

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/12688664>

# Linguistic styles: Language use as an individual difference

Article in *Journal of Personality and Social Psychology* · January 2000

DOI: 10.1037//0022-3514.77.6.1296 · Source: PubMed

CITATIONS

1,135

READS

20,760

2 authors:



James W. Pennebaker

University of Texas at Austin

385 PUBLICATIONS 57,744 CITATIONS

[SEE PROFILE](#)



Laura A King

University of Missouri

141 PUBLICATIONS 16,876 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Measurement and psychosocial factors of well being [View project](#)



Collective memory and social representations of history [View project](#)

## Linguistic Styles: Language Use as an Individual Difference

James W. Pennebaker  
University of Texas at Austin

Laura A. King  
Southern Methodist University

Can language use reflect personality style? Studies examined the reliability, factor structure, and validity of written language using a word-based, computerized text analysis program. Daily diaries from 15 substance abuse inpatients, daily writing assignments from 35 students, and journal abstracts from 40 social psychologists demonstrated good internal consistency for over 36 language dimensions. Analyses of the best 15 language dimensions from essays by 838 students yielded 4 factors that replicated across written samples from another 381 students. Finally, linguistic profiles from writing samples were compared with Thematic Apperception Test coding, self-reports, and behavioral measures from 79 students and with self-reports of a 5-factor measure and health markers from more than 1,200 students. Despite modest effect sizes, the data suggest that linguistic style is an independent and meaningful way of exploring personality.

That people differ in the ways they talk and write is hardly a novel observation. Even when the content of the message is the same, individuals express themselves verbally with their own distinctive styles. As with spoken language, written language is also unique from person to person. Teachers and researchers can often decipher the identity of students when grading their exams or can determine the reviewers of their manuscripts simply by the words and phrases the writers use. Indeed, forms of "linguistic fingerprinting" have been suggested in research and the media for generations. Analyses of word frequencies have been used to distinguish letters written by soldiers in the 1800s (e.g., Broehl & McGee, 1981), to understand speaking styles of political leaders (e.g., Hart, 1984), to establish the identities of the authors of biblical and literary works and, most recently, to determine the anonymous author of a best-selling book surrounding the presidency of Bill Clinton (Foster, 1996). Thus, at least to some extent, the ways people talk and write have been recognized as stamps of individual identity.

Despite the assumption that people verbally express themselves in stable, unique ways, language use has rarely been considered to be a bona fide individual difference. Certainly, language and word use have been shown to vary tremendously from topic to topic and from one social context to another (e.g., Forgas, 1985; Giles &

Wiemann, 1993). Powerful situational constraints on language may be one reason why relatively little attention has been given to language use as an individual difference. However, a number of both practical and conceptual issues have impeded this problem even more.

One of the most daunting problems in assessing linguistic styles is in deciding on the appropriate dimensions of language and, once selected, determining the best unit of analysis. Since the 1950s, several attempts have been made to devise comprehensive text-analytic strategies to efficiently study multiple psychological dimensions of speech and text. These approaches have focused on phrases, thought units, and even the word choices that people adopt.

Two early strategies dominated text analysis in psychology. The first, which was strongly influenced by physician Louis Gottschalk (e.g., Gottschalk & Gleser, 1969), is heavily psychoanalytic in orientation and requires trained raters to evaluate each clause of a sentence (for a recent computerized version, see Gottschalk, Stein, & Shapiro, 1997). Speech samples are generated by individuals who are asked to talk aloud in a stream of consciousness mode for 5 min. The majority of Gottschalk's work has been oriented toward distinguishing individuals with different medical diagnoses. A similar, yet more straightforward coding procedure was used by another physician, Walter Weintraub (e.g., 1989). The Weintraub approach involved coding words and phrases into 15 general categories and, as in the Gottschalk strategy, comparing the use of these linguistic dimensions across people with different medical diagnoses and, in later years, historical personalities. Stiles (1992) has developed a method of communication analysis called verbal response modes, which also requires rating on a phrase-by-phrase basis. The technique is more oriented to the form and intent of disclosures than to tapping the underlying emotional and cognitive dynamics. Unlike the Gottschalk method, Stiles transcribes more natural interactions, such as those between physician and patient and, more recently, between psychotherapist and patient.

---

The research reported in this article was funded, in part, by National Institutes of Health Grants MH52391 and MH54142. We are deeply indebted to Geoffrey Stevenson for providing data for the first reliability study, to Shannon Wiltsey and Tom Lay for help in data collection for the third reliability study, to Janel Seagal for help in data collection for the introductory psychology writing samples, to Pat Randall and Tor Neilands for statistical advice, and to Kelli Keough for comments on an earlier version.

Correspondence concerning this article should be addressed to James W. Pennebaker, Department of Psychology, University of Texas, Austin, Texas 78712, or to Laura A. King, Department of Psychology, Southern Methodist University, Dallas, Texas 75275. Electronic mail may be sent to pennebaker@psy.utexas.edu or lking@mail.smu.edu.

A second text analysis strategy has been a word-based counting system. Such an approach assumes that one or more general concepts that a person is attempting to express can be captured by examining the specific words used to make up the concepts. Consequently, if individuals are verbally expressing sadness, they would be more likely to use words such as *sad*, *cry*, *loss*, or *alone*. One of the first and most successful word-based approaches was developed by Stone, Dunphy, Smith, and Ogilvy (1966) and called the General Inquirer. Originally, the General Inquirer technique relied on the Harvard psychosociological dictionaries that tapped specific words thought to be linked to need states, motives, social and cultural roles, and different aspects of general distress. By the 1970s, the Dartmouth adaptation of the General Inquirer had been updated for more general use on mainframe computers (e.g., Rosenberg, Schnurr, & Oxman, 1990; Rosenberg & Tucker, 1979). Many other word-based text analysis programs followed, a number of which were guided by psychoanalytic themes (Bucci, 1997; Horowitz, Milbrath, & Stinson, 1995; Mergenthaler, 1996). Interestingly, the majority of studies relying on these computer-based text analysis methods focused on distinguishing medical and psychiatric diagnoses. For example, several studies indicate that text analyses can accurately distinguish somatization disorders (Oxman, Rosenberg, Schnurr, & Tucker, 1988), schizophrenia (Taylor, Reed, & Berenbaum, 1994), cancer proneness (Spence, Scarborough, & Ginsberg, 1978), suicidal tendencies (Thomas & Duszynski, 1985), and depression (Schnurr, Rosenberg, & Oxman, 1992).

Ironically, social and personality psychologists have been somewhat resistant to studying naturalistic language using word-based or phrase-based text analysis procedures. There have been interesting exceptions, such as Peterson's (e.g., Peterson, Seligman, & Vaillant, 1988; Peterson & Ulrey, 1994) work on attributional style, Tetlock's (e.g., Tetlock, Peterson, & Berry, 1993) investigation of integrative complexity, and Stein and Trabasso's (e.g., Stein, Folkman, Trabasso, & Richards, 1997) work on goal processes. Most of these approaches, however, have used text analysis in a highly focused way in an attempt to study a circumscribed number of psychological processes. In other words, most text-analytic strategies have been adopted to test particular theoretical orientations rather than examine the nature of language use per se. Similarly, personality approaches to studying language have considered self-reports of personality as the independent variable (e.g., self-esteem and Type A), with particular language domains as convenient dependent measures (e.g., Furnham, 1990).

The studies we present here examined the language that people use in a variety of written contexts to look for evidence of individual consistency across these contexts. Dating from Mischel's (1968) critique of the role of personality in the prediction of behavior, establishing the cross-situational consistency of behaviors has been a focus of research in personality. This research has focused primarily on "large-scale" behaviors such as helping or being friendly. In this investigation, we examined the cross-situational consistency of language use in the way that Allport (1961) described "stylistic behavior." Allport defined stylistic behavior as "one's manner of performing adaptive acts" (p. 461). Allport contrasted adaptive and stylistic behaviors as the difference between what one is doing (adaptive behavior) and how one is doing it (stylistic behavior). Style, Allport noted, refers to the "person idiom" that marks one's behaviors. Allport stated that

every act has an expressive aspect, an aspect that is not intentional or purposive but, rather, spontaneous and difficult to change. Although some stylistic aspects of behaviors may seem obvious—how people walk, dress, and smile—others may be less obvious.

We propose that the vocabulary that people habitually use to express themselves is a stylistic behavior. Consider, for example, how a group of fashion models would respond to the broad query "Tell me a little about yourself." Although many of the models might answer the question with some conceptual similarities, the way that each gives his or her description will differ in terms of how much is said and also the precise words that are used. The *what* of their answers will probably overlap a great deal, but the *how* may vary considerably. The purpose of the present investigation was to examine whether individual differences in language are stable across writing contexts and to explore whether such differences might relate to other, more established individual-difference characteristics.

Research that has used verbal expression as an important source of data about individuals (e.g., autobiographical memories and private wishes) assumes that the way people talk about things reveals important information about them. However, these strategies begin with the underlying constructs (such as motives or traits) that may be manifested in these narratives. Analyses of most written protocols typically rely on what individuals have written, so different ways of saying the same thing are treated as largely equivalent. Making sense of writing or speech samples also requires the luxury of a human rater, someone who is able to recognize the themes that pervade verbal material. In contrast, detecting the patterns of word usage over different topics of writing may demand attention to an aspect of text that most ignore as they read for meaning.

The adoption of a word usage approach to the analysis of naturally occurring written or spoken language is fraught with problems. Virtually all text analysis programs that rely on word counts are unable to consider context, irony, sarcasm, or even the problem of multiple meanings of words. Many of the traditional problems studied in communication, then, such as in-group-out-group status, formality of settings, and requests, are not easily detected with word counts (cf. Krauss & Fussell, 1996). In a discussion of the potential shortcomings of a computer program such as the General Inquirer, Zeldow and McAdams (1993) questioned whether lower level word counts can have true psychological meaning. For example, if a speaker routinely uses a high number of articles or adjectives, can this provide insight into the speaker's personality or psychological world? One difficulty in answering this question, the authors pointed out, was that the requisite empirical and theoretical work on language (as measured by word count programs) and personality had yet to be done.

In recent years, we have become concerned with developing efficient ways to examine a large number of text files of distressed individuals to learn whether the words people use to express themselves can predict long-term psychological and physical health (e.g., Pennebaker, 1997). As suggested by our literature search, most available text analysis programs adopted theoretical stances that were either outmoded or simply inconsistent with our own. Because we sought a system that would provide a wide range of psychological dimensions, we developed our own word-based text analysis program, called Linguistic Inquiry and Word Count, or LIWC. Unlike most previous text analysis strategies, however

LIWC's subjective dictionaries were independently rated by judges. In addition to careful construction of language categories, most LIWC dimensions were subsequently validated by having judges rate hundreds of files of written text; judges' ratings were then compared with comparable LIWC analyses of the same text files (see Pennebaker & Francis, 1996, 1999, for a description of the creation and validation of the LIWC program).

Now that the LIWC program is completed, we have become interested in establishing some of the basic psychometric properties of written language. In the present research, we investigated the degree to which language use is reliable across time and topic, possesses a reliable factor structure, and exhibits good construct and divergent validity in comparison with established personality measures. In the first phase of the work dealing with the reliability of language use, three very different writing samples were used: 18 diary entries from each of 15 inpatients in an addiction treatment unit who wrote in journals as part of their treatment, 10 class writing assignments from each of 35 health psychology summer students, and 15 psychology journal abstracts from 40 members of the Society for Experimental Social Psychology (SESP) who were first or solo authors of journal articles. These data were subjected to the LIWC word count program, and the possibility of consistent individual differences was explored.

The second phase of the research examined the factor structure of writing samples among a sample of 838 introductory psychology students. In a separate sample of 381 students, confirmatory analyses demonstrated the reliability of the four factors. The third phase of the research sought to establish the relationships between dimensions of language and traditional individual-difference measures, including the three need dimensions associated with the Thematic Apperception Test (TAT; Morgan & Murray, 1935) and the five-factor approach, as well as selected behavioral markers of health and college performance among introductory psychology students.

## PHASE 1: ESTABLISHING RELIABILITY OF LANGUAGE USE

Three separate studies analyzed multiple written language samples by different individuals to assess the reliability of language use. For ease of presentation, we first introduce the text analysis method before discussing the methods and findings of each of the separate studies.

### Text Analysis Procedures

Each essay for all of the studies was analyzed separately via the LIWC program (Pennebaker & Francis, 1999). LIWC relies on a dictionary of more than 2,200 words and word stems. These words are related to several broad categories, including language composition, psychological processes (including emotional, cognitive, and social processes), and current concerns. Each of these broad categories is composed of several specific dictionaries that tap particular constructs. For example, within the emotion category, specific dictionaries for overall negative emotion word use, anger words, positive emotions, and so forth were computed. The selection of words for each category was based on independent ratings by expert judges (see Pennebaker & Francis, 1996, and Pennebaker, Mayne, & Francis, 1997, for psychometric information).

The *percentage of words* in each person's writing sample that matched the words in the specific dictionary then was computed. Hence, all but three of the LIWC categories (total word count, words per sentence, and percentage of sentences ending in question marks) are represented as percentage of total words rather than raw word counts.

As depicted in the Appendix, LIWC categories represent 72 dimensions of language use. Each word or word stem defines one or more word categories or subdictionaries. For example, the word *cried* is part of four word categories: sadness, negative emotion, overall affect, and past-tense verb. Hence, if it is found in the target text, each of these four subdictionary scale scores will be incremented. Note also that many of the categories in LIWC are hierarchically arranged. For example, all of the words in the affect category come from the negative emotion and positive emotion subcategories. The negative emotion dimension is made up of words from the anxiety, anger, and sadness subcategories together with a substantial number of words with negative emotional connotations that may not fit into these subcategories (e.g., *guilty* and *stupid*). A brief description of the main categories follows.

### Language Composition

The subcategories within the language composition category include basic descriptive information about each file: total number of words, mean number of words per sentence, percentage of sentences ending with question marks, percentage of unique words used within the essay (also called type-token ratio), percentage of words captured by at least one of the LIWC dictionaries, and total percentage of words longer than six letters. Several of the remaining language composition variables are traditional parts of speech and are expressed as percentages of total words within each essay: pronouns (broken down separately by first, second, and third person), articles, and prepositions. A final group of subcategories includes negations (e.g., *no*, *not*, and *can't*), assents (found more often in oral transcripts; e.g., *yes* and *OK*), written numbers (e.g., *six* and *hundred*), and swear words.

### Psychological Processes

The linguistic dimensions associated with psychological dynamics were created to tap emotional, cognitive, sensory, and social processes. Unlike the more objective language composition categories, the words in the various dictionaries within the psychological processes dimensions were rated by groups of independent judges. All of the word categories in these dimensions are computed as percentage of total words.

The affective or emotional process dimensions tap many of the traditional categories found in other psychologically based text analysis programs (e.g., Stone et al., 1966). Positive emotion words include the broadest level of positive feeling dimensions (e.g., *happy* and *love*) as well as positively valenced words (e.g., *beautiful* and *nice*). Words from the optimism-energy category (e.g., *exciting* and *win*) are also part of the positive emotion dimension. As noted earlier, the negative emotion category includes a variety of negatively valenced terms (e.g., *ugly* and *hurt*) as well as words from the more specific categories of anxiety-fear, anger, and sadness-depression.

The cognitive processes dimension attempts to capture words that tap active thinking. Causal words, for example, connote attempts to explain causes and effects (e.g., *because*, *reason*, and *why*). The insight, or self-reflection, category captures words suggestive of learning or understanding (e.g., *know* and *realize*). Together, causal and insight word usage has been found to be related to mental and physical health (Pennebaker et al., 1997). Words such as *should*, *would*, and *could* are examples of the discrepancy category, which Higgins, Vookles, and Tykocinski (1992) have linked to health. Inhibition (e.g., *hold* and *restrain*), tentative (e.g., *maybe* and *perhaps*), and certainty (e.g., *absolute* and *always*) are the remaining categories within the cognitive processes dimension.

The sensory and social processes categories are self-explanatory. The sensory dimensions measure the degree to which individuals use seeing, hearing, and feeling language. Social process language includes references to other people through communication (e.g., *talk*, *listen*, and *share*) and through pronoun use (all pronouns except first-person singular), as well as references to friends, family, and other humans (e.g., *child* and *man*).

### Relativity

Generally, the relativity dimensions refer to time, space, and motion. Unlike psychological processes, however, these dimensions are not consistently hierarchical. The general time dictionary is composed of words that make temporal references (e.g., *hour*, *month*, *o'clock*, and *Tuesday*). The subcategories of the time dimension are simply temporal references based on verb tense: past, present, and future tense verbs. Space categories, too, are a mix of dimensions. The overall space dictionary is composed of prepositions (e.g., *above* and *beyond*), adjectives (e.g., *large* and *deep*), and other parts of speech depicting reference to space or location. Two hierarchically arranged subcategories include the dimensions of up (*high* and *above*) and down (*below*). Inclusive and exclusive prepositions and conjunctions are part of the space dimension, but not all words are part of the broader space category. The inclusive dimension is made up of words people use in trying to encompass or join categories (e.g., *and*, *with*, and *include*). Exclusive words are used to distinguish what is included in a category and what is not (e.g., *exclude*, *but*, and *without*). Finally, motion words simply refer to movement (e.g., *went*, *drive*, and *run*).

### Current Concerns

The current concerns dimension reflects many of the traditional categories seen in content analysis schemes. As can be deduced from the Appendix, the categories are self-explanatory: occupation (school, work, and achievement), leisure (home, sports, television, and music), money, metaphysical issues (religion and death), and physical states (body, sexuality, eating-drinking, sleeping, and grooming). Each dimension is hierarchically structured and, as might be assumed, highly dependent on the topic about which the person is writing.

### Reliability Studies

The three reliability studies sought to determine the degree to which the language people use is consistent across multiple writing

samples. Because the data from three separate samples were treated and analyzed in comparable ways, the method and results are presented together.

## Method

### Participants and Procedure

*Sample 1: Daily diaries by inpatients in an addiction treatment center.* Daily writing samples from 15 residential patients in a substance abuse and addiction treatment center were collected as part of a treatment program in England (Stephenson, Laszlo, Ehmann, Lefever, & Lefever, 1997). The participants, 11 women and 4 men, were consecutively admitted patients over a 3-month period; the primary diagnoses were alcohol-drug abuse ( $n = 8$ ), eating disorders ( $n = 4$ ), or a combination of these or other addictive behaviors, such as gambling ( $n = 3$ ).

As part of the treatment program, patients were asked to complete a "significant event sheet" at the end of each day wherein they were to write about the most significant events of the day. The sheets, which were typically written in a diary format, were turned into the counseling staff each morning and served as a gauge of each person's progress through the program. Because patients remained in the program for varying amounts of time, we analyzed only the first 18 writing assignments (the number of days for which we had writing samples for all 15 participants). It is important to note that the writing assignments began on the participants' first full day of treatment. Hence, one person's first day of writing was never another person's first day. Participants were aware that their writing samples would be read by the treatment staff and would be available to researchers, with the understanding that all data would be anonymous and confidential.

*Sample 2: Daily class assignments by Taos summer school students.* Between 1993 and 1996, the first author taught a 2-week summer school course on health psychology at the Fort Burgwin campus of Southern Methodist University (SMU) in Taos, New Mexico. Class sizes varied from 5 to 12 students. Thirty-four students (29 women and 5 men), ranging in age from 18 to 67 years ( $M = 26.4$ ,  $SD = 11.1$ ), completed daily writing assignments for 10 consecutive days, excluding weekends. The writing assignments were not graded per se, but failure to turn in a writing assignment resulted in a lower course grade. With the exception of one student (whose data were not included), all summer school students completed all writing assignments.

As part of the course requirement, students were assigned daily writing topics that were due the following day. Each day's writing assignment was typed on one of the university's computers, and a digital copy was given to the instructor. For all assignments, students were expected to write a minimum of 20 min per day. The specific assignments varied considerably from day to day and included the following: stream of consciousness writing, discussion of the students' own theory of mind and body, review of the biological bases of the stress response, significant childhood issues, the most positive experiences in their lives, the most negative experiences in their lives, thoughts after a relaxation homework assignment, thoughts after expressive movement homework, evaluation of a class trip to a local healer, and an overall course evaluation. The instructions varied from year to year, as did the order of several of the assignments. Nevertheless, the data were entered by day rather than by topic.

*Sample 3: Published abstracts by prominent social psychologists.* Forty social psychologists were randomly selected from the 690 names listed on the membership roster of the elite social psychology professional organization, SESP. To be included in the study, the prospective author had to have had at least 15 first-author journal publications between 1967 and March 1998 that were included in the PsycINFO database. If one of the randomly selected participants did not meet the criteria, adjacent names on the list were systematically examined. Because the majority of SESP members did not meet criteria, 71 names were excluded to provide a full

sample of 40 authors. The final sample was composed of 33 men and 7 women.

PsycINFO includes abstracts from more than 1,300 psychologically oriented journals in more than 20 languages. The vast majority of non-English journal articles also include English abstracts. In the present study, the 15 most recent journal abstracts from each author were downloaded and stored as separate text files. To be included in the study, the abstract had to meet each of the following criteria: written in English, SESP participant as first author, and standard journal article (i.e., no book chapters or journal corrections). Theoretical and empirical articles, comments, critiques, and replies to previous articles were included.

### Results

Each writing sample for each participant was transcribed into a computer text file and analyzed with the LIWC program. Basic psychometric information for each sample is listed in Table 1, and more detailed reliability statistics for all variables are included in the Appendix. For each of the 72 language dimensions, reliability analyses were computed across the writing samples within each sample of participants. In other words, percentages of pronouns for Writing Sample 1, Writing Sample 2, and Writing Sample 3 were correlated with each other across the participants in each group.

As can be seen in Table 1, the mean Cronbach alpha coefficient across all 72 language variables was .59. Indeed, of the 46 language composition variables that were used at least 0.1% of the time within any given study (e.g., pronouns and articles), approximately 80% had coefficients of .60 or higher. Of the 75 language composition alpha coefficients in the three studies, 60% were .60 or greater. In the more specific categories of relativity and current concerns, fewer reliabilities reached the .60 level (47% of relativity items and 50% of current concerns).

Table 1  
Summary of Reliability Studies

Study characteristic	Sample 1: Inpatients	Sample 2: Taos summer school	Sample 3: SESP abstracts
Number of writers	15	34	40
Writing samples per writer	18	10	15
Words per sample			
<i>M</i>	166.2	516.7	118.3
<i>SD</i>	114.2	265.5	35.3
Mean correlation	.15	.16	.12 <sup>a</sup>
Mean $\alpha$	.64	.56	.56 <sup>a</sup>
Percentage of language categories with Cronbach alphas $\geq .60^b$			
Overall	69% (70%)	49% (50%)	47% (59%)
Language composition	83% (83%)	83% (88%)	39% (69%)
Psychological process	72% (72%)	48% (48%)	60% (60%)
Relativity	60% (60%)	40% (40%)	40% (40%)
Current concerns	58% (57%)	21% (22%)	47% (70%)

Note. SESP = Society for Experimental Social Psychology.

<sup>a</sup> Based on 68 of the 72 Linguistic Inquiry and Word Count (LIWC) dimensions. Four categories—question marks, second person, swear words, and sleep words—had no variance. <sup>b</sup> Percentages in parentheses exclude LIWC categories that were never or rarely used (i.e., mean usage less than 0.1% of total words).

What is particularly striking about these data is the fact that word category usage is remarkably stable across time and writing topic. The inpatient sample, for example, wrote on consecutive days about very different topics as they were going through the first 2.5 weeks of intensive substance abuse treatment. Some days they wrote about family issues, other days they wrote about other inpatients, and yet others they wrote about mundane topics such as food. The summer school students wrote on assigned topics that did, in fact, range from very general (e.g., their personal philosophy of the links between mind and body) to highly specific and concrete (e.g., tracing the neural and hormonal pathways of the stress response). There were large and significant day main effects for the majority of categories, including word count, unique words, dictionary words, six-letter words, pronouns, self-references, articles, prepositions, positive and negative emotions, cognitive mechanisms, and social and physical word categories (all  $ps < .01$ ). Striking examples of the variations from day to day included self-references (highest day  $M = 11.4\%$ , lowest day  $M = 4.7\%$ ), articles (8.4% vs. 4.9%), positive emotions (3.3% vs. 2.1%), negative emotions (2.6% vs. 1.4%), and social words (12.8% vs. 6.6%). These differences simply illustrate that the types of words people used varied tremendously depending on the assigned topic.

In many ways, the most striking evidence for the reliability of language use involved the SESP sample. The abstracts for every author were written over several years. Unlike the more free-writing styles allowed in the inpatient and summer school samples, the implicit and explicit rules of journal abstract writing preclude certain word categories (e.g., swear words and self-references). For certain rarely used categories, then, reliabilities will necessarily be low. Similarly, the traditional content and current concerns categories focus on highly specific topics that may be the focus of only a few investigators. If one or more of the SESP authors were studying sleep, for example, the reliabilities of this category would have been quite high.

A potential criticism of the SESP study is that the reliabilities were artificially inflated because some members study affect, other members focus on cognition, and yet others focus on pure social processes. If this is the case, the language dimensions are merely demonstrating that, over their career, the same people are studying the same topics. Closer inspection of the data, however, tends to argue against this. For example, among the 40 SESP members randomly selected were Paul Ekman and Myron Rothbart. Ekman, of course, is well known for his work on emotion and Rothbart for his contributions to social cognition. Averaging across their 15 abstracts, Rothbart used slightly more negative emotion words than Ekman (1.7% vs. 1.5%) and a comparable number of cognitive words (6.1% vs. 6.0%). Ekman did, in fact, use more positive emotion words than Rothbart (3.9% vs. 1.4%). However, even Ekman's use of positive emotion words was eclipsed by others, such as Shalom Schwartz (5.2%) and Leigh Thompson (4.9%), whose contributions are not automatically associated with either emotion or positive mood.

### Discussion

Taken together, the three reliability studies offer encouraging evidence that language use is a reliable individual difference. Across writing instructions and topics with varying degrees of

imposed constraint, reliability coefficients were respectably high. As is apparent, reliability coefficients vary by language dimension and by study. Furthermore, these patterns raise some intriguing questions about the nature of linguistic styles and the measurement of individual differences in general.

The LIWC program was not originally designed to examine the stability of writing dimensions across individuals. It is not surprising, then, that some linguistic categories are more internally consistent than others. As with any individual-difference measure, one would assume that the more specific the response category, the less reliable the category across response domains. Indeed, across the three studies, the highly specific current concerns categories had the lowest mean correlations and reliabilities. As depicted in Table 1, only 5 of the 19 current concerns categories (26% or, if the more liberal criterion is used, 37%) had alpha coefficients averaging .60 or higher. For the broadest category, language composition variables, 12 of the 18 categories (67% or, with the more liberal criterion, 83%) evidenced coefficients of at least .60. The psychological processes and relativity categories, which could be viewed as moderately specific relative to the other two, demonstrated midrange percentages of coefficients greater than .60 (56% and 40%, respectively).

Although the results of the reliability studies indicate that linguistic styles can function psychometrically as stable individual differences, an equally important question concerns the basic factor structure of word usage and linguistic styles. To begin to address these issues, we undertook a second large-scale project to examine word use among more than 800 introductory psychology students who completed four writing samples over the course of a semester.

## PHASE 2: EXPLORING THE FACTOR STRUCTURE OF LANGUAGE USE

The first research phase provided a general perspective on the reliability of multiple dimensions of language. Phase 2 sought to examine the 15 most promising language dimensions identified in the first studies and assess the structure of these dimensions with much larger samples. Although a limited number of studies have attempted to factor analyze language use (e.g., Biber, 1988), their results have generally distinguished writing styles across different genres (e.g., newspaper stories vs. romance novels) rather than structure within comparable writing samples.

Two strategies to the factor analyses were adopted. Using the first sample of more than 800 introductory psychology students from 1992–1997, we were able to establish the basic factor structure of the 15 language dimensions. The writing samples of a second group, of 381 introductory psychology students in the fall 1998 class, were factor analyzed to determine the degree to which the original factor structure was confirmed.

### Method

#### Participants

*Original student samples.* During the fall semesters of 1992 through 1997, students enrolled in the first author's introductory psychology class were required to turn in four writing assignments: two during the first 2 weeks of classes and another two during the last 2 weeks of classes. The writing requirement accounted for 10% of the final grade, with the grade

being dependent only on turning in the assignment; content was not considered. From 1992 through 1996, students were in classes of less than 100 at SMU. In 1997, the students were enrolled in a large class of 485 students at the University of Texas (UT).

Only students who turned in at least two of the four writing assignments were retained for the analyses (approximately 94% of the enrolled students). The numbers of students turning in at least two writing assignments by year were as follows: 1992, 84; 1993, 55; 1994, 62; 1995, 64; 1996, 90; and 1997, 467. Sex was identifiable from school records or self-reports for 96.5% of the 822 students. Of these students, 478 (60.3%) were female and 315 (39.7%) were male. Across the 6 years, all four essays were available for 72.8% of the students, three essays were available for 9.3%, two essays were available for 16.0%, and one essay was available for 1.9%. Of the possible 387 essays that were not turned in, approximately 30% were attributable to students dropping the course; 20% were due to problems with computers, disks, word-processing programs, or other technical difficulties; and the remaining approximately 50% were simply not turned in. Twenty-three additional essays were excluded from the analyses because students wrote only one of the assigned four essays ( $n = 16$ ) or failed to write at least 30 words on at least two of their essays ( $n = 7$ ), suggesting that they were not taking the assignment seriously.

*Sample for confirmatory analyses.* All confirmatory analyses were based on the fall 1998 introductory psychology course taught by the first author at UT. Of the 386 students enrolled in the course, 381 turned in at least two usable essays. The final sample was 52.2% female, and 70.2% were 1st-year students; the mean age was 18.7 ( $SD = 2.13$ ) years. Overall, 87.2% turned in all four essays, 6.3% turned in three, and 6.0% turned in two. Data from 2 students who turned in only one assignment and an additional 2 who failed to turn in at least two assignments with at least 30 words were excluded.

### Procedure

On the 1st day of class for all samples, students were told about the course writing requirement. The first two writing assignments—which were to be completed on separate days—were to be handed in by the end of the 2nd week of classes. The first assignment was described as writing in a stream of consciousness mode. Its purpose was to give students a sense of how their mind moved from topic to topic. The second assignment was for people to write about their thoughts and feeling about coming to or being in college. The purpose of the coming to college essay was for students to introspect about their new lives in college.

Between 1992 and 1996 (the SMU samples), all writing assignments were turned into the instructor on diskettes in text format. In the 1997 original and 1998 confirmatory samples collected at UT, students were given an address on the Internet. Once students logged onto the Internet site, the specific writing assignment was presented on the computer screen, along with a form on which they were to write their assignments. When they had completed the assignment, they simply pressed the "submit" button on the screen, which sent the essay to the instructor's account. For the SMU samples, the specific writing instructions were included in a handout; for the UT samples, the instructions were posted on the Web site. The stream of consciousness instructions for all samples were as follows:

For this assignment you should think about what your thoughts, feelings and sensations are at this moment. Write about your thoughts as they come to you. If you prefer, you can write your thoughts and feelings on paper and then transcribe them onto this form. Your goal in this assignment is to reveal in your writing the way your mind works naturally. The important thing is for you to write continuously and don't worry about grammar, sentence structure, or spelling.

The coming to college instructions read:

Express in your writing what it has been like for you coming to college, and explore your thoughts and feelings of being in college in

general. How does this relate to other aspects of your life? For example, how does being in college relate to your childhood, your relationships with family, friends, etc.? How does it relate to who you are, who you have been in the past, and who you want to be in the future? The goal of this assignment is to explore the psychological experience of being in college. It is important for you to write continuously and don't worry about grammar, sentence structure, or spelling.

It was emphasized in class that students would not receive any feedback on their writing and that, in fact, their essays might not be read for months. Hence, any expressions of concern about their class performance, mental health, and so forth should be made directly to the instructor. Students were also informed that their essays would be analyzed as part of a larger research project dealing with "the nature of language" and that other researchers would have access to the writing samples but with all identifying information (e.g., name and social security number) removed.

Two weeks before the semester ended, the instructor announced the requirements for the last two writing assignments. Students were again asked to write in a stream of consciousness mode and about coming to college for 20 min each on 2 separate days. Writing assignments were due on the last day of classes.

## Results

### Factor Analyses for the Original Samples

Each essay for each participant was analyzed with the LIWC program. On the basis of a number of considerations, 15 of the original 72 variables were retained for subsequent factor analyses. First, only those variables from the Phase 1 studies with mean reliabilities of .60 or greater were deemed eligible for inclusion (see Appendix). Second, categories were included only if they did not substantially overlap with other included categories. For example, large percentages of inclusive and exclusive words were

prepositions. Prepositions, in this case, were not included. Similarly, first-person singular (*I*, *me*, and *my*) words were used rather than total first person because first-person plural (*we*, *us*, and *our*) overlaps with total social words. Third, linguistic categories that did not refer to features or meanings of specific words (e.g., total word count, words per sentence, or dictionary words) were excluded. Similarly, all of the current concerns variables were excluded on the basis that they were more topic relevant than process related. Finally, only those remaining LIWC variables that had mean usage levels of at least 1% per essay were included. The means and individual reliability coefficients across topics and time for the resulting 15 LIWC variables meeting each of the preceding criteria are listed in Table 2.

It is of interest to compare the mean interessay correlations for the present study with the comparable means in the three reliability studies (see Appendix). The mean interwriting sample correlation ( $r$ ) for the 15 language variables in the Phase 1 studies was .20, which was significantly lower than the mean interitem correlations in the present study (overall  $r = .29$ ),  $t(18) = 4.16$ ,  $p < .001$ . Note also that Table 2 includes the Cronbach alpha coefficients based on the four writing samples. The coefficients for these variables were lower than those found in Phase 1 ( $M = .58$  vs.  $.73$ ); these patterns are not surprising, however, given the limited number of writing samples.

To examine the underlying dimensions of the 15 LIWC variables, we conducted an exploratory factor analysis on the mean of the four writing samples. Diagnostic tests indicated that a factor model was appropriate for the data ( $KMO = .633$ , Bartlett's test of sphericity = 2,831,  $p < .001$ ). In addition, the measures of sampling adequacy were reasonably high, ranging from .26 for present tense to .86 for negative emotion. Thus, although these data are quite a bit different from the typical factor analytic data, a factor analysis was deemed appropriate. Examination of the scree plot indicated that a

Table 2  
Psychometric Information for Factor Analysis Sample

Dimension	<i>M</i>	<i>SD</i>	Reliability coefficient			
			Stream of consciousness	Coming to college	Cross-topic	$\alpha$
Words of more than 6 letters	13.06	2.26	.35	.40	.41	.74
First-person singular	10.63	1.98	.46	.33	.38	.69
Negations	2.18	0.61	.21	.22	.25	.51
Articles	4.73	1.03	.33	.26	.29	.60
Positive emotions	3.38	0.81	.16	.21	.14	.45
Negative emotions	1.80	0.69	.28	.18	.27	.53
Causation	1.10	0.42	.16	.22	.16	.45
Insight	2.47	0.67	.23	.12	.17	.44
Discrepancy	2.84	0.80	.29	.22	.23	.53
Tentative	2.84	0.77	.25	.20	.18	.47
Social processes	6.51	1.83	.30	.15	.21	.49
Past tense	3.79	1.31	.19	.19	.21	.44
Present tense	13.95	2.26	.35	.29	.32	.63
Inclusive	5.95	1.04	.25	.30	.25	.57
Exclusive	4.21	0.97	.30	.27	.36	.61

*Note.* All means are expressed as percentage of total words within the essays ( $n = 822$ ). The stream of consciousness ( $n = 677$ ) and coming to college ( $n = 626$ ) columns refer to the reliability of the language domains from the beginning of the semester to the end of the semester. Cross-topic correlations ( $n = 743$ ) are based on the mean of the two cross-topic correlations for each of the two times. A correlation of .073 or higher was required for significance ( $p \leq .05$ , two-tailed).  $\alpha$  refers to Cronbach alpha coefficient for the four writing assignments ( $n = 610$ ).



Table 3  
*Rotated Factor Loadings for Exploratory Analysis of LIWC Dictionaries*

Dictionary	Factor 1: Immediacy (22.4% variance)	Factor 2: Making Distinctions (10.3% variance)	Factor 3: The Social Past (9.8% variance)	Factor 4: Rationalization (8.6% variance)
First-person singular	.823			
Articles	-.765			
Words of more than 6 letters	-.683			
Discrepancies	.485	.427		
Exclusive		.674		
Tentativity		.644		
Negations		.579		
Inclusive		-.463		
Past tense			.856	
Present tense	.593		.596	
Positive emotion			-.469	
Social			.425	
Insight				.627
Causation				.598
Negative emotion				-.443

Note. Only loadings of .20 or above are shown.  $N = 838$ . LIWC = Linguistic Inquiry and Word Count.

four-factor or seven-factor solution would best fit the data (six factors had eigenvalues above 1). Principal-components analysis extracted four factors, and varimax rotation was used to facilitate interpretation of the factors. For the four-factor solution, inclusive words and negative emotions had communalities of .26 and .31. All other variables had communalities above .36.

Rotated factor loadings are shown in Table 3. Dictionaries loading on the first factor (eigenvalue = 3.35) included the total number of first-person singular words (*I*, *me*, and *my*), fewer articles (*a*, *an*, and *the*), fewer long words, present-tense verbs, and discrepancies (e.g., *would*, *should*, and *could*). This factor was labeled Immediacy (cf. Mehrabian's, 1967, verbal nonimmediacy construct). The second factor (eigenvalue = 1.55) included such dictionaries as exclusive words (e.g., *but*, *without*, and *except*), tentative words (e.g., *perhaps*, and *maybe*), negations (e.g., *no*, *not*, and *never*), fewer inclusion words (e.g., *and* and *with*), and a secondary loading on discrepancies. This factor was termed Making Distinctions. The third factor (eigenvalue = 1.47) was characterized by high use of past tense verbs, fewer present tense verbs, fewer positive emotion words, and high social references. Thus, the third factor was termed The Social Past. Finally, the fourth factor (eigenvalue = 1.29) was characterized by causation words (*because* and *reason*), insight words (*understand* and *realize*), and fewer negative emotion words. This factor was termed Rationalization.

Factor scores were computed by averaging the standard scores for the dictionaries within the factors. Correlations were then computed among the factor scores. Factor 1 (Immediacy) was positively correlated with Factor 2 (Making Distinctions;  $r = .51$ ,  $p < .001$ ) and negatively correlated with Factor 3 (The Social Past;  $r = -.45$ ,  $p < .001$ ). Factor 3 was modestly correlated with Factor 4 ( $r = .09$ ,  $p < .02$ ) and negatively correlated with Factor 2 ( $r = -.23$ ,  $p < .01$ ).

#### *Congruence of Factor Analyses for Confirmatory Sample*

There are multiple ways to determine the degree to which factor structures are replicable across independent samples. Sim-

ple principal-components factor analysis with varimax rotation yielded four factors (based on scree analyses) that were virtually identical to the factors from the original sample. Coefficients of congruence (from Harmon, 1976) for the four pairs of factors ranged from .83 to .99, in comparison with the off-factor coefficients, which ranged from  $-.30$  to  $.34$  ( $M = .012$ ), indicating a high degree of congruence from the original to the confirmatory factor. Similarly, the computed factor scores based on the unit-weighted algorithm from the original sample were correlated with the factor analytically derived factors from the confirmatory sample (a confirmatory strategy recommended by Tabachnick & Fidell, 1996). The correlations for the four factors were as follows: Factor 1, .93; Factor 2, .91; Factor 3, .83; and Factor 4, .75 (all  $dfs = 379$ ,  $ps < .001$ ).

Finally, LISREL confirmatory factor analyses indicated that 23% and 27% of the residuals were significantly greater than zero for the original and confirmatory samples, respectively. This implies that the language data were not appropriately suited to the LISREL analyses.

#### PHASE 3: VALIDITY STUDIES

Results thus far indicate that language use is a reliable individual difference. In addition, the factor structure of the LIWC dictionaries is reasonably robust. An obvious question remains: What do these dictionaries mean, psychologically, and what might this information provide that is different from or complementary to other types of assessment? The following studies were conducted to provide an initial examination of the relations of the LIWC language dimensions first to motive measures and then to trait measures.

This inquiry into the validity of the LIWC scales presents a dilemma. Although, traditionally, validation of a measure is conducted by looking for relations with measures of theoretically relevant constructs (or the same construct measured in a different way), the validity studies presented here were, by necessity, exploratory. Although word usage is a highly reliable phenomenon, the psychological meaning of this phenomenon is unknown. We

cannot make strong predictions about the relations of this measure to other constructs because the psychological constructs that underlie language usage are certainly not obvious. Rather than making specific predictions about the relations we expected to emerge, we conducted these studies with a general hypothesis: Language use, as a reliable individual difference, is a behavior that ought to relate to personality characteristics. In accord with the projective hypothesis, we assume that the way people talk or write is an expression of underlying psychological processes or characteristics. The following studies sought to examine what those process or characteristics might be.

### Language Use and Measures of Motivation

Comparing LIWC with more established measures of motivation is particularly intriguing because a long tradition of research and theory has made use of content-analytic procedures to measure human motivation (cf. Smith, 1992). This research has relied on thematic analysis of imaginative stories as a means of revealing the nonconscious motives that orient attention and drive behavior (McClelland, 1985). Two problems that have plagued this area of inquiry are the unreliability of TAT measures and the effort and time required to do the coding. Although it is doubtful that a word count strategy could ever totally replace the use of human raters, it remains an empirical question whether computer word counts might provide information about human motivation that is comparable to that found via content analysis (Schnurr, Rosenberg, Oxman, & Tucker, 1986). Certainly, the idea that the way people express themselves in language provides psychologically important information is a central assumption of both the traditional TAT approach and the LIWC strategy presented here. In addition, the LIWC dictionaries have already proven themselves to be quite reliable across topics and testing occasions. The tendency of TAT motive scores to vary across stimulus pictures has been discussed by a variety of authors (e.g., Atkinson, 1982).

The purpose of the first study in this phase of research, then, was to compare the LIWC factors derived from TAT protocols that had previously been rated by judges for the three TAT motives of achievement, affiliation, and power. Furthermore, this data set included self-reports of the motives as well as self-reported and actual behavioral markers of the motives collected over the course of a semester. The project, then, was able to establish the links among LIWC scores and each of the other measurement domains.

### Method

#### Participants

Participants were 62 women and 17 men who volunteered to earn extra credit in a psychology class at the University of California, Davis.<sup>1</sup> Ages ranged from 18 to 50 years ( $M = 20.99$ ,  $SD = 4.61$ ). Data were collected throughout the course of the academic term.

#### Procedure

Participants completed a version of the TAT (Morgan & Murray, 1935) using the slides of the pictures used by McClelland and colleagues in a number of studies (McClelland, 1985). The pictures in this version of the test include a man and a woman sitting by a river, a man sitting at an office desk, two women working in a laboratory, an older man and a younger woman walking with horses, and a man and a woman on a trapeze (cf.

Smith, 1992). TAT sessions were held in a small seminar room; 1 to 7 participants took part in a given session. Instructions followed Lundy's (1988) recommendation to create a relaxed atmosphere and to use very loose (although standardized) instructions. Participants were shown a slide of each picture for 10 s and then had 10 min to write stories.

During the term, participants also completed the three 16-item scales from the Personality Research Form (PRF; Jackson, 1974) pertaining to achievement, affiliation, and dominance. On the PRF, individuals judge, on a true-false scale, the extent to which items describe them. The reliability and validity of the PRF scales were discussed by Jackson (1974).

*Criteria for motive-relevant behavior.* Motive-relevant behavior was measured through self-report and observation of classroom behavior. First, participants completed a questionnaire assessing a number of motive-relevant criteria. Activities and preferences that have been demonstrated to show relationships with the motives, as well as some that appear to have face validity with regard to the motives, were included. Two criteria per motive were included. For achievement, participants were asked whether they preferred to have their grade based on their own work or based on their work relative to that of others in the class (if it had no bearing on their final grade) and whether they would be interested in a task involving teaching themselves moderately difficult computer programming skills. For affiliation, participants recorded how many friends they came to class with on the 1st day of class, and they were asked whether they would prefer the class to have solo or group projects. For power, participants were asked whether they had ever run for or attained a campus office and whether they would be interested in learning to perform hypnosis.

*Classroom behavior.* The number of times a student spoke in class was recorded on a seating chart on 8 days throughout the term. These charts were kept by the class teaching assistant, who was not informed of the purpose of the study but was told that participation was part of the students' grades. The teaching assistant kept track of three types of participation: questions, comments, and answers. Additional research assistants, who were given a brief explanation of the purpose of the study but were unaware of the identities of the participants or their motive scores, attended class on 2 days during the terms and rated participation at that time. These research assistants scored only 1 point for every instance of participation regardless of the type of participation. Agreement between the ratings of the research assistants and the teaching assistant for those days when multiple ratings occurred was 98.5% for the summed categories. The teaching assistant's ratings were used as the final scores for the students, each of whom was given a summed participation score. In addition, class attendance was recorded on daily seating charts, and the final grade in the course was obtained from class records.

*Motive coding.* Two trained coders, having demonstrated the ability to score within 96% of expert coding for TAT stories and 85% for running text scoring, scored the stories according to Winter's (1989) running text scoring procedure for social motives. The scoring system has been shown to converge with the original scoring systems and to successfully discriminate target groups for each motive (Winter, 1989). Overall interrater reliability according to Spearman-Brown estimations was .82. Individual reliabilities were .70 for achievement, .99 for affiliation, and .76 for power. The coders met and recoded each story together. Any disagreements were resolved by discussion. Next, the total score for each motive was divided by the total number of words in the protocol, and then this proportion was multiplied by 1,000. Thus, the TAT scores referred to motive imagery per 1,000 words (Winter, 1989). Finally, the TAT stories were transcribed into text files and run through the LIWC system, and factor scores were computed as described earlier.

<sup>1</sup> Although the TAT data were originally reported in King (1995), the LIWC analyses of this sample are new.

Table 4  
Correlations Among LIWC Factors and TAT and PRF Measures

LIWC factor	TAT			PRF		
	Need for achievement	Need for affiliation	Need for power	Achievement	Affiliation	Dominance
Immediacy	-.33**	-.04	.04	.01	-.03	-.13
Making Distinctions	-.18	-.18	.12	.13	-.30**	.02
The Social Past	.23*	-.25*	.15	-.01	-.21†	-.01
Rationalization	-.20†	.10	-.03	-.16	-.12	-.14

Note.  $N = 79$ . LIWC = Linguistic Inquiry and Word Count; TAT = Thematic Apperception Test; PRF = Personality Research Form.

†  $p < .10$  (marginally significant). \*  $p < .05$ . \*\*  $p < .001$ , two-tailed.

### Results and Discussion

First, correlations were computed between the LIWC and the TAT and PRF measures of motivation. Correlations among the TAT and PRF measures have been previously reported by King (1995). Results are shown in Table 4. With regard to the TAT, the LIWC factors were most systematically correlated with need for achievement. Immediacy was significantly negatively correlated with achievement motivation. The Social Past was significantly positively correlated with achievement, and Rationalization was negatively correlated with achievement, with marginal significance. Talking about the social past was significantly negatively correlated with affiliation motivation. No significant relations emerged for need for power. With regard to the PRF scores, fewer systematic relations emerged. Making Distinctions and talking about The Social Past were negatively correlated with self-reported affiliation motive.

Next, correlations were computed between the LIWC and other measures and the self-reported criterion behaviors. Results are shown in Table 5. With regard to achievement, Immediacy and Rationalization were negatively correlated with interest in teaching oneself computer programming. PRF achievement was marginally

correlated with having one's grade determined only by one's own performance. No other significant relations emerged. For affiliation, the Rationalization factor and TAT affiliation were both significantly positively correlated with preference for working on a group project rather than alone. Immediacy was marginally related to coming to class with friends on the 1st day, whereas TAT affiliation was actually negatively related to this affiliative criterion. None of the power measures related to these criteria.

Because of the relations that emerged between the LIWC factors and the TAT achievement scores, we created a LIWC achievement motive measure by averaging the LIWC dictionaries along the lines of their correlations with need for achievement (see Table 4). The standard scores for the factors were averaged; The Social Past was weighted positively, and Immediacy and Rationalization were weighted negatively. This new composite LIWC-derived achievement motive measure correlated significantly with TAT-measured achievement ( $r = .44$ ,  $p < .001$ ) but was unrelated to the PRF measure of achievement ( $r = -.05$ ).

To establish whether this LIWC-derived measure might relate to criteria independent of the hand-scored TAT measure of achieve-

Table 5  
Correlations of LIWC Factors and TAT and PRF Measures With Self-Reported and Behavioral Criteria

Measure	Achievement		Affiliation		Power		Classroom behaviors		
	Grading preference	Teaching oneself a skill	Number of friends on first day	Group (vs. solo) project	Sought political office	Interested in learning hypnosis	Participation	Missed classes	Final grade
LIWC factor									
Immediacy	-.13	-.30**	.25†	-.24†	.06	.06	.40**	.16	.30*
Making Distinctions	-.02	-.08	.13	-.22	-.15	.15	.34*	-.09	.15
The Social Past	.21	.25†	-.04	-.03	-.22	-.09	-.18	-.13	-.12
Rationalization	-.12	-.33**	.23	.37*	.00	.00	-.13	.10	.21
TAT measure									
Need for achievement	.01	.21	-.21	.04	.00	-.03	-.14	-.12	-.08
Need for affiliation	-.01	-.14	-.24†	.36**	.25†	-.11	.01	-.06	.04
Need for power	.14	.06	.08	.01	-.08	-.05	.11	.33*	.06
PRF measure									
Achievement	.24†	.18	-.21	.03	-.04	-.05	.15	-.01	-.08
Affiliation	-.08	-.28*	.06	-.13	.28*	-.12	.31*	-.08	-.10
Dominance	.15	.06	-.15	.19	.13	-.04	-.07	-.01	-.08

Note.  $N = 57$ . LIWC = Linguistic Inquiry and Word Count; TAT = Thematic Apperception Test; PRF = Personality Research Form.

†  $p < .10$  (marginally significant). \*  $p < .05$ . \*\*  $p < .001$ , two-tailed.

ment, we examined the relation of this new composite with the achievement criteria. First, as a means of improving the reliability of the achievement criterion measure, the two items were averaged together (interitem  $r = .55, p < .001$ ). Correlations were computed among the LIWC achievement motive measure, the TAT and PRF, and the criterion composite. The LIWC measure was significantly correlated with self-reported achievement behavior ( $r = .42, p < .004$ ). The PRF measure was marginally correlated with the achievement measure ( $r = .23, p < .10$ ); need for achievement was not related to the achievement criterion ( $r = .13, p > .30$ ). Next, a regression equation was computed regressing the achievement motive criterion on the LIWC achievement measure, the TAT measure, and the PRF measure. Both the LIWC measure and the PRF contributed significantly and positively to the equation (standardized  $\beta = .42$  for the LIWC,  $p < .006$ ;  $\beta = .29$  for the PRF,  $p < .05$ ), multiple  $R = .54, R^2 = .30, F(3, 42) = 5.88, p < .002$ . The TAT measure did not significantly add to the prediction of the composite ( $\beta = .06, ns$ ). Thus, the LIWC factor composite predicted the achievement criteria over and above the TAT and PRF measures of achievement.

Finally, correlations were computed between the LIWC factor scores, the TAT motives, and the PRF scores and classroom behavior. Results are shown in Table 5. Overall, few significant relations emerged. Immediacy, Making Distinctions, and PRF affiliation were significantly related to increased participation in class. TAT-measured need for power was positively correlated with missing class. Immediacy was also positively related to final course grade.

These results provide a fascinating glimpse of the psychological processes that are revealed in language use. In particular, some evidence for the relation of LIWC factor scores to achievement motivation clearly emerged. It is interesting to note that the coding instructions for need for achievement in the running text system include scoring for positive evaluations of performance, positive emotion related to excelling, and negative emotion related to failing to excel. None of these coding dimensions are reflected in the LIWC dictionaries; indeed, the kinds of contingencies required for TAT coding are impossible to detect with a word count strategy. Yet, the LIWC-derived measures of achievement motivation related to a measure of achievement preferences over and above the contribution of TAT or self-report measures.

### Language Use, Traits, and Demographic Variables

The overlap of LIWC factors with traditional measures of motivation is important in demonstrating the relative links to behaviors as well as to the judge-derived measures themselves. In evaluations of the LIWC approach, it is also critical to compare LIWC factors with standard self-reported trait and other individual-difference markers. Specifically, we sought to learn how the LIWC variables were correlated with the five-factor measure of personality, various demographic measures, several health and health-related behavior markers, and assorted commonly used self-report personality instruments.

### Method

#### Participants

LIWC and other data were collected from the participants described in Phase 2. The LIWC analyses were based on 1,203 introductory psychology

students: 822 students from 1992–1997 and 381 students from the 1998 confirmatory sample.

### Procedure

At both SMU (beginning in 1993) and UT, all students enrolled in introductory psychology were given the opportunity to complete a battery of questionnaires during the 1st week of classes. The actual questionnaires that were administered differed from semester to semester. In addition, the instructor had students complete occasional questionnaires to illustrate points (e.g., the workings of the five-factor personality dimensions and attitude and behavior questionnaires to demonstrate how correlations worked) or for extra credit. At no time were students preselected to complete questionnaires. Among the questionnaires that were used in the present study were the Five Factor Inventory (John, Donahue, & Kentle, 1991), administered between 1993 and 1998; the Positive and Negative Affect Schedule (PANAS), which taps both positive affect and negative affect (Watson, Clark, & Tellegen, 1988), administered in 1993–1995 and 1997; the Rosenberg (1979) Self-Esteem Scale, administered in 1993, 1994, 1996, and 1997; the Need for Cognition Scale (Cacioppo & Petty, 1982), administered in 1997; and self-reports of alcohol, smoking, and related health behaviors within the previous week as part of the College Attitudes and Behavior Questionnaire (Pennebaker, Colder, & Sharp, 1990) in 1994 and 1997. The following data were collected as well: number of health center visits for illness in the year of the students' participation in the course (1992 and 1995–1997), self-reported Scholastic Aptitude Test (SAT) scores (1994 and 1997), parents' education (1997), and participants' sex, age, and exam performance (1992–1998). Note that health center data were released only for students who completed the official release form at the beginning of the semester (approximately 95% of students signed the form each year).

### Results

#### Correlations With the Five-Factor Personality Dimensions

Table 6 lists the correlations between the language dimensions and the five personality factors completed by the participants. These correlations point to modest but reliable effects of personality on language use. Inspection of Table 6 indicates that three patterns stand out. First, virtually all of the Immediacy variables are negatively correlated with Openness to Experience. That is, the more immediate and simple people's writing, the lower they rate themselves on openness. Second, individuals whose writing is high on Making Distinctions tend to be less extraverted (or, by extension, more introverted). Third, self-ratings of conscientiousness tend to be negatively correlated with Making Distinctions.

Many of the systematic effects that appear in Table 6 would be predicted by the very definitions of the five factors. For example, Neuroticism was positively correlated with use of negative emotion words ( $r = .16$ ) and negatively correlated with positive emotion words ( $r = -.13$ ). Extraversion, as would be predicted, correlated with positive emotion words ( $r = .15$ ) and total social references ( $r = .12$ ). Agreeableness was positively correlated with positive emotions ( $r = .07$ ) and negatively correlated with negative emotions ( $r = -.07$ ). Note that although these correlation coefficients were quite modest, they were all statistically significant given the sample size of 841 participants ( $p < .05$ , two-tailed tests).

Table 6  
*LIWC Factors and Simple Correlations With Five-Factor Scores*

LIWC factor	Five-factor dimension				
	Neuroticism	Extraversion	Openness	Agreeableness	Conscientiousness
<b>Immediacy</b>	<b>.10*</b>	<b>.04</b>	<b>-.16**</b>	<b>.07*</b>	<b>-.02</b>
First-person singular	.13**	.04	-.13**	.07*	.01
Articles	-.09*	-.09*	.13**	-.15**	-.04
Words of more than 6 letters	-.03	-.04	.16**	-.03	.06
Present tense	.06	.01	-.15**	.04	.00
Discrepancies	.05	-.03	-.01	-.02	-.07*
<b>Making Distinctions</b>	<b>.05</b>	<b>-.14**</b>	<b>.06</b>	<b>-.05</b>	<b>-.13**</b>
Exclusive	.00	-.08*	.10*	-.06	-.08*
Tentativity	.06	-.14**	.11**	-.02	-.06
Negations	.05	-.12**	.00	-.04	-.15**
Inclusive	-.01	.07*	.01	.03	.06
<b>The Social Past</b>	<b>.04</b>	<b>.00</b>	<b>.08*</b>	<b>-.02</b>	<b>-.04</b>
Past tense	.03	.04	-.03	.06	-.06
Social	-.01	.12**	.02	.00	.02
Positive emotion	-.13**	.15**	-.06	.07*	.07*
<b>Rationalization</b>	<b>-.06</b>	<b>.02</b>	<b>-.03</b>	<b>.07</b>	<b>.04</b>
Insight	.03	-.02	.07*	.05	-.01
Causation	.03	-.08*	-.08*	.00	-.07*
Negative emotion	.16**	-.08*	.05	-.07*	-.15**

Note.  $N = 841$ . Two variables are coded on two factors: Present tense is also a part of The Social Past; discrepancy is also part of Making Distinctions. The following variables are negatively loaded on their respective factors: articles, words of more than 6 letters, inclusive, present tense (for The Social Past only), and negative emotion. LIWC = Linguistic Inquiry and Word Count.

\* $p < .05$ . \*\* $p < .001$ , two-tailed.

### *Correlations With Other Demographic, Health, and Individual-Differences Variables*

As can be seen in Table 7, the LIWC factors correlated in various ways with sex, age, parental education, self-reported SAT scores, and course test grades. Interestingly, the Immediacy dimension was consistently related to each of these dimensions. That is, people who were high in Immediacy were more likely to be female and young, to have lower SAT verbal scores and exam grades, and to have parents with lower levels of education.

Particularly intriguing are the correlations with the various illness-related indexes. Three of the health variables involved self-reports (alcohol consumption, cigarette smoking, and number of physician visits for illness in the previous month), and one, health center visits for illness, involved data released by the student health services of both SMU and UT. Because of the differing ways of defining health center visits across the two institutions and across years, health center visits were converted to binary data (0 = did not visit the health center and 1 = did visit at least once during the school year for illness). As shown in Table 7, each of the illness-related behaviors was negatively correlated with the Making Distinctions category. That is, the more that people made distinctions in their writing, the better their health or health-related behaviors (less alcohol, fewer cigarettes, and fewer doctor visits). It should be noted that this pattern is consistent with recent reanalyses of our disclosure research, which also indicates that the more people make distinctions over the course of writing, the better their subsequent physical health (Pennebaker, 1997).

Finally, Table 7 includes correlations of the LIWC factors with selected individual difference measures commonly used in social and personality psychology. Consistent with the five-factor data, individuals high in positive affect made fewer distinctions, and those high in negative affect were lower on the Rationalization factor. Those high in need for cognition tended to write in a less immediate fashion (perhaps attempting to sound smarter, using bigger words), and those high in self-esteem scored higher on the Rationalization dimension.

### *Predicting Illness-Related Behaviors With LIWC and Five-Factor Models*

An important question addressed earlier with the TAT analyses was the degree to which the LIWC factors were comparable to more traditional measures in predicting significant behaviors. Of particular interest is the degree to which LIWC factors can predict health and illness-related behaviors relative to the more traditional five-factor model. Thus, the four illness-related behaviors listed in Table 7 were standardized and averaged to yield an illness-behavior index. A multiple regression analysis was then computed in which the illness index was used as the dependent measure. In the forced-entry regression, the five-factor scores were entered initially, followed by the block entry of the four LIWC factors. This allowed us to assess the independent contribution of the two classes of predictor variables.

Overall, 469 students had complete data sets including the illness index, the five-factor measures, and the four LIWC factors. The five-factor and LIWC variables significantly and indepen-

Table 7  
*LIWC Factor Correlations With Demographic, Illness-Related,  
 and Selected Self-Report Variables*

Variable	LIWC factor dimension				<i>n</i>
	Immediacy	Making Distinctions	The Social Past	Rationalization	
Demographic characteristics					
Sex	.22**	.03	.06	.04	1,176
Age	-.07*	-.05	.00	.02	985
Parents' education	-.16*	-.15*	.01	-.06	296
SAT verbal	-.19**	.06	.04	-.05	266
Exam grades	-.09*	.04	-.02	-.03	1,144
Total friends	-.08	-.17**	.00	-.10	344
Illness-related behaviors					
Alcoholic beverages	-.11*	-.13*	.03	-.07	345
Cigarettes	-.08	-.08	.02	-.07	345
Health center visits	-.01	-.09*	.03	-.04	617
All medical visits (self-report)	-.07	-.20**	.10	-.03	363
Individual differences					
Positive Affect	.03	-.12*	.05	.02	462
Negative Affect	.01	.01	-.05	-.14*	462
Need for Cognition	-.15*	-.02	.09	.05	386
Self-esteem	.03	.02	-.08	.10*	499

Note. LIWC = Linguistic Inquiry and Word Count; SAT = Scholastic Aptitude Test.

\* $p \leq .05$ . \*\* $p \leq .001$ , two-tailed.

dently predicted illness behaviors. When the five factors were entered first, they accounted for 2.9% of the variance,  $F(5, 463) = 2.79$ ,  $p = .02$ , with the LIWC variables adding 3.3%,  $F(4, 459) = 4.04$ ,  $p = .003$ . Virtually identical results occurred when the order of independent variables was switched, attesting to their true independence. Inspection of the full regression model indicated that two five-factor beta coefficients and one LIWC coefficient were statistically significant (Openness, .09; Agreeableness, -.13; and Making Distinctions, -.18).

### Discussion and Summary of Validity Results

The data in this phase of the research indicate that the language factors are uniquely related to judges' ratings, self-reported personality, and meaningful behaviors. In the study that relied on the TAT responses, it was striking to see that text analyses of the TAT essays predicted TAT-related behaviors beyond what the judged TAT protocols were hypothesized to predict. In many ways, this is striking given that the LIWC domains were never designed to tap any of the motivational states associated with the TAT.

Relationships between the four LIWC factors and the five-factor self-report instrument provided further insight into the relative independence of these two classes of measures. Although selected five-factor variables were significantly correlated with their presumed linguistic referents, the absolute magnitudes of the relationships were often extremely low. For example, Neuroticism correlated only .16 with negative emotion word usage. More striking was that Openness to Experience was consistently and negatively related to linguistic markers of immediacy and that Extraversion was negatively correlated with markers of making distinctions. These relationships suggest that these self-report domains may well be tapping the very ways that people are cognitively organizing their worlds, as reflected through their word choices.

Finally, and perhaps most important, the LIWC dimensions were related to overt behaviors in the classroom as well as a variety of demographic and illness-related behaviors. Indeed, regression analyses indicate that these LIWC dimensions predict such things as illness as well as or better than the five-factor dimensions. Again, the magnitude of the relationships was modest but significant. However, it is important to appreciate that we predicted these overt and socially meaningful behaviors on the basis of class writing exercises that had nothing to do with health, illness, or other criterion-related measures.

### GENERAL DISCUSSION

Taken together, the data demonstrate that the ways people express themselves in words are remarkably reliable across time and situations. The dimensions of language that are reliable are impressively diverse as well. From the use of highly specific articles to general emotional language, both positive and negative emotion words, use of large words, and even verb tense are stable across individuals. Within each study, the writing topics that people addressed changed as well, although perhaps not dramatically. Within the various reliability studies, students wrote on different assigned topics; in the SESP study, we assumed that researchers publish different ideas and data over their career; in large introductory psychology classes, students wrote on two general topics but, presumably, in different ways from assignment to assignment. Despite the different ranges of writing topics, word use remained reliable.

The factor analyses in the second phase of the research further demonstrated the utility of thinking of language styles as individual differences. Interestingly, two of the four factors—Immediacy and, to a less extent, The Social Past—are similar to another factor analytic approach to language adopted by Biber (1988), an English

professor, who has been more concerned with parts of speech than psychological dimensions per se. It is not coincidental that the Immediacy dimension is composed of fairly concrete parts of speech. One problem in understanding the meaning of the four LIWC factors in the present research is that we do not know how they hold up in very different writing forms. In short, the factor structure of language may well be dependent on writing topic, setting, or implicit or explicit writing rules. Furthermore, we are not yet able to generalize the current factors to natural speech.

On a broader level, the findings support the idea that reliable linguistic styles can be identified. These styles, as suggested by Allport, help provide an understanding of how individuals convey their thoughts and feelings. Participants, for example, who use a high percentage of articles in their speech by definition are referring to concrete and impersonal objects or events. Use of articles may not reveal which objects or events a person is talking about, but it provides a sense of the person's general approach to the world.

As suggested in the third phase of the research, the linguistic style strategy overlaps in several interesting ways with projective tests. The participants may or may not be talking about themselves. Even if the writing focuses on other topics, their characteristic speech patterns may emerge. Use of negative emotions in language may serve as a good example. People who are consistently harsh in their comments on friends, the weather, and their use of time will use a high rate of negative emotion words. Even if they claim that they are not angry or upset, their use of negative emotion words may (or may not) betray an underlying sense of distress. Indeed, it is interesting that use of negative emotion words was significantly correlated with both alcohol and tobacco use, whereas self-reports of neuroticism were uncorrelated with these behaviors. Placing the present data in the context of a "person-situation" debate, we find strong evidence for the influence of the person on how the person writes, yet not much evidence for the influence of the traits of the person on the person's characteristic style of verbal expression.

An important question that is addressed in the literature on consistency is the degree to which behavior is discriminative versus consistent across situations (e.g., Greene, 1989; Shoda & Mischel, 1994). The results of the present research only begin to speak to the issue of discriminativeness. However, these results do show enormous consistency, on par with the reliabilities of traditional questionnaires. The correlations of language use across topics and over time are remarkably high, given the constraints that a particular topic may place on a person. Thus, language use may be thought of as an arena in which the impact of the person is unavoidable. An important consideration for future work is to identify whether a person's characteristic style of expression reveals personality dispositions or merely verbal ability. Importantly, the relations between the LIWC dictionaries and SAT verbal scores were modest. It does not appear that the individual consistency apparent in these data is due to intelligence or verbal acuity alone. Still, the psychological meaning of language use remains somewhat ambiguous.

A project such as this raises the traditional question in personality of what is the "gold standard" of measurement. The Big Five approach has certainly emerged as the leading contender in defining the central domains of self-assessed personality. That language dimensions are only weakly correlated with the Big Five could be

greeted with skepticism—if one is certain that the Big Five approach is the true one—or optimism. Our perspective is that the language strategy offers another comprehensive perspective in tapping individual difference styles.

If future investigators are to examine the linguistic style approach, programs such as LIWC must be examined quite closely. The LIWC computer program was designed to study the language people use when writing about traumatic experiences. Although it underwent extensive testing in the creation of the LIWC categories, the categories themselves were somewhat arbitrarily created by the authors and their colleagues. Future investigations must expand the net of language categories to provide a broader sense of how linguistic styles are interrelated.

The present project represents a first step in determining the degree to which the linguistic style approach is a valid strategy by which to study personality. The ability to isolate reliable language patterns from person to person allows some insight into the ways people perceive the world and naturally express themselves. Such an approach also allows for a very different approach to personality assessment, one that is objective (in the sense of counting words), has social relevance, and provides a window into how individuals are organizing their psychological worlds.

## References

- Allport, G. W. (1961). *Pattern and growth in personality*. New York: Holt, Rinehart & Winston.
- Atkinson, J. W. (1982). Motivational determinants of thematic apperception. In A. Stewart (Ed.), *Motivation and society* (pp. 3–40). San Francisco: Jossey-Bass.
- Biber, D. (1988). *Variations across speech and writing*. New York: Cambridge University Press.
- Broehl, W. G., & McGee, V. W. (1981). Content analysis in psychohistory: A study of three lieutenants in the Indian Mutiny, 1857–58. *Journal of Psychohistory*, 8, 281–306.
- Bucci, W. (1997). *Psychoanalysis and cognitive science: A multiple code theory*. New York: Guilford Press.
- Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, 42, 116–131.
- Forgas, J. P. (Ed.). (1985). *Language and social situations*. New York: Springer.
- Foster, D. (1996, February 26). Primary culprit: Who is anonymous? *New York Magazine*, 29, 50–57.
- Furnham, A. (1990). Language and personality. In H. Giles & W. P. Robinson (Eds.), *Handbook of language and social psychology* (pp. 73–95). New York: Wiley.
- Giles, H., & Wiemann, J. M. (1993). Social psychological studies of language: Current trends and prospects. *American Behavioral Scientist*, 36, 262–272.
- Gottschalk, L. A., & Gleser, G. C. (1969). *The measurement of psychological states through the content analysis of verbal behavior*. Berkeley: University of California Press.
- Gottschalk, L. A., Stein, M. K., & Shapiro, D. H. (1997). The application of computerized content analysis of speech to the diagnostic process in a psychiatric outpatient clinic. *Journal of Clinical Psychology*, 53, 427–441.
- Greene, J. O. (1989). The stability of nonverbal behavior: An action production approach to problems in cross situational consistency and discriminativeness. *Journal of Language and Social Psychology*, 8, 193–220.
- Harmon, H. H. (1976). *Modern factor analysis* (3rd ed.). Chicago: University of Chicago Press.

- Hart, R. P. (1984). *Verbal style and the presidency: A computer-based analysis*. New York: Academic Press.
- Higgins, E. T., Vookles, J., & Tykocinski, O. (1992). Self and health: How "patterns" of self-beliefs predict types of emotional and physical problems. *Social Cognition*, 10, 125-150.
- Horowitz, M. J., Milbrath, C., & Stinson, C. H. (1995). Signs of defensive control locate conflicted topics in discourse. *Archives of General Psychiatry*, 52, 1040-1047.
- Jackson, D. N. (1974). *Manual for the Personality Research Form*. Goshen, NY: Research Psychology Press.
- John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The Big Five Inventory—Versions 4a and 4b* (Technical Report). Berkeley: Institute for Personality and Social Research, University of California.
- King, L. A. (1995). Wishes, motives, goals, and personal memories: Relations and correlates of measures of human motivation. *Journal of Personality*, 63, 985-1007.
- Krauss, R. M., & Fussell, S. R. (1996). Social psychological models of interpersonal communication. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 655-701). New York: Guilford Press.
- Lundy, A. (1988). Instructional set and Thematic Apperception Test validity. *Journal of Personality Assessment*, 52, 141-145.
- McClelland, D. C. (1985). *Human motivation*. Glenview, IL: Scott, Foresman.
- Mehrabian, A. (1967). Attitudes inferred from non-immediacy of verbal communications. *Journal of Verbal Learning and Verbal Behavior*, 6, 294-295.
- Mergenthaler, E. (1996). Emotion-abstraction patterns in verbatim protocols: A new way of describing psychotherapeutic processes. *Journal of Consulting and Clinical Psychology*, 64, 1306-1315.
- Mischel, W. (1968). *Personality and assessment*. New York: Wiley.
- Morgan, C., & Murray, H. A. (1935). A method for investigating fantasies: The Thematic Apperception Test. *Archives of Neurology and Psychiatry*, 34, 289-306.
- Oxman, T. E., Rosenberg, S. D., Schnurr, P. P., & Tucker, G. J. (1988). Diagnostic classification through content analysis of patients' speech. *American Journal of Psychiatry*, 145, 464-468.
- Pennebaker, J. W. (1997). Writing about emotional experiences as a therapeutic process. *Psychological Science*, 8, 162-166.
- Pennebaker, J. W., Colder, M., & Sharp, L. K. (1990). Accelerating the coping process. *Journal of Personality and Social Psychology*, 58, 528-537.
- Pennebaker, J. W., & Francis, M. E. (1996). Cognitive, emotional, and language processes in disclosure. *Cognition and Emotion*, 10, 601-626.
- Pennebaker, J. W., & Francis, M. E. (1999). *Linguistic Inquiry and Word Count: LIWC*. Mahwah, NJ: Erlbaum.
- Pennebaker, J. W., Mayne, T. J., & Francis, M. E. (1997). Linguistic predictors of adaptive bereavement. *Journal of Personality and Social Psychology*, 72, 863-871.
- Peterson, C., Seligman, M. E. P., & Vaillant, G. E. (1988). Pessimistic explanatory style is a risk factor for physical illness: A thirty-five-year longitudinal study. *Journal of Personality and Social Psychology*, 55, 23-27.
- Peterson, C., & Ulrey, L. M. (1994). Can explanatory style be scored from TAT protocols? *Personality and Social Psychology Bulletin*, 20, 102-106.
- Rosenberg, M. (1979). *Conceiving the self*. New York: Basic Books.
- Rosenberg, S. D., Schnurr, P. P., & Oxman, T. E. (1990). Content analysis: A comparison of manual and computerized systems. *Journal of Personality Assessment*, 54, 298-310.
- Rosenberg, S. D., & Tucker, G. J. (1979). Verbal behavior and schizophrenia: The semantic dimension. *Archives of General Psychiatry*, 36, 1331-1337.
- Schnurr, P. P., Rosenberg, S. D., & Oxman, T. E. (1992). Comparison of TAT and free speech techniques in eliciting source material in computerized content analysis. *Journal of Personality Assessment*, 58, 311-325.
- Schnurr, P. P., Rosenberg, S. D., Oxman, T. E., & Tucker, G. J. (1986). A methodological note on content analysis: Estimates of reliability. *Journal of Personality Assessment*, 50, 601-609.
- Shoda, Y., & Mischel, W. (1994). Intraindividual stability in the organization and patterning of behavior: Incorporating psychological situations into the idiographic analysis of behavior. *Journal of Personality and Social Psychology*, 67, 674-687.
- Smith, C. P. (1992). *Thematic content analysis for motivation and personality research*. New York: Cambridge University Press.
- Spence, D. P., Scarborough, H. S., & Ginsberg, E. H. (1978). Lexical correlates of cervical cancer. *Social Science and Medicine*, 12, 141-145.
- Stein, N., Folkman, S., Trabasso, T., & Richards, T. A. (1997). Appraisal and goal processes as predictors of psychological well-being in bereaved caregivers. *Journal of Personality and Social Psychology*, 72, 872-884.
- Stephenson, G. M., Laszlo, J., Ehmann, B., Lefever, R. M., & Lefever, R. (1997). Diaries of significant events: Socio-linguistic correlates of therapeutic outcomes in patients with addiction problems. *Journal of Community and Applied Social Psychology*, 7, 389-411.
- Stiles, W. B. (1992). *Describing talk: A taxonomy of verbal response modes*. Newbury Park, CA: Sage.
- Stone, P. J., Dunphy, D. C., Smith, M. S., & Ogilvy, D. M. (1966). *The General Inquirer: A computer approach to content analysis*. Cambridge, MA: MIT Press.
- Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics* (3rd ed.). Reading, MA: Addison-Wesley.
- Taylor, M. A., Reed, R., & Berenbaum, S. (1994). Patterns of speech disorders in schizophrenia and mania. *Journal of Nervous and Mental Disease*, 182, 319-326.
- Tetlock, P. E., Peterson, R. S., & Berry, J. M. (1993). Flattering and unflattering personality portraits of integratively simple and complex managers. *Journal of Personality and Social Psychology*, 64, 500-511.
- Thomas, C. B., & Duszynski, K. R. (1985). Are words of the Rorschach predictors of disease and death? The case of "whirling." *Psychosomatic Medicine*, 47, 201-211.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54, 1063-1070.
- Weintraub, W. (1989). *Verbal behavior in everyday life*. New York: Springer.
- Winter, D. G. (1989). *Manual for scoring motive imagery in running text*. Unpublished manuscript, University of Michigan, Ann Arbor.
- Zeldow, P. B., & McAdams, D. P. (1993). On the comparison of TAT and free speech techniques in personality assessment. *Journal of Personality Assessment*, 60, 181-185.



## Appendix

## Summary of Reliability Statistics

Dimension	Diary		Taos		SESP		Mean <i>r</i>	Mean $\alpha$
	<i>r</i>	$\alpha$	<i>r</i>	$\alpha$	<i>r</i>	$\alpha$		
Language								
Word count per essay	.57	.96	.42	.87	.15	.72	.38	.85
Words per sentence	.39	.91	.52	.92	.06	.51	.32	.78
Question marks (%)	.12	.67	.19	.71	—	— <sup>a</sup>	.16	.69
Unique (type-token)	.57	.96	.54	.92	.17	.75	.43	.88
Dictionary words	.24	.85	.35	.81	.15	.72	.25	.79
Words of more than 6 letters	.44	.92	.49	.89	.23	.82	.39	.88
Total pronouns	.30	.88	.31	.79	.12	.66	.24	.78
First person singular	.25	.86	.13	.61	.03	.13 <sup>a</sup>	.14 (.19)	.53 (.74)
First person plural	.29	.86	.07	.45	.03	.18 <sup>a</sup>	.13 (.18)	.50 (.66)
Total first person	.20	.83	.11	.56	.02	.13 <sup>a</sup>	.11 (.16)	.51 (.70)
Total second person	.03	.28	.16	.62	—	— <sup>a</sup>	.10	.45
Total third person	.16	.72	.26	.72	.10	.63	.17	.69
Negations	.18	.79	.30	.81	.03	.30	.17	.63
Assents	.10	.53	.11	.55 <sup>a</sup>	.06	.55 <sup>a</sup>	.09 (.10)	.54 (.53)
Articles	.23	.84	.21	.63	.20	.78	.21	.75
Prepositions	.13	.71	.24	.76	.06	.48	.14	.65
Numbers	.07	.60	.14	.60	.03	.32	.08	.51
Swear words	.13	.52	.32	.81	—	— <sup>a</sup>	.23	.67
Psychological Processes								
Affective processes	.15	.69	.16	.64	.19	.78	.17	.70
Positive emotions	.23	.84	.11	.54	.17	.75	.17	.71
Positive feelings	.13	.74	.11	.52	.19	.79	.14	.68
Optimism	.10	.62	.10	.53	.04	.47	.08	.54
Negative emotions	.24	.84	.19	.68	.25	.84	.23	.79
Anxiety	.17	.77	.06	.31	.10	.56	.11	.55
Anger	.05	.35	.17	.67	.32	.88	.18	.63
Sadness	.17	.78	.03	.24	.31	.87	.17	.63
Cognitive processes	.23	.84	.25	.76	.13	.69	.20	.76
Causation	.08	.56	.20	.72	.17	.76	.15	.68
Insight	.15	.77	.18	.69	.12	.65	.15	.70
Discrepancy	.10	.67	.24	.74	.09	.62	.14	.68
Inhibition	.13	.74	.10	.53	.03	.37	.09	.55
Tentative	.06	.50	.28	.80	.08	.56	.14	.62
Certainty	.06	.61	.08	.43	.08	.55	.07	.53
Sensory processes	.21	.83	.09	.50	.04	.39	.11	.57
Seeing	.02	-.20	.04	.21	.04	.43	.03	.15
Hearing	.12	.66	.18	.69	.03	.32	.11	.56
Feeling	.29	.88	.12	.61	.02	.16	.14	.55
Social processes	.13	.74	.30	.78	.16	.74	.20	.75
Communication	.01	.10	.13	.59	.12	.67	.09	.45
Other references	.22	.83	.25	.75	.10	.62	.19	.73
Friends	.02	.38	.09	.49	.06	.32	.06	.40
Family	.04	.29	.03	.48	.21	.77	.09	.51
Humans	.07	.63	.09	.48	.17	.77	.11	.63
Relativity								
Time	.16	.81	.10	.50	.05	.45	.10	.59
Past tense	.37	.91	.17	.67	.15	.74	.23	.77
Present tense	.11	.69	.17	.68	.19	.78	.16	.72
Future tense	.11	.68	.11	.48	.02	.20	.08	.45
Space	.06	.46	.13	.59	.07	.54	.09	.53
Up	.04	.06	.07	.45	.04	.33	.05	.28
Down	.01	.08	.03	.27	.11	.56	.05	.30
Inclusive	.16	.78	.29	.80	.17	.75	.21	.78
Exclusive	.17	.77	.41	.87	.12	.66	.23	.77
Motion	.03	.38	.08	.50	.06	.50	.06	.46

(Appendix continues)

## Appendix (continued)

Dimension	Diary		Taos		SESP		Overall <i>r</i>	Overall $\alpha$
	<i>r</i>	$\alpha$	<i>r</i>	$\alpha$	<i>r</i>	$\alpha$		
Current concerns								
Occupation	.08	.70	.22	.75	.14	.71	.15	.72
School	.13	.74	.23	.76	.11	.63	.16	.71
Job	.20	.85	.03	.22	.22	.80	.15	.62
Achievement	.05	.47	.11	.52	.08	.58	.08	.52
Leisure	.07	.57	.01	.02	.03	.42	.04	.34
Home	.11	.65	.05	.35	.04	.09 <sup>a</sup>	.07 (.08)	.36 (.50)
Sports	.06	.55	.11	.49	.06	.43 <sup>a</sup>	.08 (.09)	.49 (.52)
Television	.05	.55 <sup>a</sup>	.01	.17	.07	.52	.04 (.04)	.41 (.34)
Music	.04	.44	.03	.21	.10	.60	.06	.42
Money	.02	.20 <sup>a</sup>	.03	.04	.10	.65 <sup>a</sup>	.05 (.05)	.30 (.31)
Metaphysical	.11	.67	.20	.66	.06	.25 <sup>a</sup>	.12 (.16)	.53 (.67)
Religion	.11	.66 <sup>a</sup>	.27	.71	.10	.38 <sup>a</sup>	.16 (.27)	.58 (.71)
Death	.12	.65 <sup>a</sup>	.07	.32 <sup>a</sup>	.00	-.03 <sup>a</sup>	.06	.31
Physical states	.17	.71	.07	.41	.21	.79	.15	.64
Body	.07	.47	.15	.56	.17	.73	.13	.59
Sexuality	.10	.63 <sup>a</sup>	.15	.58	.24	.86	.16 (.17)	.69 (.72)
Eating	.15	.69	.01	-.10	.38	.85 <sup>a</sup>	.18 (.08)	.48 (.30)
Sleeping	.11	.63	.02	.30	—	— <sup>a</sup>	.07	.47
Grooming	.14	.51	.05	.37	.08	.52 <sup>a</sup>	.09 (.09)	.47 (.44)
Samples/participant		18		10		15		
<i>N</i>		15		34		40		

*Note.* Correlations are mean values;  $\alpha$  refers to Cronbach alphas. SESP = Society for Experimental Social Psychology. Numbers in parentheses refer to adjusted correlations and alpha coefficients excluding low-usage categories.

<sup>a</sup> The overall mean of this category was less than or equal to 0.1% of words used. If the cell contains a dash, more than half of the time periods had values of zero.

Received August 31, 1998

Revision received April 25, 1999

Accepted May 24, 1999 ■

### Low Publication Prices for APA Members and Affiliates

**Keeping you up-to-date.** All APA Fellows, Members, Associates, and Student Affiliates receive—as part of their annual dues—subscriptions to the *American Psychologist* and *APA Monitor*. High School Teacher and International Affiliates receive subscriptions to the *APA Monitor*, and they may subscribe to the *American Psychologist* at a significantly reduced rate. In addition, all Members and Student Affiliates are eligible for savings of up to 60% (plus a journal credit) on all other APA journals, as well as significant discounts on subscriptions from cooperating societies and publishers (e.g., the American Association for Counseling and Development, Academic Press, and Human Sciences Press).

**Essential resources.** APA members and affiliates receive special rates for purchases of APA books, including the *Publication Manual of the American Psychological Association*, and on dozens of new topical books each year.

**Other benefits of membership.** Membership in APA also provides eligibility for competitive insurance plans, continuing education programs, reduced APA convention fees, and specialty divisions.

**More information.** Write to American Psychological Association, Membership Services, 750 First Street, NE, Washington, DC 20002-4242.