announcements across the campus. Participants were paid $20 for the initial laboratory-based assessment, were entered into a lottery with the opportunity to win $100 for completing ESM surveys (with a 1 in 10 chance of winning if all ESM surveys were completed), earned an additional $20 for wearing the EAR, and received a "time capsule" that contained feedback on how their personality had changed across the seven waves of the study.

The sample size of the original study was determined by the stopping rule of ending data collection when we reached the end of a semester and had recruited at least 400 participants. The number of ESM observations per participant was determined by our subjective impression of how many repeated measures we could obtain from each participant without compromising the quality of the data or the participants' goodwill. After exclusions (described in the Data Exclusions section), the final subset of 248 participants (173 women, 74 men, 1 gender not reported) used in the current analyses ranged in age from 18 to 29 years (M = 19.17 years, SD = 1.8) and identified as Caucasian (n = 141), Asian (n = 57), Black (n = 24), American Indian or Alaska Native (n = 1), or other or multiple races (n = 18) or did not disclose their race (n = 7).

Procedure

Here, we describe the measures and procedures relevant to the current article. Codebooks for all measures in the larger study are available at https://osf.io/akbfj.

ESM measures. Four times per day (at 12 p.m., 3 p.m., 6 p.m., and 9 p.m.) for 15 days, participants received a text message notification and were e-mailed a link to a survey that contained ESM measures of Big Five personality states in the target hour (11 a.m.–12 p.m., 2 p.m.–3 p.m., 5 p.m.–6 p.m., and 8 p.m.–9 p.m.). Using the nine items that we adapted from the Big Five Inventory (John, Naumann, & Soto, 2008), participants reported their state extraversion ("quiet" [reverse-scored]; "outgoing, sociable"), agreeableness ("considerate, kind"; "rude" [reverse-scored]), conscientiousness ("reliable"; "lazy" [reverse-scored]), and neuroticism ("worried"; "relaxed" [reverse-scored]; "depressed, blue") in each target hour (e.g., "From 11am–noon, how [outgoing, sociable] were you?"). Responses were made on a 5-point scale (1 = *not at all*, 5 = *very*). Participants completed the agreeableness items only if they reported that they were around other people during the target hour. We did not include ESM measures of Big Five openness in this wave of the study because we previously believed that the openness items on the Big Five Inventory would not translate well to a state measure (an issue we changed our mind about in later waves of data collection).

EAR protocol. During the first week (6–8 days) of the ESM protocol, 311 participants wore the EAR, implemented through the iEAR app using an iPod Touch device. The EAR component of the study was optional, was offered only during nonsummer months of the study, and was not an option when all of the researchers' iPod Touches were in use by other participants. The EAR was programmed to record 30-s audio snippets of participants' ambient sounds every 9.5 min from 7 a.m. to 2 a.m. Participants were encouraged to wear the EAR as much as possible and to wear it clipped to a waistband or the outside of their pockets (i.e., not inside a bag or pocket). Although participants had no way to tell when the device was recording, they were told that they could decide to not wear the EAR at any time for any reason.

After 3 to 4 days, participants returned to the lab to upload their data (because of device memory limitations) and then continued wearing the device before returning it after another 3 to 4 days. After returning the device, participants received a compact disc with their recordings so that they could listen to and erase any files they did not want the researchers to hear. Only a few participants (n = 15) chose to erase a total of 99 files. After these files were deleted, along with files from 6 participants who withdrew and 1 participant who had only silent recordings (suggesting that the microphone malfunctioned), 152,592 usable recordings from 304 participants remained.

EAR codings. From September 2013 until February 2018, research assistants from Washington University in St. Louis (n = 8) and the University of California, Davis (n = 100), listened to the audio files recorded during the same hours as the ESM reports (11 a.m.–12 p.m., 2 p.m.–3 p.m., 5 p.m.–6 p.m., and 8 p.m.–9 p.m.). For each of their assigned participants, coders listened to the six or seven 30-s files from each ESM-matched hour (3–3.5 min total); rated participants' levels of state extraversion, agreeableness, conscientiousness, and neuroticism during that hour (as part of a larger survey); and then moved on to the next ESM-matched hour for that participant.

Coders' ratings were made using the same items and 5-point scale (1 = *not at all*, 5 = *very*) that participants used in their ESM self-reports, with a few minor differences: (a) The items were worded in terms of how the participant seemed (e.g., "In this hour, the participant seemed [quiet]"), (b) coders completed the agreeableness items only if they believed that the participant was interacting with other people (not just around others) during the target hour, and (c) coders had the option to select "No way to tell" (rather than a number on the scale).

Because research assistants joined and left the lab at different times, each participant was coded by a different set of coders. Initially, we aimed to have each participant coded by three coders. However, as the interrater reliabilities based on three coders were low, we decided to add three more coders, so that each participant was coded by at least six coders. Between the two sets of codings, we made minor changes to the coding protocol (see the Supplemental Material available online), in hopes of increasing interrater reliability.

Transcripts. After seeing the results of the key quantitative analyses, we decided to supplement these analyses with qualitative data from transcripts of the EAR files. These transcripts were obtained through a separate coding task, in which participants' utterances were transcribed by a different research assistant from the one who provided their observer ratings. Transcribers were trained to recognize the participant's voice; to handle ambiguities such as repetitions, filler words, nonfluencies, and slang; and to use special characters to indicate when participants were singing or acting (see "Transcription Guide" at https://osf.io/kd8b3).

Data exclusions

ESM exclusions. In line with exclusion criteria applied in previous articles that used the PAIRS ESM data (Finnigan & Vazire, 2018; R. E. Wilson et al., 2015; R. E. Wilson et al., 2016), we excluded ESM reports (a) if they were

completed more than 3 hr after the notification was sent, (b) if participants completed fewer than 75% of the items, (c) if participants used the same response option for at least 70% of the items, and (d) if participants indicated that they were sleeping during the target hour. We also excluded practice ESM surveys that were completed during each participant's initial laboratory session. After these exclusions, 10,949 reports from 406 participants remained.

EAR exclusions. Coders rated participants' personality states only in hours that contained sufficient acoustic information. We kept only the hours that at least three coders rated as being informative (for details, see the Supplemental Material). On the basis of these criteria, 807 out of 5,222 hr (15.45%) were uninformative (and excluded from further analyses).

Minimum number of matched observations. Of the remaining 4,415 EAR observations, 3,050 observations had a corresponding ESM report (from 289 participants). We excluded 112 observations from 41 participants who had fewer than five matched observations (i.e., time points that contained both ESM and EAR data), resulting in 2,938 observations from 248 participants. Because there were some missing data (especially for agreeableness, as responses to these items were conditional on either being around other people [ESM reports] or interacting with other people [EAR observations]), we applied the five matched observation inclusion criteria for each personality state. This ensured that each analysis included only participants who had at least five time points containing both ESM and EAR data for the focal personality state. Beyond this minimum, we retained time points that had either ESM or EAR data to allow Mplus to use all available information. This left final sample sizes of 2,938 observations for the extraversion, conscientiousness, and neuroticism analyses and 2,519 observations for the agreeableness analyses.

Data analyses

Key analyses. The data had a multilevel structure, with observations (Level 1) nested within participants (Level 2). We used B. O. Muthén and Asparouhov's (2009) general multilevel-structural-equation-modeling (MSEM) framework, implemented in Mplus (Version 8.1; L. K. Muthén & Muthén, 2017), to model the within-person agreement between self-reported and EAR-coded personality states. MSEM enables the modeling of both random effects (to allow for individual differences in state self-knowledge) and latent variables (so that variables are not assumed to be measured without error). This means that differences in effect sizes for the four personality states will not be due to differences in measurement reliability. MSEM also allows for Level 1 and Level 2 effects to be simultaneously estimated, so that within-person effects are not conflated with between-person effects. Thus, we estimated latent variables and effects at both the within-person level and the between-person level but focus on the within-person effects in this article (for between-person correlations, see Table S2 in the Supplemental Material). We ran separate models for each of the four personality states (see Fig. S1 in the Supplemental Material).

Measurement models. Each personality state was modeled as a latent variable. For the ESM latent variables, the indicators were the two or three items for the personality state. For the two-item measures (for agreeableness, conscientiousness, and extraversion), in order for the model to be locally identified, we fixed both item loadings to 1 (and allowed all variances to be freely estimated). For the three-item neuroticism measure, we fixed the first item loading to 1 and allowed the other two factor loadings (and all variances) to be freely estimated.

For the EAR latent variables, we used coders as indicators. Some hours were coded by more than six coders, but to reduce model complexity, we included data from up to only six coders for the latent variables (for details, see the Supplemental Material). To create the indicators, we computed scale scores (i.e., the average of the two or three items for each scale) for each of the six coders. Then we used these six scale scores as indicators. Thus, every latent variable had six indicators (with each indicator representing a scale score from a given coder for a given participant). For a given participant (e.g., Participant 1), all ratings from Coder 1 were from the same coder (e.g., Research Assistant 1). However, for a different participant (e.g., Participant 2), Coder 1 could have been a different research assistant (e.g., Research Assistant 2). To model the interchangeability of coders, we fixed all loadings for the six indicators to 1, constrained the six residual variances to be equal, and allowed the within- and between-person variances of the latent EAR variables to be freely estimated.

We conducted multilevel confirmatory factor analyses (Geldhof, Preacher, & Zyphur, 2014; Shrout & Lane, 2012) on these measurement models to obtain level-specific omega (ω) reliability estimates. The within-person ω (ω_{WP}; see Table 1) estimates the reliability of change, which is the proportion of within-person variability due to meaningful changes in the personality state from one moment to the next, as assessed by two or three items (for the ESM latent variable) or six coders (for the EAR latent variable). The between-person ω (ω_{BP}; see Table S1 in the Supplemental Material) estimates the proportion of between-person variability due to true between-person differences on participants' average personality states.

Structural models. For each personality state, for the within-person models, we regressed the EAR latent variable

Table 1. Descriptive Statistics

	Self-views (ESM reports)					Observed behavior (EAR codings)				
Personality state	*M*	SD_{WP}	SD_{BP}	1 – ICC(1)	ω_{WP}	*M*	SD_{WP}	SD_{BP}	1 – ICC(1)	ω_{WP}
Extraversion	2.88	1.03	0.37	.89	.82	2.77	0.95	0.38	.86	.93
Agreeableness	3.96	0.26	0.29	.45	.26	4.17	0.24	0.14	.74	.62
Conscientiousness	3.62	0.52	0.46	.57	.47	3.72	0.39	0.20	.79	.76
Neuroticism	2.16	0.71	0.57	.61	.72	1.82	0.28	0.09	.90	.73

Note: All ratings were made on Likert-type scales ranging from 1 to 5. Values for 1 – intraclass correlation coefficient(1), or ICC(1), reflect the proportion of total variability attributable to within-person variability; ω_{WP} reflects the within-person estimate of the reliability of change. Variance estimates for the average within-person standard deviation (SD_{WP}), between-person standard deviation (SD_{BP}), and 1 – ICC(1) are based on the measurement models (i.e., latent variables). Means were obtained by computing the aggregate mean (from observed scores) for each participant and then computing the mean of these means (such that all participants were weighted equally). ESM = experience-sampling method (i.e., self-reports); EAR = Electronically Activated Recorder (i.e., observer reports).

onto the ESM latent variable, with random slopes and random intercepts for each participant. In other words, this model allowed each participant to have a different mean level on each personality state and a different association between self-reported and EAR-coded states. We also modeled the between-person path from the ESM latent variable to the EAR latent variable, although this is not the focus of this article.

Estimation and inference criteria. Because of the computational demands of these models, we used the Bayes estimator in Mplus (B. O. Muthén & Asparouhov, 2012), with the default set of diffuse (i.e., noninformative) priors. We used the 95% credibility interval (CI) around the standardized effects (βs) as inference criteria for the range of plausible population values of the effect sizes.

Qualitative analyses. To provide a sense of what was happening when participants and observers disagreed about participants' personality states, we report the transcripts that correspond to the largest discrepancies between ESM self-reports and EAR observer ratings for those Big Five states that showed the largest self–other disagreement. We conducted this supplemental analysis for each Big Five domain for which we judged self–observer agreement to be low.

To do this, we standardized the ESM self-reports and EAR observer ratings within each person across the same time points included in the key analyses (but using observed variables instead of the latent variables that were used in the key analyses). We then matched up the ESM and EAR data with the hour-level transcripts (i.e., all of the decipherable words across the six or seven 30-s recordings in each hour) from the 121 participants who gave permission to publish their transcripts, retaining only the hours that contained transcripts (i.e., excluding hours in which participants did not speak or had no decipherable speech). Next, we subtracted the EAR observer ratings from the ESM self-reports and selected the 50 target hours associated with the largest discrepancies between ESM self-reports and EAR observer ratings (25 in each direction, for each personality state). The transcripts from these target hours are shown in Tables S3 to S6 in the Supplemental Material.

Results

Descriptive statistics

Descriptive statistics for all variables are reported in Table 1. The intraclass correlations, ICC(1)s, show that there was substantial (≥ 45%) within-person variability for each of the personality states, as captured by both the ESM self-reports and EAR codings. Note that the within-person reliabilities for state agreeableness (both EAR and ESM) and ESM conscientiousness were relatively low but that MSEM corrects for attenuation due to measurement error.

State self-knowledge analyses

Do people have self-knowledge of their personality states in everyday life? To test this, we examined the correspondence between self-views (ESM reports) and observed behavior (EAR codings) using the models described above. Positive slopes reflect agreement between self-reports and observed behavior, which we interpret as evidence of self-knowledge, whereas weak or flat slopes may or may not reflect lack of self-knowledge (for more details regarding interpretation, see the Discussion section). Figure 1 shows the individual slopes and the average slope (bold line) for each Big Five domain (for all unstandardized and standardized estimates, see Table S1).

The average slopes were positive and nonzero for all four domains. However, the effects were quite a bit

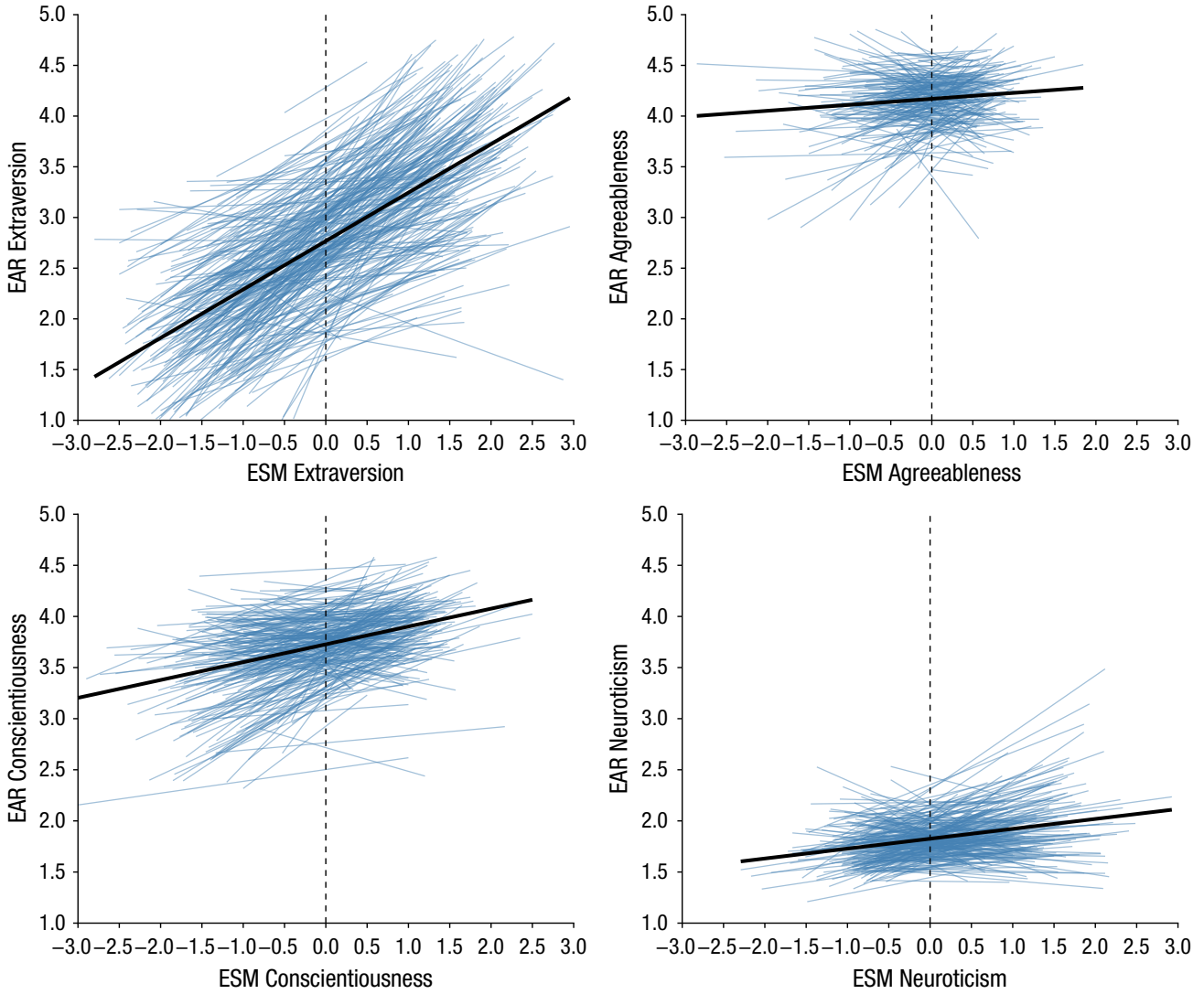

Fig. 1. Spaghetti plots depicting the within-person associations between self-views (*x*-axes) and observed behavior (*y*-axes), separately for each of the four personality states. Self-views were obtained using the experience-sampling method (ESM), and observed behavior was coded from Electronically Activated Recorder (EAR) recordings. Each colored line represents the slope for a different participant, and the black line shows the average within-person effect. Strong positive slopes reflect self–observer agreement, which we interpret as evidence of self-knowledge. Weak or flat slopes may or may not reflect lack of self-knowledge. The *x*-axes show deviations from each person's mean self-reported personality state, whereas the *y*-axes show the uncentered range of EAR codings from 1 to 5.

larger for extraversion (β = 0.63, 95% CI = [0.60, 0.66]) and conscientiousness (β = 0.47, 95% CI = [0.40, 0.55]) than for neuroticism (β = 0.27, 95% CI = [0.21, 0.32]) and agreeableness (β = 0.20, 95% CI = [0.11, 0.32]). When participants rated themselves as being more extraverted or conscientious than they usually were, the EAR coders also rated them as more extraverted or conscientious than their typical levels. This suggests that participants had self-insight into their momentary fluctuations in extraversion and conscientiousness.

In contrast, when participants rated themselves as more neurotic or agreeable than usual, this only weakly corresponded to how the EAR coders rated them. In addition, the 95% CIs for self–observer agreement on neuroticism and agreeableness, while excluding zero, did not overlap with those for extraversion and conscientiousness. In short, agreement was substantially weaker for neuroticism and agreeableness than for extraversion and conscientiousness, even though the models accounted for differences in measurement reliability across constructs. These results are more complicated to interpret than those for extraversion and conscientiousness and may or may not imply that participants lacked self-insight into how neurotic and agreeable they were

in the moment. We will return to this challenge of interpretation in the Discussion section.

In the meantime, to allow readers to gain a better sense of what was happening when participants and observers disagreed about participants' momentary agreeableness and neuroticism, we provide the transcripts from the hours with the largest self–observer discrepancies in these states in Tables S3 through S6 in the Supplemental Material (for those participants who gave consent to share their EAR recordings). In the Discussion section, we share a few of our own observations from reading these transcripts, but we encourage readers to explore the transcripts in Tables S3 through S6, along with the full set of shareable transcripts and their corresponding self–observer discrepancy scores (posted on our Open Science Framework page, osf.io/kd8b3, the password for which is available on request).

Discussion

Our goal was to test whether people know what they are like in the moment. We found high levels of self–observer agreement for state extraversion and conscientiousness but lower levels of agreement for neuroticism and agreeableness. These results can be interpreted as accuracy estimates only if we assume that observers can detect true fluctuations in personality states through brief audio recordings of participants' everyday behaviors and environments. We believe that this assumption holds more strongly for momentary extraversion, conscientiousness, and agreeableness than for neuroticism. Thus, we interpret our results as showing that people have self-insight into their momentary extraversion and conscientiousness, that momentary neuroticism is difficult (but not impossible) for observers to judge, and that people have poor self-knowledge of their momentary agreeableness.

The findings for extraversion are consistent with a large body of literature demonstrating high self–observer agreement on trait extraversion across a wide range of conditions (for a review, see Vazire & Solomon, 2015). These findings provide new evidence that self-perceptions of state extraversion are accurate—that is, people know when they are being more or less extraverted than usual. Likewise, the substantial self–observer agreement for conscientiousness suggests that people are willing and able to report when they are acting lazy versus acting reliable.

Because the EAR is not a perfect criterion, however, the lower self–observer agreement for state neuroticism and agreeableness could suggest (a) that EAR coders could not accurately detect these states or (b) that people's self-reports are inaccurate. Although both explanations are probably partially correct, we suspect that the first explanation largely accounts for low self–observer agreement for neuroticism. Previous studies have suggested that neuroticism is quite hard to observe (John & Robins, 1993) and that people are the best judges of their own trait neuroticism (Vazire, 2010). We suspect that it was difficult for EAR coders to detect states such as being worried on the basis of only audible behaviors. Thus, for state neuroticism, the weaker self–observer agreement may not imply low self-insight.

To explore whether this interpretation is consistent with the transcript data, we looked at the transcripts from the time points with the greatest discrepancies between self-reports and observers' ratings of state neuroticism (see Tables S5 and S6). As the content of these transcripts did not seem particularly informative to us, we explored another potentially relevant indicator—quantity of speech. After looking at the ESM–EAR discrepancies across all time points (including those with no speech), we observed that many of the time points in which self-reports of neuroticism were much higher than observer reports contained no speech. This suggests that people sometimes feel quite worried or depressed without expressing it verbally, which is consistent with our interpretation that state neuroticism is difficult to pick up from acoustic information alone.

However, we believe that it is plausible that people have less self-insight into their momentary agreeableness. Kindness and rudeness (the agreeableness states measured here) are defined more by behaviors than by thoughts and feelings. Thus, fluctuations in these states should be observable in naturalistic interactions with friends, roommates, and classmates, which the EAR is optimized to capture (Mehl, 2017). Indeed, our findings for extraversion show that EAR coders can detect interpersonal behaviors. We therefore believe that the weak self–observer agreement for agreeableness casts doubt on people's self-insight into how agreeable or disagreeable they are in the moment. This is consistent with the only other study we know of that examined people's awareness of their agreeableness-related behavior (during one laboratory-based group task; Gosling et al., 1998).

Tables S3 and S4, which report transcripts for the time points with the largest self–other discrepancies for agreeableness, may help shed light on the plausibility of our interpretation. For example, we agreed with the EAR coders that the participant who said "her twin brother did not have her in his wedding, which is such bullshit" was acting disagreeable (contrary to her self-rating) and that the participant who said, "Trust me, breaking up helps. And you have a good support system here.... You can come into me and Mel's room, just have a glass of wine," was acting quite agreeable (again, contrary to her self-rating). Of course, it is easy

to cherry-pick examples that fit our interpretation (and readers may not agree that even these cherry-picked examples support our interpretation), so we encourage readers to read the transcripts and come to their own conclusions.

Our results should be interpreted with the following limitations in mind. First, the observers had only 3 to 3.5 min of recordings spread across each hour they rated and had access to only acoustic information. Second, personality states comprise more than just observable behavior and more than just the content captured by the two to three items we used per domain. Third, there was less within-person variability in both self-reports and EAR codings of agreeableness states relative to the other personality states. Thus, it was likely more difficult for participants to detect the relatively narrow fluctuations in their own agreeableness states compared with, for example, the larger fluctuations in their own extraversion states.

However, given the challenges of studying self-knowledge, we believe that our methodology stands out in several ways: (a) high realism (we measured behavior across many situations in people's everyday lives), (b) moderate to high consensus on what participants were like from one moment to the next (we had each observation coded by six coders), and (c) high precision of estimates (we had large numbers of people and observations). Thus, although these results should not be the final word about state self-knowledge, they provide a strong test of college students' self-knowledge of what they are like during everyday moments.

Self-knowledge is central to our lives (and to the way that social scientists study our lives, often relying on self-reports). Our findings show that we can probably trust what people say about their momentary levels of extraversion, conscientiousness, and likely, neuroticism. However, our findings also call into question people's awareness of when they are being considerate versus rude. This is consistent with theoretical propositions and empirical evidence that people are poor judges of their trait agreeableness (John & Robins, 1993; Paulhus & John, 1998; Vazire, 2010; Vazire & Mehl, 2008) and suggests that people may not know how agreeable they *generally* are in part because they lack awareness of how rude or considerate they are in everyday *moments*.

Being aware of one's behavior in the moment also has benefits beyond its implications for trait self-knowledge. Practically, a momentary lapse in kindness could have dire consequences that could be avoided if you quickly realize that you just acted like a jerk. Recognizing instances of behavior while they happen is also a precursor to a deeper form of self-insight that involves not only knowing that you can sometimes be contemptuous but also knowing when (and ultimately, why) that happens (e.g., being caffeine deprived and in a rush). If it is true that, as Calvin told Hobbes, "we don't devote nearly enough scientific research to finding a cure for jerks" (Watterson, 1992, p. 58; see also Sutton, 2007), perhaps a good place to start is with people's blind spots about their behavior in the moment.

Action Editor

Brent W. Roberts served as action editor for this article.

Author Contributions

S. Vazire conceived the study, acquired funding, and supervised data collection. J. Sun supervised Electronically Activated Recorder coding and curated and analyzed the data under the supervision of S. Vazire. Both authors contributed equally to drafting and editing the manuscript, and both authors approved the final manuscript for submission.

ORCID iD

Jessie Sun https://orcid.org/0000-0001-6764-0721

Acknowledgments

We are grateful to Julia Rohrer, Chris Hopwood, Rick Robins, Olivia Atherton, Rich Lucas, and Luke Smillie for comments on an earlier draft of this article; to Mijke Rhemtulla for advice on the measurement models; to the many research assistants who helped run the study and code the Electronically Activated Recorder recordings; and to Brittany Solomon, Kelci Harris, Kathryn Bollich, Robert Wilson, and Katie Finnigan for supervising data collection and coding.

Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

Funding

Data collection was supported by National Science Foundation Grant BCS-1125553 (to S. Vazire).

Supplemental Material

Additional supporting information can be found at http://journals.sagepub.com/doi/suppl/10.1177/0956797618818476

Open Practices

Design and analysis plans for this study were not preregistered. Codebooks for all measures in the larger study are available via the Open Science Framework (OSF) and can be accessed at osf.io/akbfj. Although ethical considerations prevent us from making the audio files and complete set of transcripts publicly available, the quantitative data, R scripts, and Mplus input and output files required to reproduce the analyses reported in this

article are available at osf.io/kd8b3. This OSF repository also contains a password-protected file that contains transcripts (for the time points included in the key analyses) from participants who consented to have their Electronically Activated Recorder recordings shared. Interested researchers can request the password from the first author. The complete Open Practices Disclosure for this article can be found at http://journals.sagepub.com/doi/suppl/10.1177/0956797618818476. This article has received the badges for Open Data and Open Materials. More information about the Open Practices badges can be found at http://www.psychologicalscience.org/publications/badges.

References

Beck, E. D., & Jackson, J. J. (2018). *Consistency and change in idiographic personality: A longitudinal ESM network study*. Retreived from PsyArXiv website: https://psyarxiv.com/pb92q/

Breil, S. M., Geukes, K., Wilson, R. E., Nestler, S., Vazire, S., & Back, M. D. (in press). Zooming into real-life extraversion—How personality and situation shape sociability in social interactions. *Collabra: Psychology*.

Colman, D. E., Vineyard, J., & Letzring, T. D. (2018). Exploring beyond simple demographic variables: Differences between traditional laboratory samples and crowdsourced online samples on the Big Five personality traits. *Personality and Individual Differences*, *133*, 41–46. doi:10.1016/j.paid.2017.06.023

Edwards, T., & Holtzman, N. S. (2017). A meta-analysis of correlations between depression and first person singular pronoun use. *Journal of Research in Personality*, *68*, 63–68. doi:10.1016/j.jrp.2017.02.005

Finnigan, K. M., & Vazire, S. (2018). The incremental validity of average state self-reports over global self-reports of personality. *Journal of Personality and Social Psychology*, *115*, 321–337. doi:10.1037/pspp0000136

Fleeson, W. (2001). Toward a structure- and process-integrated view of personality: Traits as density distributions of states. *Journal of Personality and Social Psychology*, *80*, 1011–1027. doi:10.1037//0022-3514.80.6.1011

Geldhof, G. J., Preacher, K. J., & Zyphur, M. J. (2014). Reliability estimation in a multilevel confirmatory factor analysis framework. *Psychological Methods*, *19*, 72–91. doi:10.1037/a0032138

Gosling, S. D., John, O. P., Craik, K. H., & Robins, R. W. (1998). Do people know how they behave? Self-reported act frequencies compared with on-line codings by observers. *Journal of Personality and Social Psychology*, *74*, 1337–1349. doi:10.1037/0022-3514.74.5.1337

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative Big Five trait taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (3rd ed., pp. 114–158). New York, NY: Guilford Press.

John, O. P., & Robins, R. W. (1993). Determinants of interjudge agreement on personality traits: The Big Five domains, observability, evaluativeness, and the unique perspective of the self. *Journal of Personality*, *61*, 521–551. doi:10.1111/j.1467-6494.1993.tb00781.x

McAdams, D. P. (1995). What do we know when we know a person? *Journal of Personality*, *63*, 365–396. doi:10.1111/j.1467-6494.1995.tb00500.x

Mehl, M. R. (2017). The Electronically Activated Recorder (EAR): A method for the naturalistic observation of daily social behavior. *Current Directions in Psychological Science*, *26*, 184–190. doi:10.1177/0963721416680611

Mehl, M. R., & Conner, T. S. (2012). *Handbook of research methods for studying daily life*. New York, NY: Guilford Press.

Muthén, B. O., & Asparouhov, T. (2009). Growth mixture modeling: Analysis with non-Gaussian random effects. In G. Fitzmaurice, M. Davidian, G. Verbeke, & G. Molenberghs (Eds.), *Longitudinal data analysis* (pp. 143–165). New York, NY: Chapman and Hall/CRC.

Muthén, B. O., & Asparouhov, T. (2012). Bayesian structural equation modeling: A more flexible representation of substantive theory. *Psychological Methods*, *17*, 313–335. doi:10.1037/a0026802

Muthén, L. K., & Muthén, B. O. (2017). *Mplus user's guide* (8th ed.). Los Angeles, CA: Author.

Paulhus, D. L., & John, O. P. (1998). Egoistic and moralistic biases in self-perception: The interplay of self-deceptive styles with basic traits and motives. *Journal of Personality*, *66*, 1025–1060. doi:10.1111/1467-6494.00041

Schwarz, N., Kahneman, D., & Xu, J. (2009). Global and episodic reports of hedonic experience. In R. F. Belli, F. P. Stafford, & D. F. Alwin (Eds.), *Calendar and time diary methods in life course research* (pp. 157–174). Thousand Oaks, CA: SAGE. doi:10.4135/9781412990295.d15

Shrout, P. E., & Lane, S. P. (2012). Psychometrics. In M. R. Mehl & T. S. Conner (Eds.), *Handbook of research methods for studying daily life* (pp. 302–320). New York, NY: Guilford Press.

Solomon, B. C., & Vazire, S. (2016). Knowledge of identity and reputation: Do people have knowledge of others' perceptions? *Journal of Personality and Social Psychology*, *111*, 341–366. doi:10.1037/pspi0000061

Sutton, R. I. (2007). *The no asshole rule: Building a civilized workplace and surviving one that isn't*. New York, NY: Warner Business Books.

Tenney, E. R., Vazire, S., & Mehl, M. R. (2013). This examined life: The upside of self-knowledge for interpersonal relationships. *PLOS ONE*, *8*(7), Article e69605. doi:10.1371/journal.pone.0069605

Vazire, S. (2010). Who knows what about a person? The self-other knowledge asymmetry (SOKA) model. *Journal of Personality and Social Psychology*, *98*, 281–300. doi:10.1037/a0017908

Vazire, S., & Carlson, E. N. (in press). What do we know when we know ourselves? In O. P. John & R. W. Robins (Eds.), *Handbook of personality: Theory and research* (4th ed.). New York, NY: Guilford Press.

Vazire, S., & Mehl, M. R. (2018). Knowing me, knowing you: The accuracy and unique predictive validity of self-ratings and other-ratings of daily behavior. *Journal of Personality and Social Psychology*, *95*, 1202–1216.

Vazire, S., & Solomon, B. C. (2015). Self- and other-knowledge of personality. In M. L. Cooper & R. J. Larsen (Eds.), *APA handbook of personality and social psychology: Vol. 4.*

Personality processes and individual differences (pp. 261–281). Washington, DC: American Psychological Association.

Watterson, B. (1992). *Attack of the deranged mutant killer monster snow goons*. Kansas City, MO: Andrews McMeel Publishing.

Weidman, A., Sun, J., Vazire, S., Quoidbach, J., Ungar, L. H., & Dunn, E. W. (in press). (Not) hearing happiness: Predicting fluctuations in happy mood from acoustic cues using machine learning. *Emotion*.

Wilson, R. E., Harris, K., & Vazire, S. (2015). Personality and friendship satisfaction in daily life: Do everyday social interactions account for individual differences in friendship satisfaction? *European Journal of Personality*, *29*, 173–186. doi:10.1002/per.1996

Wilson, R. E., Thompson, R. J., & Vazire, S. (2016). Are fluctuations in personality states more than fluctuations in affect? *Journal of Research in Personality*, *69*, 110–123. doi:10.1016/j.jrp.2016.06.006

Wilson, R. E., & Vazire, S. (2015). Taking personality to the next level: What does it mean to know a person? In R. A. Scott (Ed.), *Emerging trends in the social and behavioral sciences: An interdisciplinary, searchable, and linkable resource*. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118900772.etrds0327

Wilson, T. D. (2009). Know thyself. *Perspectives on Psychological Science*, *4*, 384–389. doi:10.1111/j.1745-6924.2009.01143.x

EAR Coding Protocol Details

Personality State Response Options

In the first version of the coding survey, research assistants had the option of selecting "Not applicable". In the second version, we changed this response option to "No way to tell", and asked coders to try their best to make a judgment on the 1–5 scale, and to only select the "No way to tell" option if there was no information that could be used to make a judgment on that personality state (e.g., if the sound quality of the files provided insufficient information). In addition, we slightly modified the wording of the items from how the participant "acted" or how they "were", to how they "seemed" during the hour, to remind coders that we were interested in their holistic impressions.

Informativeness Ratings

In the first version of the coding survey (i.e., roughly the first three coders per participant), coders had five options for judging how informative the hour was, with instructions for each scale point (1 = no noise, 2 = there is noise but not sure what they're doing, 3 = there is noise and you can tell what they're doing, but not what they're saying, 4 or 5 = talking; we asked coders to make a judgment about how informative the hour was between 4 and 5). We instructed coders to only complete the survey if the hour block was at least "3" on informativeness. However, several coders completed personality state ratings for hours that they rated as being uninformative (i.e., rated as 1 or 2). As these hours seemed to contain information on participants' behavior (based on the coders' open descriptions of what the participant was doing), we recoded surveys with at least some completed ratings as being informative.

To prevent confusion, in the second version of the survey (i.e., roughly the last three coders per participant), we simplified the response options to three options ((1) No noise, white noise, or sleeping in all files; (2) Uninformative noise in all files; (3) Information on participants' behaviors or situation in at least one file), and recoded the first two options as "Uninformative" and the third option as "Informative".

Number of Coders

If a coder did not finish coding all hours for a participant during their time as a research assistant, the participant was reassigned to a new coder, who coded that participant from the beginning. This meant that some hour blocks were coded by up to 14 coders. In addition, due to human error, some hour blocks were coded by fewer than 6 coders. For the current analyses, we decided to include a maximum of 6 coders as indicators for the latent variables, to reduce model complexity and convergence issues. To decide which coders to retain (out of the possible 14 coders), for each participant, we rank-ordered coders by the number of hours they had coded for that participant, then kept the 6 coders who had coded the most hours for each participant.

Model Specifications and Estimates

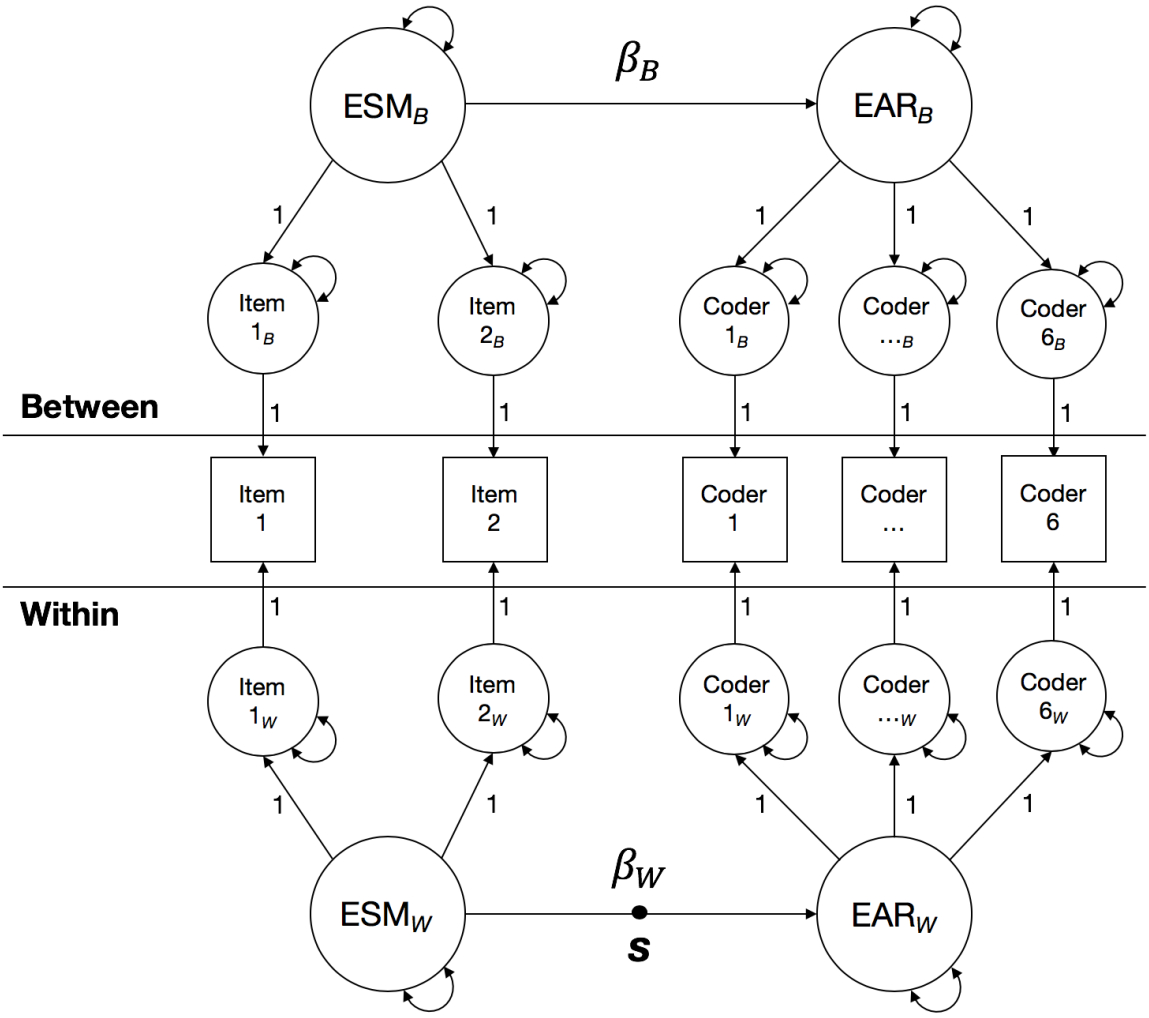

Figure 1. Multi-level structural equation models. For the neuroticism model, as the ESM latent variables had three indicators, only the first loading was fixed to 1. β_B = between-person effect, β_W = within-person fixed effect, s = random slopes. ESM = Experience Sampling Method (i.e., self-reports); EAR = Electronically Activated Recorder (i.e., observer reports)

Table S1

Within-Person Associations Between Self-Views (ESM) and Observed Behavioral States (EAR Codings)

	Unstandardized Estimates		Standardized Estimates	
Personality State	*b*	95% CI	*β*	95% CI
Extraversion	**0.58**	**[0.55, 0.62]**	**0.63**	**[0.60, 0.66]**
Quiet	0.43	[0.40, 0.46]	0.54	[0.51, 0.57]
Outgoing, sociable	0.40	[0.37, 0.43]	0.52	[0.49, 0.55]
Agreeableness	**0.29**	**[0.06, 0.57]**	**0.22**	**[0.13, 0.30]**
Considerate, kind	0.05	[0.02, 0.08]	0.13	[0.06, 0.19]
Rude	0.07	[0.03, 0.11]	0.15	[0.09, 0.22]
Conscientiousness	**0.46**	**[0.37, 0.57]**	**0.48**	**[0.41, 0.55]**
Reliable	0.06	[0.04, 0.09]	0.15	[0.10, 0.19]
Lazy	0.17	[0.15, 0.19]	0.36	[0.32, 0.40]
Neuroticism	**0.12**	**[0.08, 0.15]**	**0.27**	**[0.21, 0.32]**
Worried	0.06	[0.04, 0.08]	0.14	[0.09, 0.18]
Relaxed	0.12	[0.10, 0.15]	0.28	[0.24, 0.32]
Depressed, blue	0.04	[0.02, 0.06]	0.14	[0.09, 0.20]

Note. Standardized estimates are the within-person standardized estimates averaged over clusters, provided by Mplus. 95% CI = 95% credibility interval. Results for composites are shown in **bold**. ESM = Experience Sampling Method (i.e., self-reports); EAR = Electronically Activated Recorder (i.e., observer reports)

Table S2

Latent Between-Person Correlations Among Self-Views and Observed Behavioral States

	Self-Views (ESM)				Observed Behavior (EAR Codings)			
Personality State	E	A	C	N	E	A	C	N
ESM								
Extraversion	(.80)							
Agreeableness	.26	(.47)						
Conscientiousness	.62	.78	(.76)					
Neuroticism	–.40	–.54	–.46	(.84)				
EAR								
Extraversion	**.46**	.16	.25	–.07	(.84)			
Agreeableness	.06	**.10**	.14	.16	.05	(.37)		
Conscientiousness	.12	.34	**.39**	.02	.32	.74	(.53)	
Neuroticism	–.01	–.05	.05	**.28**	–.21	.25	.24	(.31)

Note. $|r| \geq .29$ and the correlation between ESM conscientiousness and EAR extraversion had 95% credibility intervals that excluded 0. The validity diagonal (i.e., correlations between ESM and EAR reports for the same traits) is shown in **bold**. Between-person omega reliabilities (ω_{BP}) are shown in parentheses on the main diagonal. ESM = Experience Sampling Method (i.e., self-reports); EAR = Electronically Activated Recorder (i.e., observer reports).

Transcripts for Self-Observer Discrepancies in Agreeableness and Neuroticism Ratings

Table S3

Within-Person Standardized ESM and EAR Scores and Transcripts for the 25 Largest Positive Self–Observer Discrepancies In Momentary Agreeableness Ratings

Z_{ESM}	Z_{EAR}	Z_{ESM}– Z_{EAR}	Transcript
1.63	–2.57	4.20	You like it
1.76	–1.92	3.68	And it's just pointless. Like why? | Or we could be Sam and I would be Sam. Ok, so you could be Sam and I could | What. Ok well guess we know what we're doing during
1.41	–2.18	3.59	Want that class anymore. Fuck the police, I hope this is getting recorded. | Why do you have a different college sweater? Oh. Alright. Yeah
1.35	–2.22	3.58	No they were not. That wasn't on there before. I don't know. No. You're absolutely wrong. You're doing it wrong.
2.00	–1.3	3.30	I don't know, I have
1.41	–1.83	3.25	I especially because my brother did not let did not my her twin brother did not have her in his wedding which is such bullshit. Yeah. He like hurt his like his wife just like had her own maid of honor. Like it did not even could have been a bridesmaid. That is a typical thing. You let your siblings. Yeah like | What time? I would be like. I was going to like yawn and then sing but then I just didn't so I was just like. | Like did she love her dad though or like and they are still together? | I got attached to when I thought I could hook up with him but like. I guess the the only way it could be like work as a friends with benefits is if you don't do that often. Like I hooked up with that guy who I had hooked up with at prom over the summer and the middle of the summer but like. Yes. Because that would that kind of thing I could | Well she didn't lie to us. Lied to us. I was going to let's room with I said let's room with
0.51	–2.72	3.23	Mhm. | I know
1.06	–2.16	3.22	Yeah yeah no exactly. | I put Hannah. I put Hannah. I know I was. I know. You could I'm just | She's cool she's cool. She, you're going to really like her because she's like she's very like your kind of personality and just like blunt and like down to earth and like. What? And they also call kappa kappa ginger. Yeah. I was like. Ginger?
0.71	–2.46	3.17	Yeah
0.00	–3.02	3.02	That's why I'm saying | I know I've already | You're right. I have to forgive myself for hurting you. It's just | I can't believe I ruined porn for. Oh god. Oh my god. | I haven't been. It's not like I'm, I'm not surfacing anybody. No, I'm just
1.91	–1.08	2.99	Right now and I am just like, what does that mean? I'm kinda

			scared. Oh my gosh, life is so tiring. With me, nothing. I'm just tired. | You even checked? Oh. | No, how about no.
1.15	–1.77	2.93	When the destiny calls you, you must be strong. I may not be with you, but you got to hold on, they'll see in time. I know, we'll show them together cause you'll be in my heart. You'll be | Look at these words. Look at these we off the hook. But there's the bubbles under the sea. Under the sea. Literally after I leave. Know how to play.
0.91	–1.99	2.90	Really nice office. It's a really nice office. Yeah um. Oh you just do the surgical things? Oh. Like um um. | Um ok well as a as a topic of conversation it may come up during the night that where I was last night um and it would make it a lot easier if you can if we can decide between us that I was with you. Ok. | Yeah, yeah captain. Well are you going to tell her that I was out with another woman last night? Then isn't that creating a false impression? Isn't that in fact a lie? | Definite. But um there is interest. That's half the battle. Fingers crossed. Just few minutes ago. Um you got it you got it you got it here early huh? Huh? Huh? No no it's not hot. | I know I had to get out. No no no of course not, I just had to leave. I had some place to be. Preparing for thing. You know? You know. Um yeah. I promised myself I wouldn't go into artist class. | Ask. Yes or sister, sister or brother is a good good point you know. If you did something wrong then don't tell your parents. Yeah yeah it's good. Yeah yeah. I will try I will try. | Oh really? That must be exciting! Absolutely. In contrast to? Thank you. I'm glad you like it. Fine.
1.29	–1.39	2.68	Oh ok. | They can't. I don't know, here, read it. | Ok so, I'm talking about force analysis now. Um. So, so like for like something like thread of substitutes it's just saying that.
1.81	–0.84	2.65	Yeah, but all they did was like yell at each other then be in love. This lab is so stupid. No, but this is actually stupid. Like so far it's just, you'll see. It's like the first bio lab, kind of. | Negative 6. Yeah and then it gave her like part of it and then there was like an a in it and I was like so under pressure in my situation that I could not like think of what it was. I'm sure I could have figured it out but like. But like I was too flustered
1.77	–0.87	2.64	Hi. | I don't know who's done yet. Alright, go ahead. Could we get squeezes on the third omega 5?
1.11	–1.49	2.59	This is scary. Just walk around with beers. Cassie is not here, and Jason, in a what? | It's too bad. I'm just sick of I'm just gonna stop responding to his texts. He's annoying as he just like keeps on trying to and I'm like no leave me alone. Doesn't matter. Oh Sandra | Oh so did we. No ways. Except for. So awkward that then you could do it. Oh Beyonce I heard you. You coming later.
0.83	–1.76	2.59	Yeah. No. Oh shit. Y'all bitches going down. I am perfect. I will win this. I don't lose. Oh that's that's Honolulu for the weekend. That's a | I didn't take it. I didn't go. I woke up like on time and everything and I was like I still have to shower and I was like you

			know what fuck it I'm too tired and so I just didn't go. | I don't fucking know and I don't care. I give up. What? I just fucking did it hello.
–0.86	–3.44	2.58	Because I've been here so. On the way here.
0.85	–1.70	2.55	Right? Is the | Like walking the and. I'm glad you went out just to
0.64	–1.80	2.44	Yes.
0.37	–2.05	2.42	What are you talking? I guess but it is time for where are you from she is like I am from France. | Thank you. Um I am ok. No way.
–0.24	–2.63	2.39	Oh walk of shame walk of shame.
0.66	–1.72	2.38	Ok. They're staying at a hotel, they're fine. I stayed at the hotel last night. So yeah. | I was really embarrassed. I was just like | Yeah, yeah a spider like dropped down. | Yes?
0.68	–1.70	2.37	Yeah. Helen. Helen. You look like you just woke up.

Note. Individual transcripts of 30 second recordings are separated by a vertical bar (|).

Table S4

Within-Person Standardized ESM and EAR Scores and Transcripts for the 25 Largest Negative Self–Observer Discrepancies In Momentary Agreeableness Ratings

Z_{ESM}	Z_{EAR}	Z_{ESM}– Z_{EAR}	Transcript
–2.32	1.52	–3.84	Not her just her mom and dad. | This is never cause we would not have worked. It just wasn't gonna work. Cause I was his first serious relationship. We were each other's first love. We thought it was perfect, and it wasn't. But we thought it was! And then it, I can tell you, trust me, breaking up helps. And you have a good support system here. I thought that was wrong. You can come into and mel's room, just have a glass of wine. You know. | From your first semester freshmen year. First semester freshmen year. I think that if it's something that you want, you should just work for it anyway. I'm getting like a B- in intro to psych and I want to be a psych major. But I mean like, my worst class is the class, is the subject I want to major in. I'm still gonna major in it though. | Tea makes me have to go like there's no tomorrow. 2 cups of tea makes me really have to go. I'll be back.
–2.31	1.50	–3.82	Uh-huh. Oh! Nice and warm. Uh-huh. Uh-huh. Uh-huh. Uh-huh. | Uh-huh. Maybe i'll take a break from calculus 3. And then I'm sorry come here. I'm sorry. But I'll just help you. I don't know.
–2.29	1.50	–3.79	I was already
–2.60	0.92	–3.53	It changes. Like goes for our acceptance. And see like wherever his arm is, that's, that's | done. Not there right? Yeah mhm. Here?
–1.07	2.13	–3.20	Just leave. Oh my god. Oh my god. Oh. Oh. | Um. | Me too. Mm | I don't . Um I think my TA said that he wouldn't ask about like cones or pyramids or anything. Mhm. Yeah | We can plug in 0 for that? So | Yeah I just did like a rectangle, and then that's the cross-section. Mhm. Ok. Um, so water is pumped a rate of
–2.49	0.68	–3.18	It's in the back. It's in the back of the book. Movement is balanced. Ok. I need to study these. Ok. Yeah. | No yeah. Please like stop for a while and then like incubated. No yeah I'm over it. Yeah. It lasted like 24 hours for me. Ouch. It's like a virus. What are you guys studying? Call claire. | You read all of it? Wow. I'll be doing that tonight at the library by myself probably. | Um it sounds like weird like something we'll have to study. Like what's the task's name? Retention, memory. Yeah. It's like everything you are basically. Serotonin. So it's like. | What goes with learning? Ok so emotion, impulse. I'm not sure. I'm not sure. Oh yeah sensory signals were | knockout. Is that a thing? Wait hold on let me read it again. She creates a string of mice that do not have this gene. What? Can we just say knockout because I don't feel I feel like it's not b or c.
–1.71	1.28	–2.99	Look. | Dang it.
–1.41	1.57	–2.99	Too dangerous. I feel bad. Yeah. | It's already cut and everything. Oh you're welcome. What? Well I have to go to church anyway I

			can't smoke. Oh my wallet is full of money. I haven't feel so rich forever. | No. Boy, this room will be empty.
–2.00	0.98	–2.98	No that's at ten unfortunately. No. Not on the daily. I don't know. She usually falls asleep before I do which is impressive because there's a 2 hour time difference. Mhm and year book. | What did you think she said? What do you think she said? I don't know it is like frozen. Oh she said tupperware. | I don't think people. Cookies don't sell at. Should I make them dozens of cookies? I have to make a problem for economic whatever and she wants me to make it the supply and demand of froyo. | Because it was hard and then I was like ah. And I got back. I got back at like nine after like so many classes. And then my group wanted me to work on the stupid business project like right when I came back. So then I had to do that. And then I ate at like ten and it was just like ah. And then I was all weird. | Escalator? Oh my god. I shouldn't even have said it. I was going to. I should have. Dang it. I shouldn't have said it. Every time.
–2.09	0.89	–2.97	If I can do all C's now I can get even like a B plus or something. Me too. I got that bag and I was like, oh my god that's like, oh my god then if you get sixty percent? Oh oh god that's the lowest test I've ever had. | Water what? Yeah. Ok. I get it. Ok. Good metaphor. Anyway, it's specifically in terms of productivity and economic development.
–0.93	1.76	–2.69	Shame and stuff comes into play instead of like. So I feel like in, that's just a very individualistic perspective on collectivism. So, um.
–0.54	2.13	–2.67	Do you want to write it on the board? So they can yeah. | Stuff for like children. Um oh and then Ronald McDonald um is um Ronald McDonald's house charities. It like houses residents um. It houses families who have kids in the hospital. And so we go there on an afternoon and make dinner for them. And so that's gonna be like a 4:00 to 8:00 commitment. And like that's a really cool because you guys have to cook yourselves so. Oh sorry. | Some people like to color so do what you feel like is empowering for them. Oh no. Alright and then we're also gonna pass around there's | Yes. Thank you. Oh I put my return address on it too if that's ok. Oh you don't. Oh ok. Yes you guys don't have to do that if you just want to like leave it. I hope you guys yes the whole thing too. | Yes she runs like a recipe book but still aw were you waiting for me? For you. Wait what's your name again?
–1.74	0.90	–2.64	Ok so you're going to have to it gets better. | No they don't. There's no change of shoes on the ground. | Classic, classic potheads. | I went to the arc with shawn today yeah but then he saw me and then we went to uh the no we went to the blue car. I wasn't drunk. Yeah. Yeah it was | You're good. I hope so. You can order them i'll have them, have them send you an email. There it's smart though, the way they you have to fill out this form and you have to put down

			your email so you can pick the one that you put in there and then you have to put down your parents' contact info and then
–1.09	1.54	–2.63	ID? We will give it to me? Over a bar.
–3.33	–0.73	–2.60	For an hour and a half and I'm getting paid 20 dollars to have this and I'm getting paid 40 dollars to fill out a survey. I know I know I'm like making bank man. It like no. Yeah. So far. Um only. It's like 5 percent of my day, I don't know when it records. | I'm like I want to be in my bed. No, it's listening to everything we're saying right now. It's an, it's called an EAR. No. No! It like records 10, 5 percent of my day everyday for the next like 6 days. Mhm. It's like yeah it's a psych study. I'm getting paid for it. Dollars
–1.38	1.22	–2.60	I'm gonna get it, i'll try to do it if I can.
–1.06	1.52	–2.59	You ready? It's so but yeah the next one can be nickelate. Niccolo. I was just wondering the same thing. Nickel. Ate. Yeah exactly. Ok. | I'll be back in a sec. | I mean, a lot of it's h's and those don't look very happy at all. I don't trust me. 16, 14, sorry. | Dude, it's been like a whole 10 seconds since we last talked. | Look at matt he's dying. You say STS every time. What's the team? STS is student technology services. | I'm sure we can find sugar somewhere. I think we actually have to do this now. But where? Should we go dressed up? Should we make a Target run right before? You need more fruit snacks.
–1.75	0.81	–2.56	C, e, d, e, f, g, a, b, c | d, a, c, f, e, f, e, e, c. | Yeah. Ok. Um, what you talking. Would you mind remind me what that uh one is. Yeah. Yeah.
–1.11	1.42	–2.53	Yeah just let me know where. Who wants to take over? Um. Take a peek before we get there take a peek before we get there so we know before | criminal.
–2.25	0.26	–2.51	Yeah | Can I have extra meat on them? Oh yeah provolone | There is not as much leaders like there is only 4 leaders but not any like I don't even know how much work is it. Yeah like it is not that bad.
–1.06	1.39	–2.45	They make it seem all easy. It's not. Hm? | Cool. No. Ok we go. Oh oh. Using magic? Oh god no. | Mhm. That's so cool. Mm. Cool. Thanks Piper. Oh. Ugh. Can I take it? | Oh here. Oh. Hi talk to me. Hm. | Hm. Yeah. This way? Oh here? Tanning? Oh I have those. | No. Comfort is everything. Done what? Beta is calling you. Yeah. | Yeah. Yeah. I don't like commercials like this like I really don't. I don't think you should manipulate people's emotions like that kind of you know? Especially when it's someone speaking for someone else.
–0.27	2.16	–2.43	She sounds like such like person I apologize to you in my head every time I open my drawer like my head says sorry shelfa.
–1.65	0.77	–2.42	Import them
–1.41	1.00	–2.41	Encouraging. | Advocating. | That was awesome. | Um my god but ok. So way over here. I have never.
–1.29	1.07	–2.36	Let me see. Here let me see it. Here give me. | I don't really know.

	But it's good. It's 10 dollars. I mean there's also that. Probably don't think it is though. Ok. Ready? Mhm. I got this at um | Um that's amazing right now. Can you see Laura? Oh my gosh. | Would you like a piece of gum? Cinnamint. Is that ok? Mhm.

Note. Individual transcripts of 30 second recordings are separated by a vertical bar (|).

Table S5

Within-Person Standardized ESM and EAR Scores and Transcripts for the 25 Largest Positive Self–Observer Discrepancies In Momentary Neuroticism Ratings

Z_{ESM}	Z_{EAR}	Z_{ESM}– Z_{EAR}	Transcript
2.15	–1.11	3.26	Mhm.
1.66	–1.58	3.24	Yeah. | Fuck. My bad. That's a good call. I got it I got it. I got it. Hold on I got it I got it. No I don't got it. | Kicked you off. I kicked you off. Get off my phone. Should I get that. Could we could we can we something like right now? | Damn girl. Yeah. Yeah. No!
1.50	–1.71	3.21	Did you tell what are you doing ok if you monday. Like what about wednesday? | A meeting. Yes. But all the guys who flannels with the sleeves cut off. | I did. That's totally how I talk right? Who are these people? | And then there's. There's more m&ms. Exactly this this one has dark chocolate. Mhm. It has it mixed in so I surprised them. And then I have this other bag of m&ms.
1.89	–1.29	3.18	What?
1.91	–1.10	3.02	Hi. It's ok. Um. | First. I got the first one that said emergency is over and the next one said potential threat to safety exists in down after I e-mailed Nina and told her that I'm like dying and I'm sorry if I just. I like can't see when I going to class up the stairs I got coffee in my eye and all over my face I don't understand how I can't really see right now. | I'll believe. So excited. | I don't know. I can't tell with my voice. | Chuck what would I type in to get to that? Nah, it's ok. | Madeline you started this. Madeline's like.
1.21	–1.70	2.91	2 episodes is like an hour and a half. Max and monty came in my room around like 2:40. Cause my door was propped open still. Oh. They had come back from a party. They know people. They don't go out that much, but they go to good parties. Yeah max was so drunk it was so funny. He was like laying on the floor and then he. | Yeah um they'll give you a number. So if there's a table inside and we're lucky. The past two times I've come we've gotten a table inside out of pure luck, we get there right as someone's leaving. Well if it's not, people aren't going to get up for a while | No. No one went. Yeah I don't think she's coming here but um I was like are you going with Katie and she's like going where and I explained it and she's like what are you talking like oh. Yeah. | Then it just like I realized that it wasn't gonna work. No but I didn't wanna like, then you guys would be the only parents, cause her mom would be with her friend. I didn't realize. Happy. They have a baked maple pumpkin latte. | I think I was here last sunday.
0.96	–1.94	2.90	No, we were going to Target. It's dark it's dark. Where did it go? Oh. That's why I'm kind of confused right now. Does anyone have to go? | Super into it at all but. Oh yeah. Do you?. Oh really? I didn't know. I hadn't been following it at all. No. Um I played tennis for 4 years. Yeah. Yeah true. Yep. Um so what's your plan for. | Like

			ever. Alright I’m sitting in the middle. Someone has to scoot over, my butt's not that small. We were genuinely concerned, we thought the test like literally made your brain explode. Because everyone was back and we were like, wait where's Andrea? Oh god. On that one? | Just no. Can I. Get your shit together. Yeah. Yeah yeah. True. | Just to make me feel awkward. Thanks Garrett. Yeah we're getting tickets. | That's good. That's good. You're good. You good. | A Molly thing. No.
1.71	–1.18	2.89	I don't like my job this for like 3 hours and 40 minutes and then spend 20 minutes like refilling printers and re-checking staplers and that stuff. Yes. Yeah if you think about it I am kind of getting paid 120. | Also day dot net seems good. | Mexico, Chile, Spain or are we doing 4 now? What about Chile? Alright why did I hear Portugal? Ok. Portugal wouldn't be that special. Right from Spain. | I think you used to | Person. | I can't help it. It's too good. | It wasn't too bad. I took like 20 minutes a a day for a week and then I was like really slow for a week and then I now I’m faster than I was before.
1.33	–1.40	2.74	You got to be kidding me. | [[you'll be falling.]] How can there be a less of a chance?
0.94	–1.76	2.70	I just noticed that like one of Mulan's never had it so she like flings her long hair it's like super powerful when she cuts it off it's like she | No | Well like whenever I talk about that she always gets really defensive. And then like and then my mom reflects it. And then when I ask why she says people ask me
1.07	–1.58	2.65	Wait, why was I not aware of this. I'm like, I’m probably just, stop what you're saying. We were talking about so much it's impossible to remember. | Had all of them in, uh, the bottom drawer that, of my stuff, and there was nothing in my bottom drawer, like, it was completely empty. But I haven't worn anything. I was like, this is the first time I get to wear my WashU sweatshirt, and then I, I didn't have it. I don’t know. I'm'a going to have to clean up my room anyways. We'll find stuff. This is | Jewish parts of New York, there's like, this one place that's like, just a pickle place, and that's all they do. And they have like. I don’t know. But, it's like you go into like, the basement, and then they just have lots of jars of pickles, and it's like, really cool. And I went, I haven't gone a lot. | wait. I haven't | I didn't know your names. It doesn't go like that. | home-dog. Home-skillet, uh. Yeah. Uh.
0.55	–2.05	2.60	I think she'll get me fabric because I need to be able to give some and I have some wow. | You know. | No, this is sparta. It was a great movie. I love Batman. Yes | Where'd you go? What? I mean, that happened to my friend but she went to some school in New York. Um well that makes sense. Where do you live? That makes sense. | I want to hike the Appalachian trail. Casual. | We're gonna go in here. Oh here, really are you leaving those paper towels there as decoration or? Sorry. I thought you would do something with the

			paper towels. Ok.
1.63	–0.94	2.58	Sweet. Food. Sweet.
1.73	–0.79	2.52	Oh you're gonna cook it. Yeah. But you also can't go at shit. Alright. I'm gonna go with that.
0.62	–1.89	2.52	Vegetarian
1.56	–0.94	2.50	So much email. So what's with ok. Mhm. Mhm.
1.71	–0.74	2.45	Yeah. Yeah. For the fundraising yeah for fundraising dinner | On the next big street, over from, like it's on the other side, at the end of the campus. So like here, and you keep walking, away from the on forest park, it's the next big intersection. And there's like a weird looking building on the side that sells high-tech and then there's like two Big Ben like metro stops on either side of the streets, and then high-tech it looks like they are part of the metro station, but it's not. It's really weird.
2.35	–0.09	2.45	Really funny. Um. It's funny. I think I kinda want to run the 200 like like cause like I feel I'd do good for like like but I do not know. Do you think I should just tell coach like
0.69	–1.75	2.44	Is it even going on. | B over b must be proud of this batch! | Scary! Stop!
2.56	0.13	2.44	Ok so we can say people justify it by saying they can further like cultural understanding and like um create new audiences because
0.66	–1.77	2.43	Hey good. | I thought you rushed though didn't you? Oh ok. That's cool flying solo. I dig it. Yeah? That whole stereotype was that involved in the decision? Actually though. I'm kind of stressed out about that. And like joining a sorority | Yup. Um do you have a cold or the cold weather? Ok. Yeah that's what it is. | What time is it? No | Yeah. You have to put it in here? Ok. Mhm, mhm. | Buddhists. I have a little Buddha sitting on my desk. I probably shouldn't have told you that. We are not.
2.19	–0.21	2.40	Yeah I want to say number nine is. | Those are | I don't know how to do this one. I'm throwing in the towel with it. Yeah. I'm so confused.
0.87	–1.52	2.39	Oh, but stores. And when, instead of throwing out the bread that's about to expire, they recollect it all and they go like pack it like the package up. Conversation with, just because it seems like he | No it's ok. Wait you. Yeah, I ran for a while then | What's the difference between peanut butter and jam? Peanut butter and jam. No. No it's a joke. I said one's | Forest park. We're gonna get like burgers, steaks, baked potatoes and just like not like super formal but just relaxed. Like you know what I mean I'm super stressed out the night before like. So yeah. Um. I mean once I get a job I'll be sure to request
1.63	–0.70	2.33	I don't think I brushed them last night.
1.71	–0.56	2.28	On Tuesday April 5th. So yeah, almost. It's craziness, I know. What is it? Oh yeah. I just attempted to stuff a thousand envelopes. Oh multiple times, and I'm like my skin is all dry and gross from. | Ok, I'll get it. Yeah. It's for the people | Yeah. Why? And also how are

	they gonna know where they are going to be? Unless, you have the itinerary? Nine. Yeah. Yeah. | Just chatting. There's Nessie. Hi! It's ok. How's it going? | I still think you should get the 8 or just like so it's like it's like super cheap and super cute. Yeah. Yeah. So it's like | what? Yeah, Hannah five-fifty. Five-fifty. Can I join your study room? | Oh ok. Me too but apparently, that's what needed to happen. Ok.

Note. Individual transcripts of 30 second recordings are separated by a vertical bar (|).

Table S6

Within-Person Standardized ESM and EAR Scores and Transcripts for the 25 Largest Negative Self–Observer Discrepancies In Momentary Neuroticism Ratings

Z_{ESM}	Z_{EAR}	Z_{ESM}– Z_{EAR}	Transcript
–0.69	3.52	–4.21	Right now and I am just like, what does that mean? I'm kinda scared. Oh my gosh, life is so tiring. With me, nothing. I'm just tired. | You even checked? Oh. | No, how about no.
–1.33	2.31	–3.64	And it's just pointless. Like why? | Or we could be Sam and I would be Sam. Ok, so you could be Sam and I could | What. Ok well guess we know what we're doing during
–1.27	2.30	–3.57	What happened? | Go Bears! | Yeah probably. See you that kid. I've never met a happier person in my life yeah. Like a positive happy not like a. | Do you want to just get ice cream now?
–1.54	1.96	–3.50	It's been a long week dude. Long day, long weekend, long | It's one thing if you go to class, take notes, read the book and you can't get it. It's another thing if you like have, like you know haven't started until like 10pm, Thursday night. That's your fault but then. I don't know why he is doing it. | Cultural elitist city snobs. So.
–1.11	2.22	–3.33	I'm so sleepy.
–0.67	2.59	–3.26	Pardon? Is this general chem? Um stop that was weird. I don’t know you have never really understand. It's just like you go to lecture then you don't understand and then you just do the lab and you don't understand. | Yeah.
–1.02	2.21	–3.23	I'm like I don't know! I went through the books, I went through all my notes like I honestly don't know what other material I could've studied like that that's one of the reasons why I’m so frustrated is because I don't know what else I could've studied. This is also the teacher I have like complained about because he's just so horrible. Like everyone walked out of there and was like he did not teach us any of that. Like everyone walked out of there and was like I that could have been. | No, no, no. What are we putting in this food this is horrible. Actually 12.5 grams of saturated fat. I'm not I’m not even a healthy eater but like that's disgusting. | Alright!
–0.34	2.89	–3.23	Let's not speak of it for the rest of this night. It was bad. I'm going to go to every single help session from now on. Every one. No more, no more on my own. Ok. For the most | I feel like it's starting to sink in on him it's like going to work out. It's taken 3 years but it's starting to sink in. Yeah. We had such. This is confession from I accidentally came across his diary once and I read a few pages. I'm a bad sister. And I never did it again. I felt so bad for that. | If I could get my hands on some I will how do you get them? Really? | What if we did our just like smile you can't say no. It's real. How do you do it well?
–1.57	1.59	–3.16	Hello Dean it's uh I just called to see if uh my letter of recommendation will be ready by the end of this week and see if

			there's any progress on it. Um if not, if you don't have time.
–1.79	1.29	–3.08	Ok no problem. And I put them on computer e which is one of the ok computers. Ok great cool. So
–1.13	1.93	–3.06	Yeah they kept my feet dry and uh yeah and I was at the metro station and I'm standing there you know on the platform waiting for the train to show up and this guy next to me and I like I distance myself from everybody by like 10 feet and this guy like walked over to me and he's like where are your shoes from? And I'm like here? And he's like girl they
–0.73	2.33	–3.06	I don't know. I don't know. I mean she was like she kept saying like if you're on the waitlist if you're on the waitlist if you're on the waitlist and I was just like oh does that mean I'm on the waitlist but like she really like I mean. No she said she'll tell us by tuesday before thanksgiving and if not we'll find out.
–1.86	1.19	–3.05	Ok um well I'm so right now. No no no. But like said 10:30 and it's been like 2 hours so we got to like. Right. | Oh no no. No it's not no. Guys no no no no no it's just a natural like back ache correct-al.
–1.36	1.68	–3.04	Oh my god.
–0.86	2.14	–3.00	What's up with your phone? Oh. It's gotten a lot worse. That sucks. | ok, well we'll help that out. Like a scrub. | I need something to wash this down. What is that shit? Sprite? That shit is good. | Um.
–0.47	2.49	–2.95	Hello. Hi, good. Um ok um kind of.
–1.27	1.68	–2.94	So when you're doing the song it looked like | Um. | Is there going to be a um like during parent's weekend? It's like the weekend of October 25th. | It's 1/16th. Wait no it's
–1.16	1.69	–2.85	Do you want to write it on the board? So they can yeah. | Stuff for like children. Um oh and then Ronald McDonald um is um Ronald McDonald's house charities. It like houses residents um. It houses families who have kids in the hospital. And so we go there on an afternoon and make dinner for them. And so that's gonna be like a 4:00 to 8:00 commitment. And like that's a really cool because you guys have to cook yourselves so. Oh sorry. | Some people like to color so do what you feel like is empowering for them. Oh no. Alright and then we're also gonna pass around there's | Yes. Thank you. Oh I put my return address on it too if that's ok. Oh you don't. Oh ok. Yes you guys don't have to do that if you just want to like leave it. I hope you guys yes the whole thing too. | Yes she runs like a recipe book but still aw were you waiting for me? For you. Wait what's your name again?
–0.72	2.12	–2.84	One I used before do you have quality on this art store place. Is it cool oh wow I know that's oh my gosh ok you. | No they are due on Thursday. Thursday. Yes. Yes. Not a week from today. Not a week from today. 2 weeks from today. 2 weeks from today. From 2 weeks from today. Because our lives suck. Golly um | Yeah cause the hand looks so washed. Oh well. We'll see how it. I'll show you how to check the roots mhm

–0.61	2.22	–2.83	They're not over here. They're. I said I moved them. Do you want one? Ok. There's only two left. There's only two left so probably mhm. You should probably leave the last one for | Did you do the problem set? Ok. I don't know it's not something you can look at like. Yeah. Do we have to write like the assumptions we make?
–1.06	1.77	–2.83	Ok. So um, well yeah I mean I got sick yesterday so I've been like coughing all day. I was hoping I would get better but it's not looking like it. So I don't want to you know be coughing all over people or like. So ok so um is there anything else I need to do like talk to Melissa directly or? Alright, well thank you Rebecca. Alright bye. | Who they hit it like you at the last second you swerve. Would it be safer for you to swerve to the left then hit the wall, or for you to swerve right then hit the wall? And because you're sitting on the left half of the car. Yeah it's my question right here. I'm trying to figure it out. The other option is to slam on the breaks but that's not going to work because your break are going to skid, and like your tires are going to skid
–1.29	1.49	–2.78	Since it says. Well it says, describe the 209. Oh oh. Look what I'm asking is how you how you draw it like how. I know how I know that it's like
–1.40	1.35	–2.75	Oh sorry | Traffic | Oh my god.
–0.50	2.24	–2.74	Yeah I know. I don't know how this TA is so far, I don't even know her name. Um. But yeah. Yeah it can. I actually I'm not going to at all. I have to do all the readings. | Yeah ok. Mhm. Yeah. Ok. Mhm. Yeah. Alright, well thank you. Um I'll be sure to email that TA. Yeah good. Ok. Thank you. | Um. Oh ok. Huh. Ok. So 1 2 3 4 um 5 6 7. 8? Ok why does that count? | I know. Ew. Ok | After you. Ok.
–0.28	2.35	–2.63	I don't know it is not like I don't know when

Note. Individual transcripts of 30 second recordings are separated by a vertical bar (|).