

Personality in Its Natural Habitat: Manifestations and Implicit Folk Theories of Personality in Daily Life

Matthias R. Mehl
University of Arizona

Samuel D. Gosling and James W. Pennebaker
University of Texas at Austin

To examine the expression of personality in its natural habitat, the authors tracked 96 participants over 2 days using the Electronically Activated Recorder (EAR), which samples snippets of ambient sounds in participants' immediate environments. Participants' Big Five scores were correlated with EAR-derived information on their daily social interactions, locations, activities, moods, and language use; these quotidian manifestations were generally consistent with the trait definitions and (except for Openness) often gender specific. To identify implicit folk theories about daily manifestations of personality, the authors correlated the EAR-derived information with impressions of participants based on their EAR sounds; judges' implicit folk theories were generally accurate (especially for Extraversion) and also partially gender specific. The findings point to the importance of naturalistic observation studies on how personality is expressed and perceived in the natural stream of everyday behavior.

Keywords: personality expression, person perception, personality judgment, naturalistic observation, Electronically Activated Recorder

One bright April morning in 1949 at 7 a.m. sharp, Joan Birch roused her 7-year-old son Raymond from his sleep. With a little encouragement from his mother, Raymond sat up and rubbed his eyes. At 7:03 he pulled on his right sock. At 7:04 he put on his right shoe. His mother asked him if he wanted an egg for breakfast. Sleepily, but without irritation or resentment, he said "No." In fact, everything Raymond did, every interaction he had, every place he entered from the moment he awoke until he drifted off to sleep at 8:33 p.m. was meticulously recorded by a team of eight observers, one taking over from another every half hour.

This unique scientific project documenting Raymond's every action was spearheaded by Roger Barker and Herbert Wright, directors of the Midwest Field Station in Oskaloosa, Kansas. Its goal was to create a comprehensive "record of what a seven-year

old boy did and of what his home and school and neighborhood and town did to him from the time he woke up until he went to sleep that night . . . a specimen of the behavior and of the cultural and psychological habitat of a child" (Barker & Wright, 1951, p. 1). The project was published in 1951 as the now-famous case study, *One Boy's Day*.

What ideas motivated such a monumental empirical undertaking? In an era when virtually all research took place in the controlled environments of experimental laboratories, Barker and Wright were convinced that, to really understand what people are like, we must look at what they naturally do in their ordinary humdrum lives. Consequently, Barker and Wright resisted the urge to "zoom in" and study highly controlled and largely decontextualized basic social processes. Instead, they aimed at escaping "the confines and constraints of the psychological laboratory. . . studying persons in the ordinary contextual influences of their everyday lives" (Craik, 2000, p. 236). As a result of their strong commitment to studying people directly in their natural habitats, Barker and Wright are now considered as two of the founding fathers of ecological psychology (Barker, 1968). Although their work has been highly praised, their ideas have not yielded a strong legacy in modern empirical psychology. More than 50 years after their famous case study, the field continues to suffer from the lack of descriptive, exploratory, and phenomenon-centered naturalistic observational research that inspired their endeavor in the first place (Craik, 1986; Funder, 2001; Rozin, 2001).

The present research follows in the conceptual footsteps left by Barker and Wright by exploiting recent advances in microtechnology to provide an in-depth analysis of the personality implications of people's naturally occurring daily lives. Specifically, we aimed to examine (a) how personality is manifested in daily life and (b) how people's implicit folk theories assume that personality is manifested in daily life. Finally, we sought to bring the first two questions together by examining the accuracy of people's implicit

Matthias R. Mehl, Department of Psychology, University of Arizona; Samuel D. Gosling and James W. Pennebaker, Department of Psychology, University of Texas at Austin.

This research was supported by National Institute of Mental Health Grants MH64527-01A1 and MH 52391 and by National Science Foundation Grant 0422924.

We thank the research assistants involved in this project for their help in collecting the data and coding, transcribing, and rating the large number of sound files: Monica Adame, Michael Adams, Rohan Ahluwalia, Hillary Bortz, Melanie Burnstein, Jennifer Carbajal, Esther Chong, Kevin Davidson, Colin Fowler, Katie Galloway, Jason Guckert, Pilar Halgunseth, Bethany Hensen, Emily Hester, Katherine Hughes, Amy Kaderka, Jennifer Kuo, Habeab Kurdi, Elizabeth Lyons, Mary Clare Madeley, Jacqueline McClellan, Mollie Mercer, Radha Nadkarni, Laura Naumann, Kim Noles, Calene Oltmann, Linsey Proctor, Kim Roaten, Amanda Rogers, Ronnie Saenz, James Sennette, Emily Shirey, Kristina Slaikeu, and Adam Webb. Special thanks go to Andrea Garcia for her help in all phases of the project.

Correspondence concerning this article should be addressed to Matthias R. Mehl, Department of Psychology, P. O. Box 210068, University of Arizona, Tucson, AZ 85721. E-mail: mehl@email.arizona.edu

folk theories of how personality is manifested in daily life. Before laying out these research questions in detail, it is necessary to review the few attempts to capture personality in daily life that have appeared since Barker and Wright's seminal study.

Studying Personality in Its Natural Habitat

Barker and Wright's (1951) analysis of Raymond Birch's day was a milestone for naturalistic observation in psychology. Although their ideas resonated strongly in the field, it took surprisingly long for them to be carried even one step further. It was the act frequency approach to personality assessment (Buss & Craik, 1983) that again put daily behaviors front and center, beginning with the premise that "the fundamental goal of all personality theories is to describe and account for regularities in individuals' actions, broadly conceived, occurring throughout the natural stream of everyday conduct" (Buss & Craik, 1986, p. 141). However, presumably because of the enormous burden associated with studying long streams of behavior in situ, the act frequency approach resorted to analyzing retrospective reports of behaviors rather than records of actual behavior (Block, 1989).

Several years later, Craik (1994, 2000) revived the naturalistic observational approach to studying personality. In the early 1990s, he developed *Lived Day Analysis*, an observational method that captures individuals in the natural pursuit of their lives by following them around with a handheld video camera. Lived Day Analysis is based on the realization that "lives are lived day by day, one day at a time, from day to day, day after day, day in day out" (Craik, 2000, p. 234). In creating permanent audiovisual records of a person's day, it constituted a major methodological improvement over Barker and Wright's behavior specimen record. In observing people directly in their natural habitat, it preserved the original spirit of *One Boy's Day*. However, like Barker and Wright's method, which resulted in hundreds of pages of field notes per person per day, Lived Day Analysis accrued many hours of fascinating but analytically overwhelming video footage, practically constraining its use to ideographic analyses. In addition, the method is naturalistic but not unobtrusive because a research team following participants around with a video camera inevitably induces reactivity from the target and the people in the target's immediate environment (Craik, 2000).

The latest methodology for unobtrusively sampling observations in naturalistic settings is the Electronically Activated Recorder (EAR; Mehl, Pennebaker, Crow, Dabbs, & Price, 2001), a modified digital voice recorder that periodically records brief snippets of ambient sounds. Participants wear the EAR attached to their belt or in a purse-like bag while going about their daily lives. The method is unobtrusive because the EAR operates imperceptibly. In recording moment-to-moment ambient sounds, it yields permanent acoustic logs of people's days as they naturally unfold. In providing authentic real-life data about people's lived days, it follows in the footsteps of Barker and Wright (1951) and Craik (2000). In sampling only a fraction of the time instead of recording continuously and by not requiring the researchers' presence at every moment, however, it makes large-scale nomothetic observational studies possible (Mehl & Pennebaker, 2003a, 2003b).

The EAR differs from other momentary assessment methods developed to capture quotidian behavior (e.g., Experience Sampling Method: Bolger, Davis, & Rafaeli, 2003; Feldman Barrett &

Barrett, 2001; Scollon, Kim-Prieto, & Diener, 2003; Rochester Interaction Record: Reis & Wheeler, 1991) practically and conceptually. Practically, as noted earlier, the EAR operates (a) non-reactively without interrupting the flow of a person's life and (b) imperceptibly without participants being aware of when exactly their behavior is recorded. Compared with traditional beeper methods where, at every prompt, participants must step back from what they are doing to respond to the researchers' questions, the EAR records unobtrusively and requires only minimal involvement on behalf of participants (Mehl & Pennebaker, 2003b).

Conceptually, the crucial distinction between the EAR and other momentary assessment methods lies in the perspective from which the data are gathered; this distinction contrasts the perspective of the agent from the perspective of observers (Hogan, 1982; Spain, Eaton, & Funder, 2000). The agent's perspective is generally a rather privileged one in that the agent (i.e., the self) has exclusive access to the interpretation of events. For example, the agent can report "I asked the man for a dime because I wanted to get to know him." The subjective background on which the behavior is displayed is known only to the agent. The observer's perspective, in contrast, captures the raw molecular behavior without its psychological context. For example, what an observer sees is a person asking a man for a dime. Whereas experience sampling studies track the agent's subjective, experiential account of an event (e.g., Bolger et al., 2003; Pawlik & Buse, 1982), the EAR provides the "outside" perspective of an unobtrusive observer (Mehl & Pennebaker, 2003b). In psychology, this observer's perspective has received comparatively little attention. Given that, in everyday life, people act as observers of other people's behaviors most of the time, the external perspective has high theoretical importance and social relevance (Hogan, 1982).

An additional critical benefit of the EAR lies in its potential to capture not only how people act in and interact with their environment, but also how people select environments in their first place (Buss, 1987). In the natural course of their lives, humans constantly display their idiosyncratic preferences for social contexts by deciding to enter one or refusing to enter another situation. Some people are better one on one; others thrive in groups. Some feel at ease in silence; others need the stimulation of noisy environments. Some go out exclusively with their same-sex peers; others socialize predominantly among the opposite sex. "These preferred environments provide opportunities for personal dispositions to be manifested and reinforced. Once individuals are in their chosen situations, their words and actions are genuine reflections of their personalities" (Ickes, Snyder, & Garcia, 1997, p. 166). The EAR naturally captures these methodologically elusive proactive person-environment interactions.

Taken together, the unobtrusiveness of its recordings, its relative ease of use (compared to other field observation methods), its "unprivileged" observer's perspective, and its ability to capture the effects of environment selection, make the EAR a valuable tool for studying personality in everyday life and a method well suited for addressing our three basic questions about everyday manifestations and implicit folk theories of personality.

Research Questions

Given the paucity of naturalistic observation research in personality psychology, this study sought to provide an in-depth analysis

of the personality implications of people's naturally occurring daily lives. In its descriptive, exploratory, and phenomenon-centered approach, it is rooted in a scientific paradigm that is inductive and discovery based rather than deductive and theory testing in nature (Furr, 2002; Rozin, 2001). Specifically, the study sought to address three basic research questions: (1) How are people's personalities manifested in their daily lives? (2) How are people's implicit folk theories of personality based on elements of daily life? (3) How accurate are people's implicit folk theories of personality?

Question 1: How Are People's Personalities Manifested in Their Daily Lives?

Funder (2001) recently observed that "after three-quarters of a century of research on traits the catalog of basic facts concerning the relationships between personality and behavior remains thin" (p. 212). Although everyday manifestations of personality have recently received increased attention (Chaplin, Phillips, Brown, Clanton, & Stein, 2000; Eaton & Funder, 2001; Oberlander & Gill, in press; Paunonen, 2003; Rentfrow & Gosling, 2003; Spain, Eaton, & Funder, 2000; Wu & Clark, 2003), there continues to be an urgent need for observational research that links aspects of people's personalities to their daily behaviors, environments, and interactions.

With their recent analysis of students' daily social environment and natural conversations, Mehl and Pennebaker (2003b) laid the methodological and conceptual foundation for such a project. Fifty-two undergraduate students wore the EAR twice for 2 days separated by 4 weeks. Information about students' moment-to-moment behaviors (e.g., listening to music, watching TV), social interactions (e.g., alone, talking to others), locations (at home, outside, in public places), moods (laughing, crying), and language use (e.g., emotion words, swear words, fillers) was derived from the sampled ambient sounds via behavioral coding and linguistic analyses. The psychometric analyses revealed substantial individual differences in students' daily lives, as well as a high degree of stability of these individual differences over a period of 4 weeks. The present project aimed to build on these purely psychometric analyses by providing psychological meaning to such reliable individual differences in daily life in a new and larger sample.

Eaton and Funder (2001) recently showed that the psychological processes underlying people's everyday emotional experiences were partially gender specific. Following their call to overcome the common practice of routinely aggregating data across gender, we decided to analyze the daily behaviors, social environments, and conversations of male and female students separately and together. On the basis of the EAR's potential to capture subtle differences in how male and female students habitually select social environments (Stewart & McDermott, 2004; Wood & Eagly, 2002), we expected manifestations of personality to be partially gender differentiated.

Question 2: How Are People's Implicit Folk Theories of Personality Based on Elements of Daily Life?

People's daily lives are the natural platforms on which social identities, public impressions, and reputations are negotiated (Goffman, 1959; Hogan, 1982). Every day, people witness other

people going about their lives. Naturally, they form impressions of others on the basis of the environments that they encounter them in and the behaviors that they observe. It is the connection between what is observed and what is intuitively inferred that reveals people's implicit folk theories of personality. What do we think of a person who spends most of the day at home? What implications does it have for how we see a person when we realize the person predominantly interacts with the opposite sex? How does it shape our impressions when we find that someone constantly uses profanity? The EAR provides an opportunity to study how observable features of daily life are associated with impressions that people develop of others. Therefore, we investigated the quotidian content of implicit folk theories of personality by having naïve judges form personality impressions on the basis of participants' (acoustic) daily lives.

Given the lack of prior research on this question, we made no prediction regarding differences in implicit folk theories of how personality is manifested in the daily lives of male and female students. The existence of gender stereotypes suggests that people's implicit folk theories would be broadly based on or consistent with perceived gender-linked roles and activities (Lippa, 1978; Lippa & Connolly, 1990; Madon et al., 1998; Swim, 1994).

Question 3: How Accurate Are People's Implicit Folk Theories of Personality?

The fact that people intuitively associate certain behaviors with certain aspects of personality (e.g., talking with Extraversion; swearing with a lack of Agreeableness), however, does not require that these implicit folk theories are accurate. There is evidence for the validity of some implicit folk theories (Lee, Jussim, & McCauley, 1995; Madon et al., 1998; Swim, 1994), but others may not be grounded in reality. Our goal here was to investigate the extent to which people's implicit folk theories about the relationship between personality and everyday behavior are accurate—that is, do people's intuitive theories about the behavioral manifestations of personality really correspond to how personality is manifested in the real world? Do they miss some existing behavior–personality links, and do they perceive some links that do not exist? And finally, do people hold correctly differentiated implicit theories for men and women?

Design of the Present Research

The three research questions required that we collect three sources of data: Information about what people are like, information about what these people do in their daily lives, and information about how these people are seen by others (on the basis of what they do in their daily lives). We collected these three sources of data in a single research project consisting of three major phases. In Phase 1, 96 target participants completed a battery of personality measures. In Phase 2, these participants wore the EAR for 2 consecutive days, and their moment-to-moment ambient sounds were coded, transcribed, and analyzed for aspects of their daily interactions, locations, activities, moods, and language use. In Phase 3, the EAR sound files were presented to a group of naïve judges who provided their impressions of the target participants' personalities.

Method

Target Participants

Ninety-six introductory psychology students at the University of Texas at Austin (49 male, 47 female; 92% college freshmen or sophomores, mean age = 18.7, $SD = 0.9$) served as target participants for the study. Participation was voluntary and for course credit.¹

Phase 1: Personality Measures

On arrival on the first day of the study, the participants were informed about the purpose of the project. Specifically, they were told that the study's goal was to use a newly developed scientific method to investigate the psychological implications of the daily lives of college students. After that, they signed the consent form and completed a large battery of questionnaires that included measures of their personalities. Our research questions here focused on the relationship between personality and daily life at a conceptually broad level. In addition, we wanted to maximize comparability with other research in this area (e.g., Borkenau et al., 2004; Funder & Sneed, 1993; Paunonen, 2003). Therefore, we assessed personality at the global level of the Big Five personality dimensions, using the 44-item Big Five Inventory (BFI; John & Srivastava, 1999). The participants rated themselves on the BFI items using a scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). The alpha reliabilities for Extraversion, Agreeableness, Conscientiousness, Emotional Stability, and Openness to Experience were .90, .77, .83, .87, and .80, respectively.

Phase 2: EAR Monitoring

EAR System. The EAR system consisted of a digital voice recorder (SONY Memory Stick ICD-MS1), an external microphone (OPTIMUS Tie Clip Microphone), and a controller microchip. The chip was programmed with a 30-s on and 12.5-min off cycle producing 4.8 sound recordings per hour. Participants carried the EAR around in a shock-protected case in one of their pockets or in a purse-like bag. An external microphone was clipped to the collar of their shirts. It was impossible for participants to sense when the EAR was recording.²

Procedure. Participants wore the EAR continuously for 2 consecutive weekdays during their waking hours. Monitoring sessions were scheduled from Monday morning to Wednesday morning ($n = 22$) and from Wednesday afternoon to Friday afternoon ($n = 74$). Participants were run in groups of 2 to 4. After having completed the questionnaire battery, they were thoroughly informed about the EAR procedure, including the fact that there would be an opportunity to review the recordings and erase any part they did not want the researchers to hear. Participants were encouraged to wear the EAR as much as possible and to remove it only when its functioning was in jeopardy. To increase the reliability of the analyses, we also asked participants to keep an end-of-day diary where they documented major activities of the day and times where they were unable to wear the EAR. The recorder was then activated and handed to participants in a sealed case.

When the participants returned to the lab 2 days later, they were given the opportunity to review their recordings and erase parts they did not want the researchers to hear. Nineteen participants (19.8%) reviewed their recordings, but only 3 students erased sound files (10 in total; <0.01%). Participants then completed a standard eight-item method evaluation questionnaire assessing their compliance with wearing the EAR as well as the device's obtrusiveness for themselves and the people around them.

Compliance. Across all 96 participants, the EAR recorded a total of 22,142 sound files. 7,695 of these files (35%) were coded as sleeping yielding 14,447 waking recordings. On average, 14% ($SD = 17\%$) of participants' waking recordings were coded as invalid and dropped from the analysis (not wearing the EAR: $M = 9\%$, $SD = 14\%$; insufficient ambient sounds: $M = 4\%$, $SD = 14\%$; bad recording quality: $M = 1\%$, $SD = 4\%$). Over the 2-day monitoring period, then, the EAR provided on

average 130 valid waking recordings ($SD = 34$) per participant reflecting a high compliance rate. The high compliance rate was consistent with participants' self-reports; participants reported having worn the EAR more than three quarters of their time awake ($M = 78\%$, $SD = 16\%$). The subjective estimates were strongly correlated with the number of valid sound files ($r = .57$, $p < .01$).

Obtrusiveness. Awareness of the EAR was rated on a 5-point unipolar scale (5 = *a great deal*), with moderate ratings for the participants themselves ($M = 2.91$, $SD = 1.01$) and the people around them ($M = 3.25$, $SD = 1.05$). Participants reported feeling only slightly constrained by the device ($M = 1.85$, $SD = 0.92$) and uncomfortable wearing it ($M = 1.96$, $SD = 1.04$). They further indicated that the recorder had only minor effects on their behavior ($M = 1.63$, $SD = 0.83$) and talking ($M = 1.53$, $SD = 0.87$), as well as the behavior of people around them ($M = 2.05$, $SD = 1.14$). Finally, research assistants coded how often the EAR was mentioned with others. On average, participants referred to the EAR in only 3.6% ($SD = 4.0\%$) of their daily conversations.

Coding of EAR sounds. A team of 8 research assistants was trained to code acoustically detectable features of participants' moment-to-moment behaviors, social environments, and conversations. The research assistants listened to all of a participant's EAR recordings and coded each sound file using a revised version of the Social Environment Coding of Sound Inventory (SECSI; Mehl & Pennebaker, 2003a). The coding system comprises four major category clusters: (a) the person's current location (e.g., in apartment, outdoors, in transit), (b) activity (e.g., listening to music, on the computer, eating), (c) interaction (e.g., alone, on the phone, talking in person) and (d) mood (e.g., laughing, crying, sighing). Intercoder reliabilities for the 32 SECSI categories were determined from a set of training EAR recordings (392 sound files) independently coded by all eight research assistants. Intraclass correlations (ICCs) based on a two-way random effects model (ICC[2,k]; Shrout & Fleiss, 1979) exceeded .70 for all categories except reading. Because of the low intercoder reliability ($r = .12$), reading was dropped from the analyses. Then we converted the raw codings into time-use estimates by calculating the percentage of a person's valid (i.e., compliant and codable) waking EAR recordings in which a SECSI category applied (e.g., percentage of sound files in which the participant was talking, at home, or listening to music). To capture preferences for specific types of interactions, the time-use estimates for participants' dyadic, group, phone, and self-talk, as well as participants' same-sex versus opposite-sex conversations were converted into relative time-use estimates by expressing them as proportions of participants' total amount of conversations.

Transcription and linguistic analysis. In addition to the sound coding, research assistants also transcribed all of the participants' utterances captured by the EAR. They received special training for how to handle ambiguities such as repetitions, filler words, nonfluencies, or slang. The transcripts were then submitted to a linguistic analysis using Linguistic Inquiry and Word Count (LIWC; Pennebaker, Francis, & Booth, 2001), an extensively validated text analysis program that operates by comparing all words of a given text to an internal dictionary consisting of more than 2,300 words. The words in the LIWC2001 default dictionary, which was used for this study, are arranged into 74 grammatical (e.g., first-person singular pronoun, article, preposition) and psychological (e.g., positive emotions, negative emotions, cognitive processes, social process) language-use dimensions. In addition to the total word count, LIWC

¹ Originally, the EAR was given to 110 students. Because of technical problems (unreliable triggering, insufficient battery power, microphone malfunctioning) the data of 14 participants (8 male, 6 female) could not be used for the analyses.

² Further information on the EAR, including a discussion of the ethics of the method and our privacy and confidentiality policies, is provided by Mehl and Pennebaker (2003b) and Mehl et al. (2001).

reports the percentage of words that fell into each of the 74 categories (e.g., the percentage of participants' words captured over a 2-day period that were pronouns, articles, or emotion words).

For this study, we limited the analysis to 23 LIWC variables that had previously been identified as reliable and relevant to the study of natural conversations (Mehl & Pennebaker, 2003a). Intertranscriber agreement was calculated from LIWC analyses of the transcripts of 76 conversations (>5 words) contained in the set of training recordings. ICCs based on a two-way random effects model (ICC[2,k]) with 8 transcribers were $\geq .85$ for all 23 language-use categories.³

Phase 3: Judges' Personality Impressions

A new team of 18 undergraduate research assistants (7 male, 11 female) served as naïve personality judges. Judges were divided into three groups of 5, 6, and 7, and each group was assigned a third of the sample. This way, 33 target participants were rated by 5 judges, 31 by 6, and 32 by 7 judges.⁴ We ensured that the judges did not know any of the participants they rated. Judges were instructed to listen to the entire set of a participant's EAR recordings before recording their personality impressions. They were allowed to fast forward only major periods of silence (e.g., the nights or periods of noncompliance). The order in which participants were rated was counterbalanced within the three groups of judges.

Judges recorded their impressions of the target participants' personalities on the BFI using a 7-point scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). The alpha reliabilities for Extraversion, Agreeableness, Conscientiousness, Emotional Stability, and Openness to Experience averaged .93, .86, .92, .88, .87, respectively across the ratings of the 18 judges. Interjudge reliabilities, calculated as ICCs based on one-way random effect models (ICC[1,k]), were strong across all Big Five dimensions (mean $r = .84$, $p < .01$). In order to remove idiosyncrasies in scale use, personality impressions were standardized within judge and aggregated across all judges.

Data Analytic Strategy

As noted earlier, the three research questions are most usefully conceptualized as a set, with the third one bringing the first two together. Therefore, we present the data on all three questions within each of the five personality dimensions (rather than addressing the five traits within each of the three questions). This organizational system allows us to highlight the most relevant findings for each personality domain. For each Big Five dimension, a series of correlational analyses were conducted to test Question 1 (Q1), the extent to which aspects of participants' daily lives were reliable manifestations of their personalities, and Question 2 (Q2), the extent to which judges' implicit folk theories of personality were associated with aspects of participants' daily lives. Tables 1 to 5 (shown later) summarize the findings of these analyses separately for each of the Big Five.

The correlations on the left-hand side of the tables reflect the relationships between participants' personalities and the SECSI and LIWC categories; they indicate the extent to which observable aspects of students' daily lives were associated with their personalities. The correlations on the right-hand side of the tables reflect the relationships between the aggregated judges' impressions and the SECSI and LIWC variables; the correlations indicate the extent to which judges' impressions were associated with observable aspects of students' daily lives. In addition to providing the findings for the entire sample, Tables 1 to 5 also show both sets of correlations separately for male and female participants.

Given the exploratory focus of the study, an empirically based strategy was adopted to guard against experimental-wise error and to preserve moderate levels of power for the analyses. Regardless of their theoretical relevance to the Big Five domains, EAR-derived variables were included in Tables 1–5 if one of the correlations in the full sample or the subsamples

of male and female participants was equal to or exceeded $l.30l$. The purpose of this strategy was to balance alpha and beta error and to ensure that the results focus on effects that are substantial in magnitude (Cohen, 1988; Hemphill, 2003). Similarly, in Tables 1–5, correlations that differed by $l.30l$ or more for men and women are marked in bold and point to variables with at least moderate levels of gender specificity.⁵

We used two methods to examine Question 3 (Q3), the accuracy of people's implicit folk theories of how personality is manifested in daily life. First, we calculated column-vector correlations to obtain a formal measure of the degree of convergence between the patterns for the behavioral-manifestation correlations and the implicit-folk-theories correlations (cf. Funder & Sneed, 1993). For this question, it is the similarity in patterns of the correlations that is important, not their magnitude. Therefore, we based the column-vector correlations on the (Fisher's r -to- z transformed) correlations for the full set of 31 SECSI and 23 LIWC cues, not just the cues with correlations above $l.30l$ depicted in Tables 1–5. Second, to account for the fact that practical limitations prevented the coding of some potentially relevant cues (e.g., paralinguistic information), we also computed a global measure of perceiver accuracy that is independent of the cues that were actually measured; this broader index of self-judge agreement was the correlation between the aggregated judges' impressions and participants' actual personalities.

Results

Extraversion

Q1: Behavioral manifestations. The left-hand side of Table 1 shows how Extraversion was manifested in participants' daily interactions, locations, activities, mood, and language use. According to the Big Five definition (Costa & McCrae, 1992; Goldberg, 1990), extraverted individuals are sociable, talkative, and enthusiastic. Consistent with this definition, correlations between participants' Extraversion and the EAR-derived SECSI and LIWC variables (data column 1) revealed that Extraversion was predominantly expressed in students' daily interactions and language use. For example, our exploratory analyses indicated that relative to introverts, extraverted participants were captured in more conversations and spent less time alone. As would be expected, extraverts also uttered more words over the 2-day period. Interesting to note, the manifestations of Extraversion showed theoretically important gender differences (data columns 2 and 3). For example, group conversations were indicative of Extraversion only among women and talking to oneself was indicative of Introversion only among men. Gender differences also emerged for the relative amount of mixed-sex conversations and the use of past-tense verbs.

Q2: Judges' implicit folk theories. The right-hand side of Table 1 shows which cues were associated with judges' impres-

³ To preserve space, only limited descriptive information is provided for the SECSI and LIWC variables. The full descriptive statistics including the reliabilities, base rates, standard deviations, and gender effects for each variable, as well as the intercorrelations among the SECSI and LIWC variables are available from Matthias R. Mehl.

⁴ The unequal number of judges in the three groups is a result of an error in the assignment process. One judge was supposed to receive the sound files of the participants for the first group of judges but accidentally received the sound files of participants for the third group.

⁵ The full sets of correlations for all EAR-derived variables are available from Matthias R. Mehl.

Table 1
Behavioral Manifestations of Extraversion in Students' Daily Lives and Judges' Implicit Folk Theories of How Extraversion Is Manifested in Daily Life

EAR-derived variables	Extraversion					
	Question 1: Behavioral manifestation			Question 2: Implicit folk theories		
	N = 96	Males	Females	N = 96	Males	Females
Interactions						
Alone	-.27**	-.33*	-.42**	-.67**	-.73**	-.70**
Talking	.30**	.33*	.37**	.67**	.72**	.64**
% Dyadic conversations	-.07	.04	-.31*	-.11	-.14	-.09
% Group conversations	.12	.07	.49**	.29**	.31*	.35*
% Phone conversations	-.01	-.07	-.13	-.19	-.32*	-.17
% Self-talk	-.23*	-.38**	-.03	-.21*	-.16	-.25
% Same-sex group conversations	.01	.01	.30*	.27**	.30*	.35*
% Opposite-sex dyadic conversations	.09	-.07	.17	-.12	-.32*	.01
% Mixed-sex conversations	.11	.06	.37**	.02	-.03	.14
Locations						
In restaurant/bar/coffee shop	.14	.12	.17	.18	.39*	.05
Activities						
Socializing	.16	.15	.26	.30**	.32*	.29*
Mood						
Laughing	.07	.06	.01	.42**	.48**	.36*
Arguing	.17	.20	.15	.21*	.30*	.16
Language use						
Sampled raw word count	.29**	.32*	.30*	.63**	.64**	.61**
Words of more than 6 letters	-.20*	-.32*	-.08	-.04	-.25	.19
Swear words	-.08	-.01	.24	.30**	.46**	.24
Positive emotion words	.05	.16	-.10	.20*	.32*	.12
Negative emotion words	.03	.03	.31*	.36**	.45**	.34*
Insight words	-.06	-.31*	.14	.04	-.18	.24
Past tense verbs	-.10	.10	-.35*	-.19	-.01	-.36*
Present tense verbs	.17	.12	.33*	.14	.03	.25
Inclusive words	-.04	-.16	-.02	-.12	-.36**	.11

Note. Behavioral manifestations are based on participants' self-reported Extraversion; implicit folk theories are based on judges' ratings of participants' Extraversion; the EAR-derived variables are computed as proportions of the total number of sampled sound files (Interactions, Locations, Activities, Mood) or words (Language use; except: sampled raw word count); %-variables are computed as proportions of participants' total amount of talking; all variables with at least one correlation $\geq |.30|$ are displayed; numbers in bold indicate gender differences of $\Delta r \geq |.30|$; $n_{\text{male}} = 49$; $n_{\text{female}} = 47$. EAR = Electronically Activated Recorder.
 * $p \leq .05$ (two-tailed). ** $p \leq .01$ (two-tailed).

sions of participants' Extraversion. The correlations for the entire sample (data column 4) suggest that judges based their impressions predominantly on aspects of targets' daily interactions and language use. For example, the degree of talking captured by the EAR was strongly correlated with ratings of Extraversion. Participants further came across as extraverted if they uttered many words over the 2-day period and used swear words frequently. Consistent with the idea that extraverts are sociable and express positive emotions, socializing and laughing also led judges to perceive participants as extraverted. In line with the gender differences in behavioral manifestations, judges' implicit folk theories also emerged as partially gender specific (data columns 5 and 6). Only among men, for example, was time spent in restaurants, bars, or coffee shops considered to be an indicator of Extraversion. These gender differences, however, did not always map closely onto the gender differences in actual manifestations of Extraversion.

Q3: Accuracy of implicit folk theories. As shown in Table 6, the strong column-vector correlation suggests that, across all participants, judges' cue use corresponded well to the actual mani-

festations of Extraversion ($r = .58$). This finding was supported by the strong overall agreement between the judges' impressions of participants and their actual levels of Extraversion ($r = .41$). Despite the fact that the gender differences did not always converge between the behavioral manifestations and the implicit folk theories of Extraversion, high column-vector correlations emerged for both male ($r = .56$) and female ($r = .60$) participants' levels of Extraversion. Similarly, self-judge agreement was strong for both male ($r = .38$) and female ($r = .44$) participants.

Agreeableness

Q1: Behavioral manifestations. Agreeable individuals tend to express sympathy and warmth, and are generally unfrontational. As shown in the left-hand side of Table 2 Agreeableness was most strongly expressed in students' language use (data column 1). Consistent with the trait definition, Agreeableness was negatively related to the use of swear words and positively correlated with the use of first-person singular pronouns (I, me, my)—a linguistic device known

Table 2
Behavioral Manifestations of Agreeableness in Students' Daily Lives and Judges' Implicit Folk Theories of How Agreeableness Is Manifested in Daily Life

EAR-derived variables	Agreeableness					
	Question 1: Behavioral manifestation			Question 2: Implicit folk theories		
	N = 96	Males	Females	N = 96	Males	Females
Interactions						
% Dyadic conversations	.00	-.01	-.06	.16	.01	.32*
% Phone conversations	.08	.02	.00	-.24*	-.20	-.39**
Locations						
Inside apartment	-.33**	-.37**	-.31*	-.29**	-.19	-.40**
In restaurant/bar/coffee shop	.06	.27	-.10	.25*	.37**	.18
In other public places	.23*	.29*	.20	.25*	.16	.39**
Activities						
Listening to music	-.21*	-.08	-.33*	-.13	-.12	-.12
Watching TV	-.15	-.20	-.02	-.27**	-.21	-.31*
In class	.15	.18	.05	.33**	.26	.39**
Mood						
Arguing	-.09	.19	-.32*	-.19	-.13	-.27
Language use						
Words of more than 6 letters	-.01	.04	-.08	.25*	.34*	.14
First-person singular pronouns	.23*	.16	.13	-.20*	-.22	-.38**
Negations	-.17	-.30*	-.01	-.18	-.15	-.20
Articles	-.23*	-.30*	-.03	.08	.07	.18
Swear words	-.28**	-.19	-.24	-.51**	-.62**	-.37**
Negative emotion words	-.16	-.11	-.07	-.44**	-.56**	-.19
Insight words	.03	.00	-.01	.34**	.30*	.36*
Past tense verbs	-.05	-.07	-.14	-.25*	-.19	-.37**

Note. Behavioral manifestations are based on participants' self-reported Agreeableness; implicit folk theories are based on judges' ratings of participants' Agreeableness; the EAR-derived variables are computed as proportions of the total number of sampled sound files (Interactions, Locations, Activities, Mood) or words (Language use); %-variables are computed as proportions of participants' total amount of talking; all variables with at least one correlation $\geq |.30|$ are displayed; numbers in bold indicate gender differences of $\Delta r \geq |.30|$; $n_{\text{male}} = 49$; $n_{\text{female}} = 47$. EAR = Electronically Activated Recorder.
 * $p \leq .05$ (two-tailed). ** $p \leq .01$ (two-tailed).

to convey personal rapport (Pennebaker, Mehl, & Niederhoffer, 2003). Agreeableness was further expressed in students' daily locations; agreeable participants spent less time at home than disagreeable participants. The separate analyses for male and female students revealed only a few gender-specific manifestations of Agreeableness (data columns 2 and 3). For example, the amount of arguing captured by the EAR was indicative of a lack of Agreeableness among women but not among men, and the amount of time spent in a restaurant, bar, or coffee shop was differentially related to Agreeableness for males and females.

Q2: Judges' implicit folk theories. As shown on the right-hand side of Table 2, the cue-use correlations for the entire sample (data column 4) suggest that, consistent with the trait definition and the identified behavioral manifestations, judges' impressions were strongly shaped by students' language use. Participants who refrained from using swear words and other negative emotion words came across as agreeable. Also, class attendance, watching little TV, and spending time in public places such as restaurants, bars, and coffee shops were related to impressions of Agreeableness. Consistent with the low degree of gender differentiation in the ways Agreeableness was manifested, judges' cue use was also similar across gender (data columns 5 and 6). Only the use of negative emotion words and a preference for dyadic interactions

was differentially associated with perceptions of Agreeableness for men and women.

Q3: Accuracy of implicit folk theories. Across the entire sample, judges' implicit folk theories matched the actual manifestations of Agreeableness fairly well (shown in Table 6, data column 1). In line with this finding, self-judge agreement was also moderately strong across all participants (data column 4). Judges were equally effective for male and female participants at basing their impressions on the correct cues (data columns 2 and 4). Self-judge agreement was also similar for male and female targets (data columns 5 and 6).

Conscientiousness

Q1: Behavioral manifestations. Conscientiousness is associated with behaving responsibly, carefully, and with self-discipline. In line with this conceptualization, Conscientiousness was most strongly expressed in participants' daily activities and language use (see Table 3). For example, across all participants, Conscientiousness was positively related to the time spent in class and negatively related to the use of swear words. Furthermore, the more conscientious students were, the more time they spent in public places other than restaurants, bars, or coffee shops (e.g., on

Table 3
Behavioral Manifestation of Conscientiousness in Students' Daily Lives and Judges' Implicit Folk Theories of How Conscientiousness Is Manifested in Daily Life

EAR-derived variables	Conscientiousness					
	Question 1: Behavioral manifestation			Question 2: Implicit folk theories		
	N = 96	Males	Females	N = 96	Males	Females
Interactions						
% Phone conversations	.07	.04	.03	-.13	-.08	-.35*
Location						
Inside apartment	-.27**	-.33*	-.23	-.36**	-.38**	-.38**
In other public places	.36**	.41**	.33*	.44**	.47**	.45**
Activities						
Watching TV	.00	-.07	.10	-.41**	-.33*	-.53**
Working at a job	-.17	-.01	-.31*	.12	.16	.06
In class	.42**	.30*	.52**	.46**	.52**	.35*
Entertainment	-.13	-.07	-.21	-.35**	-.34*	-.29*
Mood						
Sighing	-.18	.00	-.32*	.02	.16	-.15
Language use						
Words of more than 6 letters	.19	.17	.21	.30**	.34*	.25
Second person pronouns	-.05	.34*	-.42**	-.03	.15	-.28
Swear words	-.29**	-.29*	-.34*	-.61**	-.70**	-.34*
Nonfluencies	-.02	.03	-.05	.24*	.32*	.17
Filler words	.18	.14	.15	.20*	.35*	-.10
Negative emotion words	-.25*	-.19	-.31*	-.49**	-.61**	-.07
Insight words	.01	-.06	.03	.29**	.27	.30*
Exclusive words	.19	.02	.38**	.05	.00	.21

Note. Behavioral manifestations are based on participants' self-reported Conscientiousness; implicit folk theories are based on judges' ratings of participants' Conscientiousness; the EAR-derived variables are computed as proportions of the total number of sampled sound files (Interactions, Locations, Activities, Mood) or words (Language use); %-variables are computed as proportions of participants' total amount of talking; all variables with at least one correlation $\geq |.30|$ are displayed; numbers in bold indicate gender differences of $\Delta r \geq |.30|$; $n_{\text{male}} = 49$; $n_{\text{female}} = 47$. EAR = Electronically Activated Recorder.
 * $p \leq .05$ (two-tailed). ** $p \leq .01$ (two-tailed).

campus). Again, however, some manifestations of Conscientiousness differed for male and female students. For example, our exploratory analyses revealed that the amount of time spent working at a job and the amount of captured sighing were indicative of a lack of Conscientiousness among women only. Further, the use of second person pronouns (i.e., you, your) was positively related to Conscientiousness in men and negatively in women.

Q2: Judges' implicit folk theories. The cue-use analyses yielded a rather clear picture with regard to the perceptual paths underlying judges' impressions of participants' Conscientiousness. Judges correctly derived their impressions predominantly from participants' activities and language use. The less TV and entertainment exposure students had and the more time they spent in class, the more conscientious they were rated. Also, a carefully reflected word choice as indicated by a frequent use of long words, insight words (e.g., realize, understand), and a low use of swear words and other negative emotion words was related to judgments of high Conscientiousness. Whereas judges showed no gender differentiation with regard to the activities they considered indicative of Conscientiousness, gender differences did emerge for their use of language cues. For example, the use of negative emotion words in general and the use of swear words in particular, as well as the use of filler words (e.g., like, well) shaped judges' impressions of targets' Conscientiousness more for males than for fe-

males. Also, judges had an intuitive grasp of the fact that the use of second person pronouns (i.e., you, your) was positively related to Conscientiousness in men and negatively in women. This latter finding is interesting because it reveals that judges can implicitly hold assumptions about the personality implications of daily behavior that they are unlikely to report explicitly.

Q3: Accuracy of implicit folk theories. Across the entire sample, there was a strong correspondence between the ways Conscientiousness was manifested in behavior and judges' implicit folk theories (shown in Table 6). Consistent with this pattern, there was also substantial self-judge agreement for this dimension. The column-vector correlations as well as the self-judge agreement were similar in magnitude for male and female participants indicating that judges' implicit folk theories were equally accurate for both genders.

Emotional Stability

Q1: Behavioral manifestations. Across the entire sample, very few reliable manifestations of Emotional Stability emerged (see Table 4). Emotionally stable individuals are characterized as being calm, not easily upset, and low in self-consciousness. Consistent with this definition, Emotional Stability was significantly related to less time spent arguing. Emotional Stability was further related

Table 4
Behavioral Manifestation of Emotional Stability in Students' Daily Lives and Judges' Implicit Folk Theories of How Emotional Stability Is Manifested in Daily Life

EAR-derived variables	Emotional Stability					
	Question 1: Behavioral manifestation			Question 2: Implicit folk theories		
	N = 96	Males	Females	N = 96	Males	Females
Interactions						
Alone	-.16	.10	-.25	-.46**	-.51**	-.25
Talking	.14	-.03	.22	.28**	.44**	.03
% Dyadic conversations	.07	.17	.08	.04	-.09	.35**
% Group conversations	.06	-.17	.08	.35**	.25	.16
% Phone conversations	-.19	.07	-.18	-.51**	-.28*	-.49**
% Same-sex group conversations	.00	-.29*	.15	.30**	.27	.03
% Dyadic opposite-sex conversations	-.03	.09	-.02	-.23*	-.33*	-.09
Locations						
Outdoors	-.15	-.37**	.00	.17	.24	.14
In restaurant/bar/coffee shop	.01	.04	.01	.16	.30*	.16
In other public places	.16	.32*	-.02	-.01	-.17	.05
Activities						
Socializing	-.04	-.30*	.16	.29**	.40**	.15
Mood						
Laughing	.09	-.13	.35*	.24*	.39**	.34*
Crying	-.06	—	.05	-.40**	—	-.39**
Arguing	-.22*	-.02	-.30*	-.31**	.18	-.54**
Sighing	.15	-.08	.30*	.16	.20	.12
Language use						
Sampled raw word count	.22*	.05	.36*	.28**	.46**	.17
First-person singular pronouns	-.16	.00	-.07	-.41**	-.26	-.23
Tentative words	.04	.10	-.01	.15	.02	.34*

Note. Behavioral manifestations are based on participants' self-reported Emotional Stability; implicit folk theories are based on judges' ratings of participants' Emotional Stability after; the EAR-derived variables are computed as proportions of the total number of sampled sound files (Interactions, Locations, Activities, Mood) or words (Language use; except: sampled raw word count); %-variables are computed as proportions of participants' total amount of talking; all variables with at least one correlation $\geq |.30|$ are displayed; numbers in bold indicate gender differences of $\Delta r \geq |.30|$; $n_{\text{male}} = 49$; $n_{\text{female}} = 47$. EAR = Electronically Activated Recorder.

* $p \leq .05$ (two-tailed). ** $p \leq .01$ (two-tailed).

to a higher number of uttered words. Both of these correlations showed signs of gender differentiation. Overall, there was little overlap in the patterns for female and male manifestations of Emotional Stability. Only among women, for example, were laughing, sighing, and not arguing indicative of Emotional Stability. Also, only among men did socializing, particularly in same-sex groups, and spending time outdoors suggest a lack of Emotional Stability. Finally, the time spent in public places other than restaurants, bars, and coffee shops, was indicative of Emotional Stability only in male participants.

Q2: Judges' implicit folk theories. The right-hand side of Table 4 shows that, when forming their impressions of participants' Emotional Stability, judges primarily attended to information about targets' interactions and moods. For example, across all participants, judgments of Emotional Stability were negatively related to the time spent alone and on the phone and positively correlated with the amount of time spent talking in same-sex groups. Also, the more targets socialized, the more emotionally stable they appeared. In line with the pronounced gender differences in the behavioral manifestations, Emotional Stability showed a strong influence of gender on judges' cue use. For example,

whereas the total amount of talking captured by the EAR was considered indicative of Emotional Stability only for male participants, dyadic conversations were considered indicative of Emotional Stability only among female participants. Also, only female and not male participants came across as lacking Emotional Stability to the extent that they were frequently captured in arguments.

Q3: Accuracy of implicit folk theories. Across the entire sample, there was a strong correspondence between judges' cue use and the actual behavioral manifestations of Emotional Stability in participants' daily lives (shown in Table 6). Self-judge agreement was also strong when calculated across the entire sample. It is interesting to note, however, that when calculated separately for male and female participants, a strong column-vector correlation and strong self-judge agreement emerged only among female students. For male participants, judges' ratings were no more accurate than chance levels. Further, judges' pattern of cue use for men was negatively correlated with the actual pattern of behavioral manifestations. Considering that judges' cue use for Emotional Stability was substantially gender differentiated, this suggests that judges based their ratings (a) on largely accurate stereotypes about gender differences in Emotional Stability and (b) on stereotypes of

everyday manifestations of Emotional Stability that were rather accurate for women but inaccurate for men.

Openness to Experience

Q1: Behavioral manifestations. The behavioral manifestations of Openness to Experience were less consistent with the trait definition than was the case for the other Big Five dimensions. Openness to Experience is associated with being curious, imaginative, and having a wide range of interests. As shown in Table 5, Openness to Experience was predominantly manifested in participants' language use. Across the entire sample, Openness was, for example, negatively correlated with the use of third-person pronouns (e.g., he, she, they), and past tense verbs. Furthermore, the more open participants were, the more time they spent in restaurants, bars, and coffee shops. Based on our criterion of $\Delta r \geq .130$, no gender-specific manifestations of Openness in students' daily lives emerged.

Q2: Judges' implicit folk theories. The analyses for the entire sample revealed that participants' language use and locations were considered particularly indicative of their levels of Openness (see Table 5). For example, Openness was inferred from spending time in public places as compared with at home. Also, an elaborate use of language as indicated by the frequent use of long words, insight

words, and tentative words elicited impressions of Openness. Consistent with the absence of gender differences in behavioral manifestations, few gender differences emerged in the implicit folk theories of how Openness was expressed.

Q3: Accuracy of implicit folk theories. In line with the theoretically inconsistent findings regarding expressions of Openness to Experience, both the correspondence between judges' implicit folk theories and the actual trait manifestations and overall self-judge agreement were lower for Openness than for any of the other Big Five dimensions (see Table 6). When calculated separately for both genders, judges had a moderately accurate intuitive grasp of everyday manifestations of Openness for female but not for male participants.

Discussion

In the tradition pioneered by Barker and Wright (1951), this study sought to provide an in-depth exploratory investigation into the personality implications of daily life. It identified a number of concrete, molecular, and observable ways in which participants' personalities were manifested in their daily social interactions, locations, activities, moods, and language use. It further revealed a detailed portrait of how people's implicit folk theories of personality are based on observable aspects of daily life. Finally, by

Table 5
Behavioral Manifestation of Openness to Experience in Students' Daily Lives and Judges' Implicit Folk Theories of How Openness to Experience Is Manifested in Daily Life

EAR-derived variables	Openness to Experience					
	Question 1: Behavioral manifestation			Question 2: Implicit folk theories		
	N = 96	Males	Females	N = 96	Males	Females
Interactions						
% Phone conversations	-.13	-.09	-.13	-.27**	-.22	-.32*
Locations						
Inside apartment	-.06	-.08	-.04	-.37**	-.26	-.48**
In restaurant/bar/coffee shop	.24*	.31*	.21	.27**	.26	.30*
In other public places	.07	.08	.05	.31**	.19	.45**
Activities						
Watching TV	-.09	-.09	-.11	-.42**	-.36**	-.51**
In class	-.03	.12	-.15	.26**	.24	.30*
Language use						
Words of more than six letters	.03	-.07	.12	.24*	.32*	.14
Third-person pronouns	-.28**	-.14	-.36*	-.17	-.08	-.23
Prepositions	-.08	-.13	-.05	.17	-.01	.34*
Swear words	.06	.02	.05	-.17	-.32*	-.03
Insight words	.05	.02	.09	.32**	.41**	.26
Discrepancy words	-.09	-.21	.01	.23*	.09	.37**
Tentative words	.05	.16	-.07	.30**	.35*	.25
Social processes	-.21*	-.06	-.29*	-.31**	-.26	-.35*
Past tense verbs	-.26**	-.19	-.30*	-.31**	-.34*	-.32*
Present tense verbs	.10	.11	.08	.16	-.02	.30*

Note. Behavioral manifestations are based on participants' self-reported Openness to Experience; implicit folk theories are based on judges' ratings of participants' Openness to Experience; the EAR-derived variables are computed as proportions of the total number of sampled sound files (Interactions, Locations, Activities) or words (Language use); %-variables are computed as proportions of participants' total amount of talking; all variables with at least one correlation $\geq .130$ are displayed; numbers in bold indicate gender differences of $\Delta r \geq .130$; $n_{\text{male}} = 49$; $n_{\text{female}} = 47$. EAR = Electronically Activated Recorder.

* $p \leq .05$ (two-tailed). ** $p \leq .01$ (two-tailed).

Table 6
Accuracy of Judges' Implicit Folk Theories of How Personality Is Manifested in Daily Life: Column-Vector Correlations Between the Behavioral Manifestations and Judges' Cue Use and Overall Self-Judge Agreement

Question 3: Big Five dimensions	Column-vector correlations			Self-judge agreement		
	<i>N</i> = 96	Males	Females	<i>N</i> = 96	Males	Females
Extraversion	.58	.56	.60	.41**	.38**	.44**
Agreeableness	.46	.42	.39	.25**	.15	.32*
Conscientiousness	.58	.57	.40	.27**	.26*	.26*
Emotional Stability	.53	-.36	.51	.30**	-.11	.37**
Openness to Experience	.39	.08	.45	.21*	.32*	.11

Note. The column-vector correlations reflect the degree of convergence between the patterns of correlations for the behavioral manifestations of personality (left-hand sides of Tables 1–5, data columns 1–3) and the correlations for judges' cue use (right-hand sides of Tables 1–5; data columns 4–6); the column-vector correlations are based on the full set of Fisher's *r*-to-*z*-transformed correlations for the 31 SECSI and 23 LIWC variables (instead of only the correlations depicted in Tables 1–5); self-judge agreement is the correlation between participants' personalities and the aggregated judges' ratings of participants' personalities. No significance tests for the column-vector correlations are provided because the sometimes highly correlated EAR-derived variables violate the assumption of independence of observations (i.e., correlations); $n_{\text{male}} = 49$; $n_{\text{female}} = 47$. SECSI = Social Environment Coding of Sound Inventory; LIWC = Linguistic Inquiry and Word Count; EAR = Electronically Activated Recorder.

* $p \leq .05$ (one-tailed). ** $p \leq .01$ (one-tailed).

comparing judges' daily life based impressions of participants' personalities to participants' actual personalities, the study showed that people's implicit theories of personality are, to a considerable extent, accurate. An important finding emerged for the role of gender in the everyday expression and perception of personality: Many trait manifestations as well as aspects of judges' implicit folk theories emerged as gender differentiated. Below we discuss the broader implications of these findings by integrating them into existing research on everyday manifestations and judgments of personality.

Everyday Manifestations of Personality: The Role of Differential Environment Selection

Several recent studies have contributed to the "catalog of basic facts concerning the relationships between personality and behavior" (Funder, 2001, p. 212) by identifying reliable and ecologically valid trait-behavior links (Eaton & Funder, 2001; Funder & Sneed, 1993; Gifford & Hine, 1994; Oberlander & Gill, in press; Gosling et al., 2002; Paunonen, 2003; Rentfrow & Gosling, 2003; Spain et al., 2000; Wu & Clark, 2003). This study was the first to use an unobtrusive-observation approach to relate personality to daily behavior recorded in real-world contexts. Reassuringly, many of the identified manifestations were consistent with the definitions of the corresponding traits. The time spent talking, for example, was indicative of students' Extraversion, the use of swear words was negatively related to their Agreeableness, and class attendance was strongly associated with Conscientiousness.

In some cases, however, the study failed to confirm expected trait-behavior links and some nonpredicted patterns arose. For example, based on the relationship between Extraversion and positive affect (Lucas, Diener, Grob, Suh, & Shao, 2000), one would have expected the verbal or nonverbal expression of positive emotions (use of positive emotion words; laughing) to be

associated with Extraversion. Similarly, on the basis of the relationship between Neuroticism and negative affect (Watson & Clark, 1984), one would have expected the verbal and nonverbal expression of negative emotions (use of negative emotion words; crying, arguing) to be associated with a lack of Emotional Stability. Clearly, these exploratory findings need to be replicated. Specifically, further research is now needed that examines the generalizability of the findings to other populations, such as community samples whose daily environments are more diverse, to further our understanding of the processes that underlie trait expressions in people's ordinary social behaviors and interactions.

Our gender-specific analyses indicated that for all Big Five dimensions except Openness, aspects of daily life were differentially related to personality for male and female participants. For example, group conversations were indicative of Extraversion only among females, and self-talk was a marker of Introversion only among men. Similarly, only among women was laughing a sign of Emotional Stability and only among men was socializing in same-sex groups indicative of a lack of Emotional Stability. Conceptually, the idea that personality expressions are moderated by background variables such as age, culture, social class, or gender has intuitive appeal. Yet, Eaton and Funder (2001) recently noted that very few studies have reported trait manifestations broken down by gender (or other demographic variables). Methodologically, the probability of revealing gender moderation effects likely depends on how the phenomenon is approached. Many laboratory studies by necessity impose an unnatural constraint on participants by assigning them to a limited number of controlled situations instead of allowing them to actively execute their preferences for certain environments over others. Thus, lab studies may, by design, rule out proactive person-environment interactions and thereby disguise gender-specific personality implications that result from gender-linked preferences in environment selection (Buss, 1987; Ickes et al., 1997).

With their ability to capture situational preferences, naturalistic assessment methods such as the EAR may help reveal existing environment-dependent personality implications. For example, in line with this argument, gender-specific trait manifestations have recently been found in real-world studies on people's daily emotional experiences (Eaton & Funder, 2001) and physical environments (Gosling, Craik, Martin, & Pryor, 2005). Future research should build on the exploratory foundation provided here to develop a theoretical account of how gender-linked psychological processes (such as an agentic vs. communal orientation or an instrumental vs. expressive tendency) can affect everyday expressions of personality (Stewart & McDermott, 2004; Wood & Eagly, 2002).

The Architecture and Accuracy of People's Implicit Folk Theories of Personality

Our findings regarding judges' implicit folk theories of how personality is expressed in daily life mapped generally well onto how experts and laypersons understand the Big Five dimensions (Sneed, McCrae, & Funder, 1998). Talking was considered an indicator of Extraversion, cursing an expression of a lack of Agreeableness, laughing a sign of Emotional Stability, and class attendance a marker of Conscientiousness. Furthermore, judges in general considered these core trait indicators equally trait-relevant for both genders.

Frequently, however, similar behaviors affected judges' personality impressions differently for male and female targets. For example, judges' implicit folk theories about everyday manifestations of Emotional Stability evidenced rather pronounced gender differences. Whereas among women negative emotional expression in the form of arguing was perceived as a lack of Emotional Stability, among men, the same behavior tended to lead to opposite perceptions. Furthermore, whereas talking tended to be perceived as generally indicative of Emotional Stability in men, only a preference for dyadic conversations was seen as a sign of Emotional Stability in women. Although these findings may not be readily predicted from prior research on gender stereotypes (Lippa & Connelly, 1990; Madon et al., 1998; Swim, 1994), they do suggest that gender stereotypes can influence the personality implications that people ascribe to behaviors for men and women. The fact that judges' implicit folk theories were grounded in actual behavioral manifestations of Emotional Stability only among women suggests that people may hold largely incorrect assumptions about the emotional implications of the social behaviors of men.

Nonetheless, the overall findings showed generally strong accuracy in judges' implicit folk theories of personality, largely replicating existing research on everyday personality judgment in other zero-acquaintance studies (Borkenau & Liebler, 1992; Funder & Sneed, 1993; Gosling et al., 2002; Kenny & Albright, 1987; Vazire & Gosling, 2004). With the exception of ratings of male students' Emotional Stability and Agreeableness and female students' Openness, self-judge correlations were substantial for all Big Five dimensions. Also, accuracy was highest for Extraversion and Conscientiousness, a finding that could be explained by the good observability of trait-relevant behaviors for these dimensions in the contexts examined here (e.g., Funder & Sneed, 1993; John & Robins, 1993; Watson, Hubbard, & Wiese, 2000).

Previous research found lower levels of accuracy for personality judgments of evaluative traits, such as Emotional Stability (Borkenau & Liebler, 1992; Oberlander & Gill, in press; John & Robins, 1993; Markey & Wells, 2002). Yet, in this study, judges' ratings of Emotional Stability showed a high degree of accuracy among female participants. Why are EAR recordings potentially more revealing than other contexts about a person's emotional life? Recall that the EAR samples ambient sounds across the full spectrum of people's daily social encounters, many of which may not be represented in most zero-acquaintance contexts. With its fine-grained grid of observations it is able to capture behavior that hints at less publicly presentable aspects of personality. For example, the method regularly catches emotional outbursts, arguments, intimate conversations, profanity, politically incorrect, and sexually provocative remarks. In addition to documenting a person's behavior "on stage" (Goffman, 1959), it also reveals those moments where humans are caught off guard showing their usually hidden, weak, and unpolished faces.

Drawing on Gosling et al.'s (2002) distinction between intentional identity claims and nonintentional behavioral residue, people's moment-to-moment social environments and conversations are saturated with unintentional messages that leave a residue in the EAR records. The words people spontaneously use in their daily interactions are a prime example of behavior that largely escapes voluntary control. On the basis of these considerations, Vazire and Gosling (2004) speculated that a person's daily life should carry particularly rich information with regard to evaluative traits. This study confirmed their prediction in the context of Emotional Stability, albeit only for female students. Despite the abundance of expressive behavior in daily life, the EAR sounds did not provide clear clues to the emotionality of male participants. This finding is consistent with research showing that men generally express fewer overt signs of Emotional Stability than women, presumably because men tend to be raised to conceal signs of nervousness or insecurity more than women do (e.g., Brody, 1999; Fischer, Manstead, & Zaalberg, 2003).

A potential source of bias with regard to judges' implicit folk theories results from the finding that the correlations for the perceived trait-behavior links were consistently larger than the correlations for the actual trait-behavior links. For example, the amount of captured talking was correlated .67 with judges' impressions of targets' Extraversion, but it was only correlated .30 with participants' self-reported Extraversion. Similarly, the use of swear words was correlated -.52 with perceived but only -.28 with actual Agreeableness. This suggests that judges gave more weight to available cues than is warranted based on the magnitude of the actual real-world trait-behavior associations. Other research on personality judgment has found a similar asymmetry (e.g., Borkenau & Liebler, 1992; Gifford & Hine, 1994; Gosling et al., 2002).

It is possible that this asymmetry reflects an observer's tendency to overestimate the effects that a single personality variable has onto people's everyday behaviors. It might, however, also be a result of the particular study design where participants' real-world behaviors were naturally causally overdetermined but judges' impressions were, by design, causally constrained to a very limited set of information. Whereas participants' EAR-assessed behaviors were likely simultaneously caused by various factors that were not assessed in the study, judges derived their impressions exclusively

from the information that was captured by the EAR. Furthermore, the implicit-folk-theories correlations are based on variables that were derived from the same medium (i.e., the recorded ambient sounds); the correlations for the manifestations of personality, however, are based on variables that share no method variance. Thus, the asymmetry in the strength of assumed versus real trait–behavior associations is an interesting finding, but additional research is required to determine whether it is driven by substantive or artifactual factors.

Finally, although the judges had never met the participants and had no formal personality training, they were not entirely naïve about the personality implications of participants' daily lives. Both judges and participants were undergraduate students at the University of Texas at Austin and thus shared a common everyday environment. This shared meaning system (e.g., the meaning of "Longhorns," "6th Street," or "the drag" in a University of Texas student context) might have positively affected the agreement among judges and the accuracy estimates for people's implicit personality theories (Kenny, 1991; Funder, 1995). Our findings thus most readily apply to trait–behavior associations that people implicitly hold within the context of their own daily environments.

Directions for Future Research

This study provided an exploratory analysis of the personality implications of naturally occurring daily life. It addressed basic questions surrounding everyday manifestations of personality as well as people's implicit folk theories of how personality is related to observable aspects of daily life. This research provides the foundation for several avenues of future research.

First, this project focused on the daily life of college students. Undoubtedly, many base rates revealed by the EAR are sensitive to sample characteristics (e.g., time spent in class, working at a job). However, differences in base rates between populations do not necessarily translate into differences in the underlying processes. For example, if elderly people were found to have, on average, fewer daily conversations than students, individual differences in the time spent talking could still be indicative of Extraversion. On the other hand, sometimes population-based differences in daily behavior inevitably do produce differences in the expression of personality. Consequently, the generalizability of the identified personality implications of daily life needs to be tested in other populations.

Second, future research should explore behavioral manifestations of personality from a cross-cultural perspective. As noted by McCrae and Costa (1999), how "Conscientiousness is expressed in Italy is likely to be very different from how it is expressed in Iran. Ethnographic methods might be needed to identify the culturally prescribed forms in which personality factors are manifested" (p. 149). Although the EAR is not an ethnographic method, it can capture the daily life of communities in unfiltered, vivid, and contextually rich ways. Also, recording a culture's ambient sounds can help bypass a nagging problem in cross-cultural research—the potential nonequivalence of questionnaire translations (Church, 2001). Thus, it would be highly valuable to use the EAR or other naturalistic observation methods to document how the social en-

vironments into which individuals are socialized constrain the expression of personality (Ramírez-Esparza, Mehl, & Pennebaker, 2005).

Third, future research should consider moving beyond the Big Five domain when seeking to identify important personality implications of people's daily behaviors. Undoubtedly, the Big Five have proven to be extremely useful as an organizing framework for personality research and were a logical starting point for this first exploratory study. Nevertheless, they have also been the target of serious concerns (e.g., Block, 1995; Paunonen, & Jackson, 2000). Within a given study, their use should be driven by conceptual adequacy rather than by common practice. This project sought to explore manifestations and implicit folk theories of personality at a broad level. It is possible that by assessing personality at a global level but measuring behavior at a specific level, it pursued a less-than-optimal strategy for identifying real-world trait–behavior links. Future studies should incorporate other personality variables that extend beyond and are more specific than the Big Five, including Big Five facets (Costa, & McCrae, 1992) and the California Q-sort (Block, 1978). For example, a recent follow-up to this exploratory project used the current paradigm to investigate behavioral manifestations of depression in daily life (Mehl, in press).

Fourth, another important direction for future research concerns the conceptual and empirical development of the behavioral and linguistic categories. Since its original development, the SECSI categories have already undergone a series of revisions (Mehl & Pennebaker, 2003a, 2003b). Beyond adding new behavioral categories (e.g., religious activities), we have recently begun to zoom in on participants' social interactions and have developed coding systems for the topics (e.g., relationships, fashion, health) and the purposes (e.g., gossip, disclosure, school) of participants' daily conversations. The LIWC categories have been refined in similar ways (Pennebaker et al., 2001). On a broader level, however, it is important to step back and reflect how the specific coding approach we adopted, namely to extract concrete, molecular and highly specific information about people's daily lives, constrained the research findings. Valuable avenues for future research would be to explore coding systems that operate at higher, psychologically more meaningful, levels of analysis (Funder, Furr, & Colvin, 2000) and, when larger samples become available, to use multivariate methods to combine empirically related categories.

Finally, future research should further explore people's implicit folk theories of how personality is manifested in daily life. The ambient sounds provided by the EAR constitute naturalistic stimuli for testing the extent to which behaviors have different personality implications for different social groups. This study found that judges assumed some degree of gender specificity in everyday trait–behavior links. Stereotypes that people hold about social groups often contain implicit assumptions about the groups' daily behaviors, environments, or interactions. Using more diverse samples would allow tests of how behaviors are implicitly evaluated based on group membership that is identified (e.g., ethnicity) or assumed (e.g., sexual orientation) from the ambient sounds. This study provides a paradigm for testing the relative accuracy and bias of the theories that people implicitly hold about other people's daily lives.

Broader Implications of a Naturalistic Observation Approach to Studying Personality

As a tool for sampling observations in naturalistic settings, the EAR fills an important gap in the psychological study of person–environment interactions (Walsh, Craik, & Price, 2000); however, the method has obvious limitations. For example, it allows only the “acoustic observation” of daily life, and many interesting social phenomena can only be grasped visually. Also, despite generally high compliance rates, the method likely oversamples publicly displayable behaviors and underrepresents private, intimate, socially undesirable, deviant, or illegal behavior. Furthermore, the perspective adopted by the EAR does not capture how people subjectively interpret and make sense of their daily lives (Bolger et al., 2003). Finally, as any researcher in this area can attest, the data collection process involved in observational studies can consume large quantities of time and resources.

Despite the shortcomings of the EAR method, however, the study of naturalistic person–environment interactions using innovative observation technology (e.g., Craik, 1994; Intille et al., 2003; Watanabe, Izawa, Kato, Ropert-Coudert, & Naito, 2005) holds great potential for personality psychology. The question “Who should own the definition of personality?” (Hofstee, 1994) has long concerned researchers in the field. Self-reports and informant reports both contain large components of accuracy, but both also have their blind spots (John & Robins, 1993; Kolar, Funder, & Colvin, 1996; Spain et al., 2000). Kenny (1994) has argued that in an ideal world, behavioral measures of personality would be the gold standard with which to compare other sources of information about a person. These new methods bring us a step closer to this standard.

In many ways, people’s real-world interactions within their social environments are the very things social and personality psychologists want to know about. They reveal what people spontaneously do, the situations they naturally seek out or avoid, and the idiosyncratic ways in which they connect with their social worlds. The EAR is a method that allows this information to be collected with a high degree of ecological validity. The recorded ambient sounds provide a vivid sensory document of life as the person lived it—preserved in its natural historical and social context. Furthermore, the observer’s perspective in psychological assessment has high social relevance; humans act as bystanders most of the time and a person’s interpretation of an event is in most situations privileged information that is inaccessible to others. In everyday life, the only information that is readily available is the raw, molecular behavior, stripped of its psychological context—information as it is captured by the EAR. These arguments should encourage researchers to reconsider how to think about and measure personality. The efficiency of self-reports is beyond comparison. However, as Barker and Wright (1951) recognized more than half a century ago, it is the potential to document personality right where it occurs that makes naturalistic observation approaches such as the EAR a vital avenue for enhancing our understanding of human personality as it is expressed in its natural habitat.

References

- Barker, R. G. (1968). *Ecological psychology: Concepts and methods for studying the environment of human behavior*. Stanford, CA: Stanford University Press.

- Barker, R. G., & Wright, H. F. (1951). *One boy’s day. A specimen record of behavior*. New York: Harper & Row.
- Block, J. (1978). *The Q-sort method in personality assessment and psychiatric research*. Palo Alto, CA: Consulting Psychologists Press.
- Block, J. (1989). A critique of the Act Frequency Approach to personality. *Journal of Personality and Social Psychology*, *56*, 234–245.
- Block, J. (1995). A contrarian view of the five-factor approach to personality description. *Psychological Bulletin*, *117*, 187–215.
- Bolger, N., Davis, A., & Rafaeli, E. (2003). Diary methods: Capturing life as it is lived. *Annual Review of Psychology*, *54*, 579–616.
- Borkenau, P., & Liebler, A. (1992). Trait inferences: Sources of validity at zero acquaintance. *Journal of Personality and Social Psychology*, *62*, 645–657.
- Borkenau, P., Mauer, N., Riemann, R., Spinath, F. M., & Angleitner, (2004). Thin slices of behavior as cues of personality and intelligence. *Journal of Personality and Social Psychology*, *86*, 599–614.
- Brody, L. (1999). *Gender, emotion, and family*. Cambridge, MA: Harvard University Press.
- Buss, D. M. (1987). Selection, evocation, and manipulation. *Journal of Personality and Social Psychology*, *53*, 1214–1221.
- Buss, D. M., & Craik, K. H. (1983). The act frequency approach to personality. *Psychological Review*, *90*, 105–126.
- Buss, D. M., & Craik, K. H. (1986). The act frequency approach and the construction of personality. In A. Angleitner, A. Furnham, & G. van Heck (Eds.), *Personality psychology in Europe. Vol. 2: Current trends and controversies* (pp. 141–156). Berwyn, PA: Swets North America.
- Chaplin, W. F., Phillips, J. B., Brown, J. D., Clanton, N. R., & Stein, J. L. (2000). Handshaking, gender, personality, and first impressions. *Journal of Personality and Social Psychology*, *79*, 110–117.
- Church, A. T. (2001). Personality measurement in cross-cultural perspective. *Journal of Personality*, *69*, 979–1006.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Costa, P. T., & McCrae, R. R. (1992). *NEO PI-R professional manual*. Odessa, FL: Psychological Assessment Resources.
- Craik, K. H. (1986). Personality research methods: An historical perspective. *Journal of Personality*, *54*, 18–51.
- Craik, K. H. (1994). Manifestations of individual differences in personality within everyday environments (pp. 19–25). In D. Bartussek & M. Amelang (Eds.), *Fortschritte der Differentiellen Psychologie und Psychologischen Diagnostik: Festschrift zum 60. Geburtstag von Professor Dr. Kurt Pawlik* [Advances in differential psychology and psychological assessment: Festschrift on the occasion of the 60th birthday of Professor Dr. Kurt Pawlik]. Göttingen, Germany: Hogrefe.
- Craik, K. H. (2000). The lived day of an individual: A person–environment perspective. In W. B. Walsh, K. H. Craik, & R. H. Price (Eds.), *Person–environment psychology: New directions and perspectives* (pp. 233–266). Mahwah, NJ: Erlbaum.
- Eaton, L. G., & Funder, D. C. (2001). Emotional experience in daily life: Valence, variability, and rate of change. *Emotion*, *1*, 413–421.
- Feldman Barrett, L., & Barrett, D. J. (2001). Computerized experience sampling: How technology facilitates the study of conscious experience. *Social Science Computer Review*, *19*, 175–185.
- Fischer, A. H., Manstead, A. S. R., & Zaalberg, R. (2003). Social influences on the emotion process. *European Review of Social Psychology*, *14*, 171–201.
- Funder, D. C. (1995). On the accuracy of personality judgment: A realistic approach. *Psychological Review*, *102*, 652–670.
- Funder, D. C. (2001). Personality. *Annual Review of Psychology*, *52*, 197–221.
- Funder, D. C., Furr, R. M., & Colvin, C. R. (2000). The Riverside Behavioral Q-sort: A tool for the description of social behavior. *Journal of Personality*, *68*, 451–489.
- Funder, D. C., & Sneed, C. D. (1993). Behavioral manifestations of

- personality: An ecological approach to judgmental accuracy. *Journal of Personality and Social Psychology*, 64, 479–490.
- Furr, R. M. (2002). Psychology and astrophysics: Overcoming physics envy. *Dialogue*, 17, 17.
- Gifford, R., & Hine, D. (1994). The role of verbal behavior in the encoding and decoding of interpersonal dispositions. *Journal of Research in Personality*, 28, 115–132.
- Goffman, E. (1959). *The presentation of self in everyday life*. Garden City, NY: Doubleday.
- Goldberg, L. R. (1990). An alternative “Description of personality”: The Big-Five factor structure. *Journal of Personality and Social Psychology*, 59, 1216–1229.
- Gosling, S. D., Craik, K. H., Martin, N. R., & Pryor, M. R. (2005). Material attributes of Personal Living Spaces. *Home Cultures*, 2, 51–88.
- Gosling, S. D., Ko, S. J., Mannarelli, T., & Morris, M. E. (2002). A room with a cue: Personality judgments based on offices and bedrooms. *Journal of Personality and Social Psychology*, 82, 379–398.
- Hemphill, J. F. (2003). Interpreting the magnitude of correlation coefficients. *American Psychologist*, 58, 78–79.
- Hofstee, W. K. B. (1994). Who should own the definition of personality? *European Journal of Personality*, 8, 149–162.
- Hogan, R. (1982). A socioanalytic theory of personality. *Nebraska Symposium of Motivation*, 30, 55–89.
- Ickes, W., Snyder, M., Garcia, S. (1997). Personality influences on the choice of situations. In R. Hogan, J. Johnson, & S. Briggs (Eds.), *Handbook of personality psychology* (pp. 165–195). New York: Academic Press.
- Intille, S. S., Tapia, E. M., Rondoni, J., Beaudin, J., Kukla, C., Agarwal, S., & Bao, L. (2003). Tools for studying behavior and technology in natural settings (pp. 157–174). In A. K. Dey, A. Schmidt, & J. F. McCarthy (Eds.), *Proceedings of UBIComp 2003: Ubiquitous Computing, vol. LNCS 2864*. Berlin: Springer.
- 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.
- John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality theory and research* (pp. 102–138). New York: Guilford Press.
- Kenny, D. A. (1991). A general model of consensus and accuracy in interpersonal perception. *Psychological Review*, 98, 155–163.
- Kenny, D. A. (1994). *Interpersonal perception: A social relations analysis*. New York: Guilford Press.
- Kenny, D. A., & Albright, L. (1987). Accuracy in interpersonal perception: A social relations analysis. *Psychological Bulletin*, 102, 390–402.
- Kolar, D. W., Funder, D. C., & Colvin, C. R. (1996). Comparing the accuracy of personality judgments by the self and knowledgeable others. *Journal of Personality*, 64, 311–337.
- Lee, Y. T., Jussim, L., & McCauley, C. R. (Eds.). (1995). *Stereotype accuracy: Toward appreciating group differences*. Washington, DC: American Psychological Association.
- Lippa, R. (1978). The naive perception of masculinity-femininity on the basis of expressive cues. *Journal of Research in Personality*, 12, 1–14.
- Lippa, R., & Connelly, S. (1990). Gender diagnosticity: A new Bayesian approach to gender-related individual differences. *Journal of Personality and Social Psychology*, 59, 1051–1065.
- Lucas, R. E., Diener, E., Grob, A., Suh, E. M., & Shao, L. (2000). Cross-cultural evidence for the fundamental features of extraversion. *Journal of Personality and Social Psychology*, 79, 452–468.
- Madon, S. J., Jussim, L., Keiper, S., Eccles, J., Smith, A., & Palumbo, P. (1998). The accuracy and power of sex, social class and ethnic stereotypes: Naturalistic studies in person perception. *Personality and Social Psychology Bulletin*, 24, 1304–1318.
- Markey, P. M., & Wells, S. M. (2002). Interpersonal perception in Internet chat rooms. *Journal of Research in Personality*, 36, 134–146.
- McCrae, R. R., & Costa, P. T., Jr. (1999). A five-factor theory of personality. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality theory and research* (pp. 139–153). New York: Guilford Press.
- Mehl, M. R. (in press). The lay assessment of sub-clinical depression in daily life. *Psychological Assessment*.
- Mehl, M. R., & Pennebaker, J. W., (2003a). The social dynamics of a cultural upheaval: Social interactions surrounding September 11, 2001. *Psychological Science*, 14, 579–585.
- Mehl, M. R., & Pennebaker, J. W. (2003b). The sounds of social life: A psychometric analysis of students’ daily social environments and natural conversations. *Journal of Personality and Social Psychology*, 84, 857–870.
- Mehl, M. R., Pennebaker, J. W., Crow, M., Dabbs, J., & Price, J. (2001). The Electronically Activated Recorder (EAR): A device for sampling naturalistic daily activities and conversations. *Behavior Research Methods, Instruments, and Computers*, 33, 517–523.
- Oberlander, J., & Gill, A. (in press). Language with character: A stratified corpus comparison of individual differences in e-mail communication. *Discourse Processes*.
- Paunonen, S. V. (2003). Big Five factors of personality and replicated prediction of behavior. *Journal of Personality and Social Psychology*, 84, 411–424.
- Paunonen, S. V., & Jackson, D. N. (2000). What is beyond the Big Five? *Plenty! Journal of Personality*, 68, 821–835.
- Pawlik, K., & Buse, L. (1982). *Rechnergestützte Verhaltensregistrierung im Feld: Beschreibung und erste psychometrische Überprüfung einer neuen Erhebungsmethode*. [Computer-supported behavior recording in the field: Description and first psychometric test of a new data collection method.] *Zeitschrift für Differentielle und Diagnostische Psychologie*, 3, 101–108.
- Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2001). *Linguistic Inquiry and Word Count: LIWC2001*. Mahwah, NJ: Erlbaum. www.liwc.net
- Pennebaker, J. W., Mehl, M. R., Niederhoffer, K. G. (2003). Psychological aspects of natural language use: Our words, our selves. *Annual Review of Psychology*, 54, 547–577.
- Ramírez-Esparza, N., Mehl, M. R., & Pennebaker, J. W. (2005, January). *Do Mexicans and Americans behave differently? A cross-cultural study using the electronically activated recorder (EAR)*. Poster presented at the 6th annual meeting of the Society for Personality and Social Psychology, New Orleans, LA.
- Reis, H. T., & Wheeler, L. (1991). Studying social interaction with the Rochester Interaction Record. *Advances in Experimental Social Psychology*, 24, 269–318.
- Rentfrow, P. J., & Gosling, S. D. (2003). The do re mi’s of everyday life: The structure and personality correlates of music preferences. *Journal of Personality and Social Psychology*, 84, 1236–1256.
- Rozin, P. (2001). Social psychology and science: Some lessons from Solomon Asch. *Personality and Social Psychology Review*, 5, 2–14.
- Scollon, C. N., Kim-Prieto, C., & Diener, E. (2003). Experience sampling: Promises and pitfalls, strengths and weaknesses. *Journal of Happiness Studies*, 4, 5–34.
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing reliability. *Psychological Bulletin*, 86, 420–428.
- Sneed, C. D., McCrae, R. R., & Funder, D. C. (1998). Lay conceptions of the five-factor model and its indicators. *Personality and Social Psychology Bulletin*, 24, 115–126.
- Spain, J. S., Eaton, L. G., & Funder, D. C. (2000). Perspectives on personality: The relative accuracy of self versus others in the prediction of emotion and behavior. *Journal of Personality*, 68, 837–867.
- Stewart, A. J., & McDermott, C. (2004). Gender in psychology. *Annual Review of Psychology*, 55, 519–544.

- Swim, J. K. (1994). Perceived versus meta-analytic effect sizes: An assessment of the accuracy of gender stereotypes. *Journal of Personality and Social Psychology*, *66*, 21–36.
- Vazire, S., & Gosling, S. D. (2004). e-Perceptions: Personality impressions based on personal web-sites. *Journal of Personality and Social Psychology*, *87*, 123–132.
- Walsh, W. B., Craik, K. H., & Price R. H. (Eds.) (2000). *Person–environment psychology: New directions and perspectives*. Mahwah, NJ: Erlbaum.
- Watanabe, S., Izawa, M., Kato, A., Ropert-Coudert, Y., & Naito, Y. (2005). A new technique for monitoring the detailed behaviour of terrestrial animals. *Applied Animal Behaviour Science*, *94*, 117–131.
- Watson, D., & Clark, L. A. (1984). Negative affectivity: The disposition to experience aversive emotional states. *Psychological Bulletin*, *96*, 465–490.
- Watson, D., Hubbard, B., & Wiese, D. (2000). Self-other agreement in personality and affectivity: The role of acquaintanceship, trait visibility, and assumed similarity. *Journal of Personality and Social Psychology*, *78*, 546–558.
- Wood, W., & Eagly, A. H. (2002). A cross-cultural analysis of the behavior of women and men: Implications for the origins of sex differences. *Psychological Bulletin*, *128*, 699–727.
- Wu, K. D., & Clark, L. A. (2003). Relations between personality traits and self-reports of daily behavior. *Journal of Research in Personality*, *37*, 231–256.

Received March 11, 2005

Revision received September 28, 2005

Accepted October 11, 2005 ■