

Mapping Cognitive Structures and Processes Through Verbal Content: The Thought-Listing Technique

John T. Cacioppo, William von Hippel, and John M. Ernst
Ohio State University

Open-ended cognitive assessment techniques have helped illuminate the cognitive structures and processes underlying various clinical problems. The authors review a specific open-ended protocol analysis for assessing cognitive structures and processes—the thought-listing technique. They begin with a brief description of this technique and its validity, limitations, and potential clinical use. They then review representative research using the thought-listing technique in studies of psychopathology and psychotherapy. They conclude with a discussion of new and potentially useful methods (e.g., adjusted ratio of clustering scores, multidimensional scaling, implicit memory measures) for mapping cognitive representations and coping processes on the basis of data from thought listings in clinical and counseling psychology.

The flux of thoughts, ideas, feelings, and images in waking states is so continuous and mercurial that it has been likened to the flow of a fast-running mountain stream. Like a mountain stream, the stream of an individual's thoughts and feelings can differ in subtle yet dramatic ways across two different environments, while the same environment can elicit quite different streams of thoughts and images from two different individuals (e.g., Cacioppo & Petty, 1981; Ericsson & Simon, 1980; Nisbett & Wilson, 1977). The *thought-listing technique*, in which an individual is asked to list everything about which he or she is (or was) thinking, is one form of protocol analysis that can be used to assess the contents and underlying landscape of this stream.

The stream of thoughts listed by a solitary individual at a single point in time yields ambiguous information about an individual's cognitive processes or structures. In the thought-listing technique, therefore, information about cognitive processes and structures is inferred from comparisons between data obtained from measurement occasions or groups (e.g., thought listings from the same individual in different circumstances or from individuals from known groups). Consider, for instance, a situation in which individuals are presented with behaviors that are somewhat ambiguous with regard to whether they implicate depression, such as, "Did not leave the house the entire weekend." The individuals might then be asked to prepare to explain the reasons for this behavior and, afterwards, to list everything about which they had been thinking. If the individuals are chron-

ically depressed, the concept of *depression* should be accessible in memory and should be more likely to be a recurrent element in the stream of feelings and ideas elicited in this assessment context than other possible accounts (e.g., *reclusive*, *busy*, or *tired*). If the individuals are hostile and the concepts of *hostility*, *interpersonal threats*, and *conflict* are instead relatively accessible in memory, then the individuals may be more likely to list thoughts indicating a sense of being unfriendly, hostile, or excluded in response to the scenario (cf. Higgins, King, & Mavin, 1982; Smith & Branscombe, 1988). If a normal (nondepressed and nonhostile) group of individuals were also tested, an investigator could discern what was the base rate for each of these types of responses. Thus, comparison groups and the eliciting conditions for the verbal protocols are crucial to the interpretation of the responses that are obtained.

When these individuals are asked why they listed a particular thought or explanation—that is, to identify the reason for their mental contents—research has demonstrated that their explanations can be quite inaccurate (cf. Nisbett & Wilson, 1977; Wilson & Brekke, 1994). Therefore, no assumptions are made in the thought-listing technique about the individual's ability to report the motives for or causes of his or her mental contents. Nor are any assumptions made about an investigator's ability to designate a priori the relevant mental contents (e.g., beliefs) that distinguish between adaptive and maladaptive cognitive processes. As in the example above, however, the thoughts listed and their order may say something about (a) the way in which a person views the world; (b) his or her coping processes in response to different challenges, threats, and circumstances; and (c) the motives and cognitive structures from which these coping processes originate.

Finally, the thought-listing procedure does not require that individuals recall thoughts or events from distant periods of the day or week but rather only that they list the thoughts that are within their stream of consciousness and those that are readily accessible at the moment of assessment. In this respect, the thought-listing procedure is more similar to experience sampling methodologies (Csikszentmihalyi & Larson, 1987) and ecologi-

John T. Cacioppo, William von Hippel, and John M. Ernst, Department of Psychology, Ohio State University.

This research was supported in part by National Science Foundation Grant DBS-9211483 and National Institute of Mental Health Grant T32-MH19728. We thank John Skowronski for his comments on an earlier draft of this article.

Correspondence concerning this article should be addressed to John T. Cacioppo, Department of Psychology, Ohio State University, 1885 Neil Avenue, Columbus, Ohio 43210-1222. Electronic mail may be sent via Internet to cacioppo.1@osu.edu.

cal momentary assessments (Stone & Shiffman, 1994) than to assessments based on written daily or weekly diaries.

Our aim in this article is to review developments and research on the thought-listing technique, an open-response method for acquiring and categorizing mental contents (Cacioppo & Petty, 1981). By *mental contents*, we mean the reportable consequences of a person's cognitive processes; these include an individual's thoughts, feelings, ideas, expectations, appraisals, and images. We begin with a brief description of the thought-listing technique as well as strengths and weaknesses of this technique. We next review representative studies using the thought-listing technique to examine questions about psychopathology and psychotherapy. We conclude with a discussion of additional methods for quantifying data obtained using thought listings that have been used in studies of social cognition to understand specific aspects of (and individual differences in) people's cognitive representations and coping processes.

The Thought-Listing Technique in Context

Procedures for assessing maladaptive cognitions and cognitive processes fall into at least three broad categories. Behavioral (performance) measures rely on self or other (e.g., trained judges, experimenters) ratings of overt and nonverbal behavior from which deductions may be drawn regarding the observed individual's cognitive processes. Behavioral measures can be distinguished from cognitive assessments because the former rely on overt behavior or self-reports of overt behavior rather than on self-reported thoughts or feelings.

A second major type of assessment is the structured questionnaire. This relies on an individual's endorsement of thoughts, feelings, or both, presented in the form of either a survey or a structured interview. These structured assessments can take many forms. Individuals, for instance, may check off (from a predetermined list) those thoughts or beliefs they had during some explicit circumstance or period of time. In addition to noting whether a thought occurred, individuals may also be requested to estimate the frequency (or intensity) with which the thought occurred. Structured questionnaires can greatly simplify the quantification of predetermined dimensions of mental contents; but when using such assessments, one risks overlooking unspecified yet important recurrent themes present in a person's thoughts and ideation.

The thought-listing technique (or, more generally, verbal protocol analyses) is a type of cognitive assessment that is particularly useful when one either has no predetermined ideas about the cognitive dimensions that are relevant or has only a few untested hunches. The thought-listing technique is based on the assumption that the psychological significance of an individual's thoughts and feelings, as well as the underlying cognitive processes, can be examined by content analyzing the individual's reported thoughts, ideas, images, and feelings. Instead of observing a person's behavior directly, asking them to respond to questionnaires, or interviewing them, one obtains protocols from different groups of individuals (or the same individuals in different situations) for purposes of comparison. Thought listings, for instance, may be obtained in anticipation of some situation or task (e.g., listing one's thoughts following a period in which an individual has anticipated a discussion of a topic with an

argumentative partner), as they occur (e.g., taking notes about one's perceptions and thoughts during an argument with a partner), or retrospectively (e.g., listing what one had been thinking immediately following the argument with a partner). In contrast, thought listings typically have not been obtained during conditions of high cognitive load (e.g., during an effortful task or interaction) in an attempt to minimize measurement reactivity (see Cacioppo & Petty, 1981). Research in which thought listings were obtained immediately following high cognitive load conditions, however, indicates that valid data can be obtained about the cognitive processes that were occurring during the high load conditions (e.g., Petty, Wells, & Brock, 1976).

Typically, the participants or judges rate the nature of the listed thoughts (e.g., their valence, target), although listed thoughts can be scored along any dimension(s). A detailed discussion of the procedures for administering the thought-listing technique and quantifying the resulting data can be found elsewhere (e.g., Cacioppo & Petty, 1981; Tarico, Van Velzen, & Altmeyer, 1986; cf. Glass & Arnkoff, 1994).

The thought-listing technique was first used to explore possibly maladaptive cognitive processes in a study of undergraduate male students who were either high or low in social anxiety (Cacioppo, Glass, & Merluzzi, 1979). The students were instructed that they were to engage in a discussion with an unfamiliar woman and were left alone in a room for several minutes to collect their thoughts about the impending interaction. Participants were then asked to list everything about which they had been thinking:

We are now interested in everything that went through your mind about the upcoming discussion. Please list these thoughts, whether they were about yourself, the situation, and/or others; whether they were positive, neutral, and/or negative. Any case is fine. Ignore spelling, grammar, and punctuation. You will have 2.5 minutes to write. We have deliberately provided more space than we think people will need, to ensure that everyone would have plenty of room. Please be completely honest. Your responses will be anonymous. The next page contains the form we have prepared for your use to record your thoughts and ideas. Simply write down the first thought you had in the first box, the second in the second box, etc. Please put only one idea or thought in a box. (Cacioppo, Glass, & Merluzzi, 1979, p. 254)

After 2.5 min, the participants were instructed to review their listed thoughts and rate each as favorable toward themselves, unfavorable toward themselves, or neutral (or personally irrelevant). Subsequently, participants also rated themselves and the impending discussion using semantic differential scales (e.g., good-bad, active-passive). Two independent judges, who were unaware of the experimental conditions, also scored the protocols according to the valence of each listed thought to the "self," with high interrater reliability ($r = +.95$) observed between the judges.

Analyses of the number of negative, neutral, and positive self-statements revealed two results worth noting here. First, men high in social anxiety were found to generate significantly more negative self-statements (as rated by judges) and to express more negative self-regard prior to engaging in a discussion with an unfamiliar woman than did men low in social anxiety. It is not that socially anxious men thought less, or less positively,

about themselves when anticipating the interaction but that they thought more negatively about themselves. Men high in social anxiety also rated themselves more negatively and as lower in potency and activity than did men low in social anxiety, and they rated the impending discussion with the unfamiliar woman more negatively on the semantic differential than did men low in social anxiety. These results are consistent with more recent research in psychopathology suggesting that negative self-statements have a stronger and a broader influence on people's symptoms and affective states than positive self-statements and that psychological adjustment may be more closely related to the absence (or diminution) of negative thinking than the presence (or augmentation) of positive thinking (e.g., Bruch, Gorsky, Collins, & Berger, 1989; Glass & Furlong, 1990; Melchior & Cheek, 1990; see also Kendall, 1984).

Analyses of the participants' ratings of their self-statements indicated that individuals high and low in social anxiety rated their own listed thoughts similarly (Cacioppo et al., 1979). This latter finding was interpreted to mean that each group had a unique frame of reference for what constitutes a normal or favorable self-statement. Consistent with this interpretation, research has shown that high socially anxious individuals possess more negative expectations regarding social interactions, a more concentrated attentional focus on the positive attributes of the other person, and more negative generalizations about themselves than do low socially anxious individuals (e.g., Clark & Arkowitz, 1975; Mahone, Bruch, & Heimberg, 1993). Thus, although not anticipated, the thought-listing technique revealed the importance of considering the frames of reference used by participants versus those of the judges who were charged with categorizing the listed thoughts.

When individuals are not randomly assigned to groups, as when comparing clinical and control groups, only the ratings of judges who are unaware of the experimental conditions can be assumed to reflect comparable frames of reference across conditions. When participants are randomly assigned to conditions, their constitutional frames of reference can be assumed to be comparable across conditions, but their frames of reference, as well as their thoughts and ideas, may be affected by the experimental conditions (e.g., see Cacioppo, Andersen, Turnquist, & Tassinari, 1989, for a discussion of the effects of perspective differences on judgments and ratings). Of course, the discrepancies between participant and judge ratings can become interesting data because they may provide a view on the contrasting frames of reference that characterize individuals in the different groups.

It follows from the above that participant ratings of listed thoughts may also reveal group differences that are not apparent in judge ratings or in simple counts of positive or negative self-statements. Individuals with phobias, anxiety, or depression, for instance, may perceive the same negative event as more consequential than may normals. Data consistent with this possibility were reported in an interesting study by Huber and Altmaier (1983). These researchers obtained thought listings from individuals with a snake phobia and those without such a phobia following behavioral avoidance tasks. Subsequently, the participants rated the salience or intensity of each listed thought (from 1 to 3), and listed a thought they perceived to be the exact opposite of each listed thought. Each bipolar pair of thoughts

was then rated by independent judges on a 5-point scale to index the degree of threat. The participants' ratings of salience were multiplied by the judges' ratings of threat to obtain a total threat rating for each listed thought. Analyses of these indices indicated that those with a phobia and those without one did not differ in the number of positive or negative thoughts listed but that the groups did differ in terms of the average total threat rating of the listed thoughts. Thus, whether participant or judge ratings are preferable depends on the nature of the research question.

Applications and Paradigms

Within clinical and counseling psychology, the thought-listing technique has been used to study topics such as social phobia (Heimberg, Bruch, Hope, & Dombeck, 1990; Heimberg, Dodge, et al., 1990), snake phobia (Huber & Altmaier, 1983), test anxiety (Blankstein & Flett, 1990; Blankstein, Toner, & Flett, 1989), social anxiety in its various forms (Altmaier, Leary, Halpern, & Sellers, 1985; Cacioppo et al., 1979; Glass & Furlong, 1990; Mahone et al., 1993), romantic relationships (Van Lange & Rusbult, 1995), group therapy (Hines, Stockton, & Morran, 1995; Morran, Kurpius, & Brack, 1989), fraudulence (Kolligian & Sternberg, 1991), depressed mood (Elliott, MacNair, Herrick, Yoder, & Byrne, 1991; Ellis, Seibert, & Herbert, 1990; Sedikides, 1992), career barriers (Swanson & Tokar, 1991), and assertiveness training (Bruch, Hamer, & Kaflowitz-Linder, 1992).

The Cacioppo et al. (1979) research, discussed above, is an example of the ways in which thought listing has been used to understand the cognition of individuals diagnosed or dealing with a psychological disorder (see also Ellis et al., 1990; and Mahone et al., 1993). A noteworthy feature of this research is that it points to the need not only to examine the negativity and positivity of the thoughts listed by individuals but also to the targets (self vs. other) of the thoughts. A second example of this importance can be found in the work of Mahone et al. In this study, male participants completed the Cacioppo et al. thought-listing protocol twice. After viewing a picture and reading a personal statement of a future conversation partner (an attractive female confederate), participants were asked to list their thoughts regarding themselves (self-focused) and their partner (other focused). Participants then engaged in a 5-min conversation with their partner. Here, the total number of negative and positive thoughts did not predict self-report anxiety, self-efficacy, or observed anxiety behaviors. However, a greater number of negative self-thoughts predicted both lower self-efficacy and higher self-reported social anxiety. In addition, a greater number of positive other thoughts predicted a greater number of observed anxious behaviors. Hence, the focus of attention was important for understanding the nature of cognitions regarding social anxiety (note that it was not only negative self-thoughts but also positive other thoughts that were important).

Psychological disorders affect not only individuals with the disorders but also the people with whom they interact. A second way in which thought listings have been used in clinical and counseling psychology, therefore, has been to examine the social effects of psychopathologies. Elliott et al. (1991), for instance, used thought listing to probe the effects of dyadic interactions with confederates who acted either depressed or normally on

the cognitions and behaviors of normal individuals. Elliott et al. found that participants who interacted with a confederate acting in a depressed mode spoke less to, gazed less at, and listed more negative thoughts than participants who interacted with a nondepressed confederate. Ellis et al. (1990) provided complementing data in their study of the effects of mood on thought listings. Neutral and depressed moods were induced using the Velton (1968) mood procedure. Ellis et al. found that participants in a depressed mood listed a greater proportion of negative thoughts than did individuals in a neutral mood.¹

Third, the thought-listing technique has served as a dependent variable for testing the efficacy of therapeutic techniques (e.g., Altmeyer et al., 1985; Crocker, 1989; Glass & Arnkoff, 1983; Long, 1984). In a study examining therapeutic interventions for public-speaking anxiety, Altmeyer et al. examined the cognitions of participants regarding an upcoming public speaking task. These individuals were selected from a screening task as high in public-speaking anxiety. Prior to this task, two thirds of the participants had undergone three weekly 1 1/2-hr training sessions in either participant modeling or stress inoculation. The stress inoculation group was taught cognitive and behavioral coping strategies for anxiety. The participant modeling group received examples of positive public speaking and coaching on public speaking skills. A third group represented a no-training condition. Participants were asked to list all thoughts immediately after receiving instructions for an upcoming public speaking task but before they began the task. Coders rated each thought for confidence regarding the upcoming speech and intensity of the thought. Hence, thought listing was used as a criterion for measuring therapy intervention success. The experimenters found that stress inoculation individuals listed more confident thoughts than did participant modeling individuals, who in turn listed more confident thoughts than did control individuals.

Fourth, thought listing has been used for validating assessment tools. Kolligian and Sternberg (1991), for instance, developed the Perceived Fraudulence Scale (PFS) as a paper-and-pencil questionnaire for assessing the extent to which an individual perceives him- or herself to be a fraud or imposter. As a way of testing the validity of the scale, they asked college students to read six short scenarios and to imagine themselves as the main character in each scenario. The participants were then asked "spontaneously to list all thoughts and feelings that occurred to them during and in response to reading each passage" (p. 313). They had 8 min to read each passage and to list at least five thoughts. Two independent coders then rated each thought along seven dimensions, such as fraudulence, self-deprecation, anxiety, and positive emotions. After the thought-listing task, participants completed several questionnaires including the PFS and a semistructured interview. The interview later was rated on a global dimension for perceived fraudulence. The number of fraudulent thoughts was correlated with the PFS ($r = .48$), providing evidence for the validity of the PFS as a measure of perceived fraudulence. The number of self-deprecating thoughts was also correlated with the PFS ($r = .31$). In addition, the number of fraudulent thoughts was correlated with the rating of perceived fraudulence for the interview ($r = .46$). Hence, not only did the thought-listing scores correlate with the scale itself but also with a behavioral rating.

Fifth, thought listing has been used as a method for exploring

the process of therapy (Heppner, Rosenberg, & Hedgespeth, 1992; Hines et al., 1995; Morran et al., 1989). Morran et al., for example, used the thought-listing technique to examine the thoughts of counselors conducting therapy. Participants were counselor trainees and community-based professional therapists. Each participant conducted a 25-min counseling session with a college student discussing a real problem. After the session, the counselors (a) were instructed in the thought-listing technique, using the guidelines given by Cacioppo and Petty (1981); (b) viewed four 2-min segments of a videotape recorded from the session they had just conducted; and, (c) after each 2-min segment, were asked "to try and recall their exact thoughts during the session" (Morran et al., 1989, p. 506). Subsequently, a group of judges sorted the thoughts into 14 mutually exclusive categories of self-task and gave each category a label. A second group of judges then resorted the thoughts into the categories defined by the first group. Each thought had to be placed into the same category by at least 75% of the judges to be counted as reliably categorized. The raw frequency of thoughts per category was then transformed into proportion of thoughts in the category to total number of thoughts. These proportions were then submitted to an alternating least squares scaling analysis that yielded two dimensions. These two dimensions, (a) attending and assessing versus information seeking and (b) integrative understanding versus intervention planning, accounted for 72% of the variance. These data were interpreted as reflecting the underlying thought processes that counselors used while providing therapy.

Sixth, thought listing has been proposed as a tool in the development of therapeutic strategies. Halford and Sanders (1988) had 20 distressed and 20 nondistressed couples complete a 10-min problem-solving task. The participants' cognitions were then assessed using both thought listing (following directions given by Cacioppo & Petty, 1981) and video-assisted recall. Coders then rated each thought as negative self, negative other, positive self, positive other, or none of the above. Percentage scores for each type of thought were then calculated from total number of category thoughts divided by total number of thoughts. Results indicated that distressed couples reported more negative other thoughts, fewer positive other thoughts, and fewer positive self-thoughts than did nondistressed couples. More interesting, thought listing provided a pattern of results similar to that obtained using the video-assisted recall without the associated cost and time involvement. The investigators suggested that future research examine how the thoughts generated in such a task can be used to guide therapeutic interventions, such as stress inoculation training for couples.

Finally, once one has identified what dimensions are relevant and what cognitions differentiate two or more groups of interest

¹ Ellis et al. (1990) conducted a second study in which elated, neutral, and depressed moods were induced using the Velton mood procedure. Unlike their first study, however, participants engaged in an incidental memory task, after which they were instructed to list the thoughts that occurred to them while working on the task. Although the proportion of negative thoughts tended to vary with mood, Ellis et al. reported that 97% of the thoughts listed pertained to the participants' reactions to the cognitive task as distinct from their mood state. These data attest to the sensitivity of the thought-listing technique to assessment conditions.

(e.g., anxious vs. nonanxious, depressed vs. nondepressed), the thought-listing technique has been used as a basis for developing structured self-report scales. The first to use the thought-listing technique in this manner were Glass, Merluzzi, Biever, and Larsen (1982), who developed the Social Interaction Self-Statement Test (SISST). The SISST is a 30-item scale for the assessment of the cognitive correlates of heterosocial anxiety, with scale items worded in the form of self-statements. The items came from thought listings generated by participants who were responding to 10 heterosocial vignettes. Specifically, the listed thoughts were rated by independent judges on an 11-point scale from *negative* to *positive*, and the 15 most positively rated and the 15 most negatively rated self-statements were selected for inclusion in the scale. Glass et al. found that responses to these items discriminated high from low heterosocially anxious men and women and correlated with self-report measures of social anxiety and heterosocial skill. They also found that negative (but not positive) self-statement scores were correlated with judge and confederate ratings of skill and anxiety, providing additional evidence of the greater impact of negative rather than positive self-statements in maladaptive coping processes.

Psychometric Properties

Psychometric data on the thought-listing technique in clinical and counseling research were reviewed more recently by Glass and Arnkoff (1994) and Heimberg (1994; see also Arnkoff & Glass, 1989; and Martzke, Andersen, & Cacioppo, 1987). Briefly, evidence for the criterion-related validity of thought-listing data has been provided in research on social anxiety and social phobia. The number of negative thoughts, for instance, has been found to covary with lower self-evaluations (Cacioppo et al., 1979) and level of subjective discomfort (Amsel & Fichten, 1990; Myszka, Galassi, & Ware, 1986); although Glass and Arnkoff noted that when frequency counts are plagued by problems of restricted range, these associations are considerably weaker. Concurrent validity has also been obtained in contrasts of various groups known to differ in their level of shyness, anxiety, or phobia (see the review by Glass & Arnkoff, 1994).

Fichten, Amsel, and Robillard (1988) found that participants' ratings of how comfortable they would feel in hypothetical interactions were similar whether or not participants completed a thought-listing measure, suggesting that the assessment is not reactive (cf. Cacioppo & Petty, 1981; and Glass & Arnkoff, 1994). Blackwell, Galassi, Galassi, and Watson (1985), however, presented evidence from highly math-anxious students who performed a math task, suggesting that the thought-listing procedure may result in fewer listed thoughts and higher levels of anxiety than think-aloud procedures. Whether this is a problem depends on the question of interest. The evidence for convergent and discriminant validity of measures obtained from thought listings is spottier, and we are left to echo Glass and Arnkoff's call for the need for investigators to attend to the convergent and discriminant validity of these measures.

Predictable changes in thought-listing data have also been found from pre- to posttreatment across a range of interventions for social anxiety and social phobia (cf. Heimberg, 1994). For instance, Heimberg, Dodge, et al. (1990) used the thought-listing technique to evaluate cognitive-behavioral group therapy

and educational supportive group psychotherapy for social phobia. Results of the posttest indicated that individuals with a social phobia in both groups showed significant increases in the percentage of positive thoughts and significant reductions in the percentage of negative thoughts. A 6-month follow-up assessment further revealed that the increases in positive thoughts were maintained in both groups but that the reductions in negative thoughts were maintained only in the group who had received cognitive-behavioral group therapy. More interesting, correlations between change scores (6-month follow-up - pretreatment) on the thought-listing measures and assessor-rated severity of symptoms were significant for both negative ($r = +.44$) and positive ($r = -.31$) thoughts (Heimberg, Dodge, et al., 1990). Although data of this type are limited, the research on the treatment of social phobia by Heimberg and colleagues suggests that thought-listing measures may prove a valid and useful measure in studies of clinical outcomes.

Limitations

Although the studies reviewed above illustrate some of the advantages of the thought-listing technique, there are also limitations to this technique that are important to recognize. Among the more problematic of these is that individuals may be unwilling to report their thoughts accurately. This may occur when their responses are socially inappropriate and are not anonymous. Second, the thought-listing technique may be inappropriate when individuals are unable to provide relevant verbal protocols. Young children, for instance, may be unable to provide accurate retrospective reports of their internal dialogues, in which case interview or think-aloud procedures may be more appropriate for assessing mental contents (Daleiden, Vasey, & Williams, 1996). On a related note, children and adults may be unable to report on repressed aspects of their mental contents. Clever investigators may be able to minimize or circumvent these problems by using measures or establishing assessment contexts that lift the veil of self-presentations to reveal an individual's internal dialogue and underlying structure (see *Measures Derived From Thought-Listing Data to Infer Cognitive Representations and Processes*). Thus, these problems should be considered when developing experimental designs or interpreting results from retrospective verbal protocols.

Third, people's memory for thoughts, feelings, and events is imperfect, and the reconstructive nature of memory can introduce systematic biases in thought listings. Redelmeier and Kahneman (in press), for instance, contrasted the real-time and retrospective evaluations of two moderately painful medical procedures (colonoscopy and lithotripsy). Patients undergoing these procedures rated their pain every 60 s and provided retrospective evaluations. Correlational analyses showed that the peak pain reported and the level of pain reported at the end of the procedure were the best predictors of retrospective evaluations. Kahneman, Fredrickson, Schreiber, and Redelmeier (1993) reasoned that if people's cognitive contents and judgments following an event (e.g., a painful medical procedure) were colored by their peak and end experiences with the event, then paradoxical improvements in their thoughts and evaluations of a painful medical procedure would be observed by extending the duration of the procedure to end on a less painful note. Participants in

their study had two experiences, separated by 7 min: (a) immersion of one hand in 14 °C (54 °F) water for 60 s and (b) the same experience extended by 30 s, during which time the water temperature was raised to a slightly less painful 15 °C (56 °F). Subsequently, participants chose which of these two procedures to repeat. More than two thirds of the participants preferred the longer procedure, and preference for the longer procedure rose to 81% among participants who reported diminishing pain during the last 30 s of the longer trial. Although thought-listing data were not obtained in these studies, Kahneman and his colleagues' findings point to an important principle that may bias any retrospective cognitive assessment that relies on self-reported data.

Finally, when thought listings are to be scored by independent judges, the analyses of listed thoughts can also be arduous and time-consuming work. An investigator must (a) select dimension(s) to use to content analyze the listed thoughts (e.g., positive, neutral-irrelevant, or negative statements), (b) train independent judges to categorize each listed thought along the relevant dimension(s), (c) verify high interrater reliabilities, and (d) extract appropriate measures from these data. Although individuals are typically instructed to list one thought per box on a response form, the investigator may also need to *unitize* the listed thoughts (i.e., parse them into elemental units, such as a complete thought or affective expression) if the thoughts are listed in an unstructured format.

Measures Derived From Thought-Listing Data to Infer Cognitive Representations and Processes

The most common measures extracted from thought-listing data involve the scoring of the number of positive, negative, and neutral-irrelevant self-statements. In this section, we consider issues surrounding the quantification of these self-statements, additional dimensions that have proved informative in studies of psychopathology, measures of cognitive structure that can be extracted from thought-listing data (e.g., adjusted ratio of clustering [ARC] scores), and implicit memory measures.

Valence Scoring

Valence scoring of thought listings has proved useful, but investigators have differed in their treatment of the resulting measures. One method of dealing with these data has been simple frequency counts of positive, negative, and neutral-irrelevant thoughts (e.g., in a repeated measures design). A related measure that has proven informative in the clinical literature is Schwartz and Garamoni's (1986, 1989; Schwartz, 1986) state-of-mind measure. Rather than simple counts of the positive and negative thoughts listed, this model combines both positive and negative thoughts into a single index of mental health using the ratio of positive thoughts/(positive + negative thoughts).

Although most studies have either analyzed the number of various categories of thoughts (e.g., positive, negative) or state-of-mind ratios, the thought-listing data in some studies have been expressed as a difference score between positive and negative thoughts or a ratio of this difference score divided by the sum of the listed thoughts. These measures, which provide an index of the net affective disposition evident in the thought-

listing data, are consistent with a long tradition of conceptualizing affective processes as captured by a single bipolar (positive-negative) dimension (e.g., Osgood, Suci, & Tannenbaum, 1957; Thurstone, 1931). This approach, however, also treats positive and negative evaluative processes as equivalent, reciprocally activated, and interchangeable. Contrary to conceptualizations of this evaluative process as bipolar, recent evidence suggests that distinguishable motivational systems underlie assessments of the positive and negative significance of a stimulus (Cacioppo & Bertson, 1994; Cacioppo, Gardner, & Bertson, 1997). Thus, thought-listing data may reflect variations in terms of the production of positive self-statements and the production of negative self-statements. Low numbers of positive and negative listed thoughts in response to a stimulus or situation reflect neutrality or indifference, whereas high numbers of positive and negative listed thoughts would more likely reflect ambivalence or conflicted states. Simple difference scores and related measures that treat positive and negative thoughts as interchangeable mask these differing psychological states.

Accumulating research also suggests that positive and negative self-statements are distinguishable (stochastically and functionally independent). As noted above, Kendall (1984) reviewed evidence that psychological adjustment may be more closely related to the absence (or diminution) of negative thinking than the presence (or augmentation) of positive thinking—a phenomenon he termed the “power of nonnegative thinking” (p. 69). Briggs, Piliavin, Lorentzen, and Becker (1986) found that the greater an individual's perceptions of personal risk in donating bone marrow, the less willing he or she was to join a bone-marrow registry. In addition, several studies have shown that beliefs regarding the negative (rather than positive) consequences of donating blood differentiate individuals who intend versus those who do not intend to become (or continue as) blood donors (e.g., Condie, Warner, & Gillman, 1976; Edwards & Zeichner, 1985). In a related note, Parsi and Katz (1986) found that the positive and negative responses to posthumous organ donation represented two independent dimensions and that individuals' negative beliefs about organ donation were more strongly predictive of who would sign donor cards than were their positive beliefs (cf. Cacioppo & Gardner, 1993). Therefore, thought-listing data scored along a valence dimension may be represented more completely when the measures of positive and negative thoughts are treated as separable rather than as necessarily falling along a bipolar continuum.

Measurement contexts and comparison conditions are an integral part of the thought-listing technique and are features that contribute to the efficacy of the technique in studies of psychopathology. Theoreticians in clinical psychology have long held, however, that the content of readily accessible self-statements does not provide a complete analysis of cognition (Arnkoff, 1980; Kelly, 1955; McDermut, Haaga, & Shayne, 1991). Recent work on the organization of cognitive outputs (e.g., thought listings, recall) offers another means of examining cognitive processes of which an individual or client may be unaware and yet are therapeutically important. In this section, we review measures of cognitive organization that can be—but have not heretofore been—applied to thought-listing data.² Our aim is to

² Most studies on thought listings have used valence scores; we therefore focus on thought listings that have been scored along the valence

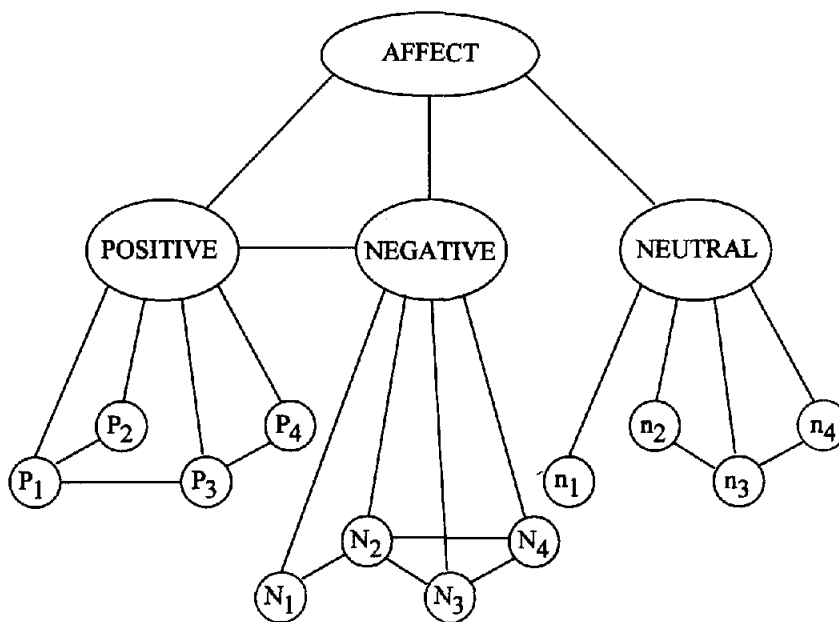


Figure 1. Affective organizational structure of yesterday's activities (links represent mental pathways, levels indicate hierarchy, and letters with subscripts indicate activities).

demonstrate how these indices can be calculated from thought-listing data and to illustrate how they may be useful in distinguishing adaptive from maladaptive memory structures and thought-production processes. Thus, our goal is to complement existing measures of cognitive structure such as Linville's (1985, 1987) measure of self-complexity, Gara and colleagues' (Gara, 1990; Gara, Rosenberg, & Goldberg, 1992; Gara et al., 1993) use of hierarchical classes cluster analysis (HICLAS) of free-response self-descriptions, and Segal and colleagues' (Segal, Hood, Shaw, & Higgins, 1988; Segal & Vella, 1990) use of Stroop tests to study depressive self-schemata. The application of individual differences multi-dimensional scaling (INDSCAL; Carroll & Chang, 1970) to thought-listing data also is discussed below.

Thought-Listing Order

If individuals are asked how they spent their time yesterday, there are a number of different ways they could answer the question. For example, they might think of all the possible things that they do in a day and select from this list those activities that they actually engaged in yesterday. Such a strategy, however, would be high in effort and low in yield. Consequently, the

strategy that most individuals adopt is to categorize their day into units based on emotional response, time of day, type of activity, accompanying people, and so forth. Instances of these categories from the previous day are then accessed sequentially, as each type of category cues different information from memory (see Figure 1). As a consequence of this sort of strategic retrieval process, it can be possible to infer the organization of the information in memory by examining the order in which that information is recalled. If a number of negative self-statements are recalled sequentially, for example, one can infer that these self-depictions are stored in the same category in memory. If individuals with depression are more likely to recall negative self-statements sequentially and positive self-statements nonsequentially than are nondepressed individuals, then an investigator would learn not only about differences between individuals with and without depression in the number of positive and negative self-statements but also about differences in the cognitive representations of these individuals.

The most commonly used measure of clustering is the ARC score, developed by Roenker, Thompson, and Brown (1971). The following is the formula for the ARC score:

$$(CR_{\text{obs}} - CR_{\text{exp}}) / (CR_{\text{max}} - CR_{\text{exp}}), \quad (1)$$

where CR_{obs} is observed number of category repetitions (number of times one item from a category follows another item from that same category), CR_{exp} is expected or chance number of category repetitions, and CR_{max} is maximum possible number of category repetitions (which equals total number of items recalled - number of categories represented in the recall protocol). CR_{exp} , in turn, is computed by summing the square of number of items recalled from each category, dividing this sum by total number of items recalled, and subtracting 1 from the final value.

dimension in the following discussions. However, content scoring can be performed on any dimension that is of interest to an investigator (e.g., source, target; cf. Cacioppo & Petty, 1981). It is worth noting, therefore, that the measures outlined below can easily be extended to whatever content dimension is of interest to an investigator. Indeed, some of the measures discussed in the following sections are designed to unveil relevant dimensions, without any a priori specifications by investigators, that are based on features of the way in which the thoughts were elicited or listed.

To provide an example, imagine that socially anxious and nonanxious individuals are asked to list their thoughts after preparing to speak with a blind date. These thoughts could be categorized by judges in terms of whether they depict positive thoughts about the self (S+), positive thoughts about others (O+), negative thoughts about the self (S-), negative thoughts about others (O-), neutral thoughts about the self (S), neutral thoughts about others (O), and irrelevant thoughts (I). Imagine that a particular individual produced the following order of 18 thoughts:

I O S O- O- O- O+ O- S+ I S- S- S- S- O- S S- I.

The observed number of category repetitions is five because category repetitions appear in the 5th, 6th, 12th, 13th, and 14th items listed above. The expected number of category repetitions is the sum of the square of number of items recalled from each category—that is, 3^2 (for category I) + 1^2 (for category O) + 2^2 (for category S) + 5^2 (for category O-) + 1^2 (for category O+) + 5^2 (for category S-) + 1^2 (for category S+)—divided by total number of items recalled (i.e., 18) - 1. Thus, the expected number of category repetitions is $(\frac{66}{18}) - 1 = 2.67$. The maximum possible number of category repetitions is 11 and is derived by taking total number of items recalled (18) and subtracting number of categories represented in the recall protocol (7). Thus, according to Equation 1, the ARC score for this recall profile would be $(5.00 - 2.67)/(11.00 - 2.67)$ or 0.28. ARC scores can range from a value of zero when clustering is at chance levels to one when clustering is perfect. Because people's internal dialogues tend to be abbreviated and mercurial, perfect category repetition is quite atypical and an ARC score of 0.28 indicates a moderate degree of clustering.

ARC scores have not been used in clinical studies to date, even though there are a variety of issues that could be addressed through their use. For example, whether anxious and nonanxious (or with a depressed and nondepressed mood) individuals show different degrees of organization of valenced self and other information could be explored. It is also possible to examine more complex questions, such as whether ARC scores increase or decrease when positive and negative thoughts about the self are collapsed into a single category. If collapsing positive and negative self-thoughts into a single category leads to an increase in ARC scores, one can infer that these thoughts are clustered together in a single category. However, if collapsing these thoughts into a single category leads to a decrease in ARC scores, one can infer that positive and negative self-thoughts are categorized separately. When we collapse the positive and negative self-thoughts together in the above thought-listing protocol, we find that the ARC score drops to 0.20. This suggests that positive and negative thoughts about the self are segregated in memory. When we collapse the positive and negative other thoughts together in the above thought-listing protocol, in contrast, we find that the ARC score rises to 0.43. This suggests that positive and negative thoughts about others are clustered together in memory. Such comparisons of ARC scores could also be computed separately for anxious and nonanxious individuals to determine whether social anxiety leads to different types of clustering.

Related techniques have been used to make inferences about

self-structure from free-response protocols. Trafimow, Triandis, and Goto (1991), for example, presented participants with Kuhn and McPartland's (1954) self-attitude task, in which participants provide 20 consecutive completions to the sentence stem, "I am . . ." Trafimow et al. then computed the conditional probability that private versus collective aspects of the self would be recalled when the previously recalled statement was either private or collective in nature. These conditional probabilities provided support for their hypothesis that private and collective aspects of the self are stored in separate locations in memory (see also Trafimow & Fishbein, 1995).³

Experiments like those of Trafimow et al. (1991) demonstrate the use of measures of clustering that are based on recall order for uncovering the structure of self-relevant information. Furthermore, due to the nature of the thought-listing task, clustering measures such as ARC scores provide important advantages to investigators who are limited by conventional methods. For example, because the ARC score does not necessitate the ability to introspect about one's own self-structure, it provides a useful complement to measures such as Linville's (1985, 1987) self-complexity score, which requires participants to sort their own characteristics into meaningful groups. In addition, because the ARC score can be computed from thought listings associated with almost any context, it provides a useful complement to measures such as Segal's (Segal et al., 1988; Segal & Vella, 1990) Stroop task and Gara's (1990; Gara et al., 1993) HICLAS analysis, which require a series of self and other ratings on a variety of trait dimensions. Finally, the ARC score has the added advantage that the investigator need not have an a priori understanding of how participants organize their thoughts because various hypotheses can be generated and tested from the items and orderings present in the thought listings themselves.

One potential problem with the ARC score is that the investigator may not accurately deduce the categories that the individual is using when retrieving information from memory. For example, if an investigator assumes that a participant is using affect-based categories to retrieve information but the participant is actually using activity-based categories, then perfect clustering by the participant might be scored as chance clustering by the investigator. This difficulty can potentially be ameliorated by asking participants (or judges) to look over the thought listing and categorize the thoughts according to any of a variety of different schemes. The categories provided by the participant (or judge) can then be used to compute ARC scores. A different sort of solution to this problem can be provided by computing *subjective* measures of clustering, which assess the extent to which items recalled adjacently on one thought-listing task are again recalled adjacently on a later thought-listing task. When items are consistently recalled in the same order across retrieval tasks, it is assumed that this order represents an idiosyncratic structure on the part of the participant (Shuell, 1969). Thus,

³ It is worth noting, however, that under some circumstances, conditional probabilities can be misleading about the underlying structure of memory, and thus ARC scores are typically preferable (e.g., see Skowronski & Welbourne, 1997). Occasionally, ARC scores result in a negative number when fewer repetitions emerge than are expected by chance. Investigators typically trim such negative numbers at zero, however, assuming that they represent clustering at chance levels.

the measurement of subjective clustering requires sequential thought-listing tasks, and relies on a revised formula for computing the subjective ARC or ARC' score (Pellegrino, 1971).

A second potential problem with the ARC score (and other retrieval-based measures of structure) is that sequential recall of related items might indicate not that the items are clustered together in memory but that order was imposed on them by the retrieval task (Anderson, 1976; Shuell, 1969). Thus, it may be that the activities of the previous day were not stored in affect-based categories, yet the affect-based retrieval strategy adopted by the individual would, nevertheless, suggest such a structure for the information in storage. For example, children, pets, money, photographs, and important documents may not be items that are typically stored in a single category in memory, yet if people were asked what they would bring out of their house in the case of a fire, these are the types of items that they would retrieve (Barsalou, 1983). Such ad hoc categories are easily created by retrieval instructions, and the difficulty lies in determining when retrieval order represents the underlying structure of storage and when it represents order induced by the retrieval process.

Although there is no simple solution to these problems, the investigator can have greater confidence that ARC and ARC' scores represent underlying structure in storage when similar structures emerge from thought-listing tasks that are cued through different types of instructions and at different times. The important point for investigators is to remain cognizant that the same information may well be structured in multiple ways in memory, and thus inconsistent clusterings produced across thought listings may each be meaningful indicators of underlying structure.

Multidimensional Scaling

Multidimensional scaling refers to a class of techniques that use proximities (i.e., a number that indicates how similar two objects are perceived to be) among objects as input. The primary output is a spatial representation of points, in which each point represents an object and the overall configuration represents the conceptual structure or organization of the objects (Kruskal & Wish, 1978). INDSCAL (Carroll & Chang, 1970) is a variant of multidimensional scaling that has been used in clinical psychology to study topics ranging from social anxiety (Goldfried, Padawer, & Robins, 1984) and unassertiveness (Rudy, Merluzzi, & Henahan, 1982) to opiate addiction (Chaney & Roszell, 1985) and smoking relapse (McDermut et al., 1991).

In an illustrative study, McDermut et al. (1991) used INDSCAL in a prospective study to explore whether exsmokers' conceptual organization of situations in which they might be tempted to smoke would predict smoking relapse. Recent exsmokers rated the similarity of 44 high-risk situations in terms of how difficult it would be to avoid smoking, and the resulting similarity judgments were subjected to INDSCAL. Two dimensions emerged in this scaling: (a) a negative-affect dimension, which reflected the extent to which the situation was associated with negative emotions or the absence of positive emotions, and (b) a social factors dimension, which reflected the extent to which the situation was characterized by social pressure versus solitary activities. Furthermore, the similarity judgments of indi-

viduals who began smoking again within the succeeding 12 months, in contrast to those who abstained continuously, placed a greater emphasis on the negative-affect dimension than the social factors dimension.

In a related vein, multidimensional scaling techniques such as INDSCAL may prove useful in mapping the extent to which participants rely on different categories in rating their thoughts and the extent to which their thoughts cluster together in a "similarity space." The use of INDSCAL may benefit from a modification of the typical thought-listing procedure because the listed thoughts differ in number and content within and across groups. Recall that Glass et al. (1982) used the thought-listing technique to identify discriminating items (self-statements) in their 30-item scale for the assessment of the cognitive correlates of heterosocial anxiety (SISST). The approach used by Glass et al. can be used to generate a set of self-statements relevant to any given clinical problem. Once generated, the set of self-statements can be used to elicit similarity judgments in an effort to map individual and group differences in the conceptual organization of the thought listings.

Response Latencies

If the thought listings are obtained in a manner that allows one to calculate when each was listed (e.g., using computers or tape recorders), listing latencies can be computed to provide a window on the categorical structure of information in memory. When people provide responses within a single category, response latencies are shorter than when people switch to a new category. For example, if individuals use an affect-based strategy to retrieve information about the previous day's activities, they might first access the general category for positive activities, before following the mental pathways that lead to the specific activities that are stored there. Once the search for information has exhausted the specific items in that category, they must mentally move back "up" to the general category level, access a new general category, and then move back "down" to the specific items within that new category. As can be seen in Figure 1, this process of moving from one category to the next requires that the individuals traverse a greater number of mental pathways than when moving between items within a category. As a consequence, less time elapses between the retrieval of two different items from the same category than between the retrieval of two different items from two different categories (Seamon & Chumbley, 1977). This interitem response interval can therefore be used as an additional indicator of which pieces of information are clustered together in memory and which are not.

Although this procedure does not address the difficulty of when structure is induced by retrieval versus based on storage, it has the advantage that the investigator does not need to be able to determine a priori which items should be categorized together. In contrast to ARC scores, which require the investigator to make decisions about relationships between items to compute the clustering measure, the latencies can simply be examined and inter- and intracategory changes can be directly inferred from the data. It is worth noting, however, that differences in latencies are likely to be continuous in nature, and thus it will not always be easy to differentiate long from short responses. Furthermore, because not every long latency is evidence of a

category change, this measure still requires interpretation concerning which items appear to be clustered together in a particular individual's memory. The advantage of the measure nonetheless remains because this interpretation can be based on idiosyncratic features of the items retrieved and their associated latencies and does not require an a priori understanding of how an individual has organized information in his or her memory. Particularly in combination with an ARC or ARC' score, this measure can provide additional evidence concerning the idiosyncratic structure of storage.

Implicit Memory Measures

The thought-listing techniques described above provide a number of different ways of probing cognitive content and organization, but they are all alike in that they rely on intentional retrieval to report the contents of internal dialogues. This intentional retrieval component has the potential to introduce bias into the measure, either through an unwillingness (e.g., thought or response suppressing) or an inability (e.g., forgetting or repressing) to report certain types of information. Recent research on various memory measures suggests that thought listings can be measured in ways that avoid the intentional retrieval component and thereby reduce or eliminate the impact of such biases. In particular, implicit measures (Graf & Schacter, 1985; Roediger, 1990) tap a person's memory and cognitions without relying on conscious or intentional recollection. To the extent that performance (e.g., a listed thought or thoughts) on a task is influenced by experience, then it must be the case that the individual has some memory for that experience, conscious or otherwise.

For instance, if individuals were asked to memorize a list of words, such as *sun*, *sand*, *wind*, *moon*, and *waves* and then were asked to list the first laundry detergent that comes to mind, a preponderance typically respond with "Tide." If the individuals were instead asked to memorize the list of words *fun*, *balloons*, *clown*, *smile*, and *happy*, they more typically would respond with "Cheer". The task of memorizing the word list created an accessible cognitive organization that predictably affected information processing in response to the question about laundry detergents. When these individuals were asked why they listed a particular detergent—that is, to identify the reason for their mental contents—most reported that they or someone they know uses that detergent, they know it is a popular brand, they recently saw an advertisement for the detergent, and so on; few if any reported that their preceding exposure to the word list primed them for the detergent they reported (cf. Nisbett & Wilson, 1977; and Wilson & Brekke, 1994).

Thought-listing and implicit memory measures have not been combined in this fashion to study psychological disorders, although doing so is fairly straightforward. This is because many implicit memory measures represent, in essence, specific methods for eliciting a thought listing or thought listings. For example, an individual might be told behaviors of a hypothetical person that are somewhat ambiguous with regard to whether they implicate social anxiety, such as "Did not speak to anyone at the party." The individual is then asked to list the trait or traits that come to mind. If the individual is chronically anxious, has been made to feel anxious, or has been induced to think

about the concept of *anxiety*, then anxiety should be accessible in memory and should be more likely to be listed as a descriptor of the ambiguous behavior, as compared with other possible descriptors such as *unfriendly* or *tired* (Higgins et al., 1982; Smith & Branscombe, 1988). A client in a depressed mood or an individual who has been made to feel or think about sadness might respond to the same stimulus with trait terms such as *sad* or *rejected*. Despite the comorbidity of depression and anxiety, careful construction of such materials may allow the researcher or practitioner to assess whether concepts related to depression or anxiety are more accessible for a particular participant or client (cf. Lehnert, Overholser, & Adams, 1996).

Other types of implicit memory measures simply request the individual to generate exemplars of different categories of events or behaviors (Rappold & Hashtroudi, 1991). Although such tasks do not explicitly request individuals to generate items from their own experiences, such items from personal experience are, nevertheless, more likely to appear in these thought-listing protocols. Thus, exemplars that people have thought about or encountered recently are more likely to be chosen for this thought-listing task than exemplars that have not been thought about or encountered recently. Consistent with this suggestion, research has shown that free association can tap individuals' memories for prior specific instances or episodes (Kihlstrom, Schacter, Cork, Hurt, & Behr, 1990; Schacter, 1985).

Although such implicit memory (i.e., thought) assessments often tap the same information that would be provided by intentional recollection tasks, adapting implicit memory tasks to elicit thought listings in patient populations has the potential to provide additional information that individuals are unwilling or unable to report. For example, when under the influence of alcohol, individuals show preserved memory on implicit but not explicit measures (Hashtroudi, Parker, DeLisi, Wyatt, & Mutter, 1984). Evidence for this preservation of implicit memory can even be seen in patients with severe anterograde and retrograde amnesia, such as is associated with closed head injuries or the alcoholic Korsakoff's syndrome. These individuals often cannot list anything about an episode in which information was learned, yet they show evidence that the learned information is accessible in thought listings when elicited using implicit memory assessments, such as the free association task (Gardner, Boller, Moreines, & Butters, 1973; McAndrews, Glisky, & Schacter, 1987; Schacter, 1987; Shimamura & Squire, 1984). Indeed, through the use of the free association task, memory has been demonstrated for information learned while patients were under surgical anesthesia (Kihlstrom et al., 1990). Finally, implicit thought assessments have also demonstrated memory across different personalities in a patient with dissociative identity disorder, who showed no ability to consciously retrieve information across those personalities (Nissen, Ross, Willingham, Mackenzie, & Schacter, 1988). Given this early evidence, the integration of the thought-listing and implicit-memory techniques may contribute to a more complete analysis of cognition in clinical and counseling psychology.

Conclusion

Sir Arthur Eddington (1939) described a hypothetical scientist who sought to determine the size of the various fish in the

sea. The scientist began by weaving a 2-in. mesh net and setting sail across the seas. Each day he would cast his net, measuring, recording, and analyzing the results of each catch. After extensive sampling, the scientist concluded that there were no fish smaller than 2 in. in the sea. The moral of Sir Eddington's analogy is that measurement instruments are shaped by developers' implicit theories of the phenomenon they wish to study and that once developed, measurement instruments shape data and theory in fundamental ways. Rating scales and closed-ended questionnaires, although efficient, have the potential cost of hindering recognition of interesting and relevant phenomena that fall outside their designed purpose. Although thought listings can be difficult and time consuming to reduce and analyze, they offer a potentially important complement to cognitive assessments that limit what aspects of a phenomenon are quantified. Accordingly, the thought-listing technique has provided an effective open-ended means of sampling mental contents in different populations; from these contents, measures of cognitive organization and processes can be calculated. The thought-listing technique is also a work in progress. As evidenced by this review, investigators in clinical and counseling psychology have been creative in the topics and paradigms in which they have used the thought-listing technique as well as in the methods of quantifying thought-listing data. This creativity, combined with provocative new methods for obtaining and quantifying thought listings to better map people's cognitive representations and coping processes, bodes well for the continued evolution of the thought-listing technique in clinical and counseling psychology.

References

- Altmaier, E. M., Leary, M. R., Halpern, S., & Sellers, J. E. (1985). Effects of stress inoculation and participant modeling on confidence and anxiety: Testing predictions of self-efficacy theory. *Journal of Social and Clinical Psychology, 3*, 500-505.
- Amsel, R., & Fichten, C. S. (1990). Ratio versus frequency scores: Focus of attention and the balance between positive and negative thoughts. *Cognitive Therapy and Research, 14*, 257-277.
- Anderson, J. R. (1976). *Language, memory, and thought*. Hillsdale, NJ: Erlbaum.
- Arnkoﬀ, D. B. (1980). Psychotherapy from the perspective of cognitive theory. In M. J. Mahoney (Ed.), *Psychotherapy process: Current issues and future directions* (pp. 339-361). New York: Plenum Press.
- Arnkoﬀ, D. B., & Glass, C. R. (1989). Cognitive assessment in social anxiety and social phobia. *Clinical Psychology Review, 9*, 61-74.
- Barsalou, L. W. (1983). Ad hoc categories. *Memory and Cognition, 11*, 211-217.
- Blackwell, R. T., Galassi, J. P., Galassi, M. D., & Watson, T. E. (1985). Are cognitive assessment methods equal? A comparison of think aloud and thought listing. *Cognitive Therapy and Research, 9*, 399-413.
- Blankstein, K. R., & Flett, G. L. (1990). Cognitive components of test anxiety: A comparison of assessment and scoring methods. *Journal of Social Behavior and Personality, 5*, 187-202.
- Blankstein, K. R., Toner, B. B., & Flett, G. L. (1989). Test anxiety and the contents of consciousness: Thought-listing and endorsement methods. *Journal of Research in Personality, 23*, 269-286.
- Briggs, N. C., Piliavin, J. A., Lorentzen, D., & Becker, G. A. (1986). On willingness to be a bone marrow donor. *Transfusion, 26*, 324-330.
- Bruch, M. A., Gorsky, J. M., Collins, T. M., & Berger, P. A. (1989). Shyness and sociability revisited: A multicomponent analysis. *Journal of Personality and Social Psychology, 57*, 904-915.
- Bruch, M. A., Hamer, R. J., & Kaflowitz-Linder, N. G. (1992). Cognitive balance and assertiveness: Testing the construct validity of the states of mind categories. *Cognitive Therapy and Research, 16*, 653-667.
- Cacioppo, J. T., Andersen, B. L., Turnquist, D. C., & Tassinari, L. G. (1989). Psychophysiological comparison theory: On the experience, description, and assessment of signs and symptoms. *Patient Education and Counseling, 13*, 257-270.
- Cacioppo, J. T., & Berntson, G. G. (1994). Relationship between attitudes and evaluative space: A critical review, with emphasis on the separability of positive and negative substrates. *Psychological Bulletin, 115*, 401-423.
- Cacioppo, J. T., & Gardner, W. L. (1993). What is underlying medical donor attitudes and behavior? *Health Psychology, 12*, 269-271.
- Cacioppo, J. T., Gardner, W. L., & Berntson, G. G. (1997). Beyond bipolar conceptualizations and measures: The case of attitudes and evaluative space. *Personality and Social Psychology Review, 1*, 3-25.
- Cacioppo, J. T., Glass, C. R., & Merluzzi, T. V. (1979). Self-statements and self-evaluations: A cognitive-response analysis of heterosocial anxiety. *Cognitive Therapy and Research, 3*, 249-262.
- Cacioppo, J. T., & Petty, R. E. (1981). Social psychological procedures for cognitive response assessment: The thought-listing technique. In T. V. Merluzzi, C. R. Glass, & M. Genest (Eds.), *Cognitive assessment* (pp. 309-342). New York: Guilford Press.
- Carroll, J. D., & Chang, J. J. (1970). Analysis of individual differences in multidimensional scaling via an *n*-way generalization of "Eckert-Young" decomposition. *Psychometrika, 35*, 283-319.
- Chaney, E. F., & Roszell, D. K. (1985). Coping in opiate addicts maintained on methadone. In S. Shiffman & T. A. Wills (Eds.), *Coping and substance use* (pp. 267-293). Orlando, FL: Academic Press.
- Clark, J., & Arkowitz, H. (1975). Social anxiety and the self-evaluation of interpersonal performance. *Psychological Reports, 36*, 211-221.
- Condie, S. J., Warner, W. K., & Gillman, D. C. (1976). Getting blood from collective turnips: Volunteer donation in mass blood drives. *Journal of Applied Psychology, 61*, 290-294.
- Crocker, P. R. E. (1989). A follow-up of cognitive-affective stress management training. *Journal of Sport and Exercise Psychology, 11*, 236-242.
- Csikszentmihalyi, M., & Larson, R. (1987). Validity and reliability of the experience-sampling method [Special issue]. *Journal of Nervous and Mental Disease, 175*, 526-536.
- Daleiden, E. L., Vasey, M. W., & Williams, L. L. (1996). Assessing children's states of mind: A multitrait, multimethod study. *Psychological Assessment, 8*, 125-134.
- Eddington, A. (1939). *The philosophy of physical science*. New York: Macmillan.
- Edwards, P. W., & Zeichner, A. (1985). Blood donor development: Effects of personality, motivational, and situational variables. *Personality and Individual Differences, 6*, 743-751.
- Elliott, T. R., MacNair, R. R., Herrick, S. M., Yoder, B., & Byrne, C. A. (1991). Interpersonal reactions to depression and physical disability in dyadic interactions. *Journal of Applied Social Psychology, 21*, 1293-1302.
- Ellis, H. C., Seibert, P. S., & Herbert, B. J. (1990). Mood state effects on thought listing. *Bulletin of the Psychonomic Science, 28*, 147-150.
- Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review, 82*, 215-251.
- Fichten, C. S., Amsel, R., & Robillard, K. (1988). Issues in cognitive assessment: Task difficulty, reactivity of measurement, thought listing versus inventory approaches, and sequences versus frequency counts. *Behavioral Assessment, 10*, 399-425.

- Gara, M. A. (1990). A set-theoretical model of person perception. *Multivariate Behavioral Research, 25*, 275-293.
- Gara, M. A., Rosenberg, S., & Goldberg, L. (1992). *DSM-III-R* as a taxonomy: A cluster analysis of diagnoses and symptoms. *Journal of Nervous and Mental Disease, 180*, 11-19.
- Gara, M. A., Woolfolk, R. L., Cohen, B. D., Goldston, R. B., Allen, L. A., & Novalany, J. (1993). Perception of self and other in major depression. *Journal of Abnormal Psychology, 102*, 93-100.
- Gardner, H., Boller, F., Moreines, J., & Butters, N. (1973). Retrieving information from Korsakoff patients: Effects of categorical cues and reference to the task. *Cortex, 9*, 165-175.
- Glass, C. R., & Arnkoff, D. B. (1983). Cognitive set and level of anxiety: Effects on thinking processes in problematic situations. *Cognitive Therapy and Research, 7*, 529-542.
- Glass, C. R., & Arnkoff, D. B. (1994). Validity issues in self-statement measures of social phobia and social anxiety. *Behavioral Research and Therapy, 32*, 255-267.
- Glass, C. R., & Furlong, M. (1990). Cognitive assessment of social anxiety: Affective and behavioral correlates. *Cognitive Therapy and Research, 14*, 365-384.
- Glass, C. R., Merluzzi, T. V., Biever, J. L., & Larsen, K. H. (1982). Cognitive assessment of social anxiety: Development and validation of a self-statement questionnaire. *Cognitive Therapy and Research, 6*, 37-55.
- Goldfried, M. R., Padawer, W., & Robins, C. (1984). Social anxiety and the semantic structure of heterosocial interactions. *Journal of Abnormal Psychology, 93*, 87-97.
- Graf, P., & Schacter, D. A. (1985). Implicit and explicit memory for new associations in normal and amnesic subjects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*, 501-518.
- Halford, W. K., & Sanders, M. R. (1988). Assessment of cognitive self-statements during marital problem solving: A comparison of two methods. *Cognitive Therapy and Research, 12*, 515-530.
- Hashtroudi, S., Parker, E. S., DeLisi, L. E., Wyatt, R. J., & Mutter, S. A. (1984). Intact retention in acute alcohol amnesia. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*, 156-163.
- Heimberg, R. G. (1994). Cognitive assessment strategies and the measurement of outcome of treatment for social phobia. *Behavioral Research and Therapy, 32*, 269-280.
- Heimberg, R. G., Bruch, M. A., Hope, D. A., & Dombek, M. (1990). Evaluating the state of mind model: Comparison to an alternative model and effects of method of cognitive assessment. *Cognitive Therapy and Research, 14*, 543-557.
- Heimberg, R. G., Dodge, C. S., Hope, D. A., Kennedy, C. R., Zollo, L., & Becker, L. E. (1990). Cognitive behavioral group treatment of social phobia: Comparison to a credible placebo control. *Cognitive Therapy and Research, 14*, 1-23.
- Heppner, P. P., Rosenberg, J. L., & Hedgespeth, J. (1992). Three methods in measuring the therapeutic process: Clients' and counselors' constructions of the therapeutic process versus actual therapeutic events. *Journal of Counseling Psychology, 39*, 20-31.
- Higgins, E. T., King, G. A., & Mavin, G. H. (1982). Individual construct accessibility and subjective impressions and recall. *Journal of Personality and Social Psychology, 43*, 35-47.
- Hines, P. L., Stockton, R., & Morran, D. K. (1995). Self-talk of group therapists. *Journal of Counseling Psychology, 42*, 242-248.
- Huber, J. W., & Altnaier, E. M. (1983). An investigation of the self-statement systems of phobic and nonphobic individuals. *Cognitive Therapy and Research, 7*, 355-362.
- Kahneman, D., Fredrickson, D. L., Schreiber, C. A., & Redelmeier, D. A. (1993). When more pain is preferred to less: Adding a better ending. *Psychological Science, 4*, 401-405.
- Kelly, G. H. (1955). *The psychology of personal constructs*. New York: Norton.
- Kendall, P. C. (1984). Behavioral assessment and methodology. In G. T. Wilson, C. M. Franks, K. D. Brownell, & P. C. Kendall (Eds.), *Annual review of behavior therapy: Theory and practice* (Vol. 9, pp. 47-86). New York: Guilford Press.
- Kihlstrom, J. F., Schacter, D. L., Cork, R. C., Hurt, C. A., & Behr, S. E. (1990). Implicit and explicit memory following surgical anesthesia. *Psychological Science, 1*, 303-306.
- Kolligian, J., Jr., & Sternberg, R. J. (1991). Perceived fraudulence in young adults: Is there an "imposter syndrome"? *Journal of Personality Assessment, 56*, 308-326.
- Kruskal, J. B., & Wish, M. (1978). *Multidimensional scaling*. Beverly Hills, CA: Sage.
- Kuhn, M. H., & McPartland, T. (1954). An empirical investigation of self-attitudes. *American Sociological Review, 19*, 58-76.
- Lehnert, K. L., Overholser, J. C., & Adams, D. M. (1996). The cognition rating form: A new approach to assessing self-generated cognitions in adolescent sentence completions. *Psychological Assessment, 8*, 172-181.
- Linville, P. W. (1985). Self-complexity and affective extremity: Don't put all of your eggs in one cognitive basket [Special issue]. *Social Cognition, 3*, 94-120.
- Linville, P. W. (1987). Self-complexity as a cognitive buffer against stress-related illness and depression. *Journal of Personality and Social Psychology, 52*, 663-676.
- Long, B. C. (1984). Aerobic conditioning and stress inoculation: A comparison of stress-management interventions. *Cognitive Therapy and Research, 8*, 517-542.
- Mahone, E. M., Bruch, M. A., & Heimberg, R. G. (1993). Focus of attention and social anxiety: The role of negative self-thoughts and perceived positive attributes of the other. *Cognitive Therapy and Research, 17*, 209-224.
- Martzke, J. S., Andersen, B. L., & Cacioppo, J. T. (1987). Cognitive assessment of anxiety disorders. In L. Michelson & M. Ascher (Eds.), *Cognitive-behavioral assessment and treatment of anxiety disorders* (pp. 62-88). New York: Guilford Press.
- McAndrews, M. P., Glisky, E. L., & Schacter, D. L. (1987). When priming persists: Long-lasting implicit memory for a single episode in amnesic patients. *Neuropsychologia, 25*, 497-506.
- McDermut, W., Haaga, D. A. F., & Shayne, V. T. (1991). Schemata and smoking relapse. *Behavior Therapy, 22*, 423-434.
- Melchior, L. A., & Cheek, J. M. (1990). Shyness and anxious self-preoccupation during a social interaction. *Journal of Social Behavior and Personality, 5*, 117-130.
- Morran, D. K., Kurpius, D. J., & Brack, G. (1989). Empirical investigation of counselor self-talk categories. *Journal of Counseling Psychology, 36*, 505-510.
- Myszka, M. T., Galassi, J. P., & Ware, W. B. (1986). Comparison of cognitive assessment methods with heterosocially anxious college women. *Journal of Counseling Psychology, 33*, 401-407.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84*, 231-259.
- Nissen, M. J., Ross, J. L., Willingham, D. B., Mackenzie, T. B., & Schacter, D. L. (1988). Memory and awareness in a patient with multiple personality disorder. *Brain and Cognition, 8*, 117-134.
- Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning*. Urbana: University of Illinois Press.
- Parsi, N., & Katz, I. (1986). Attitudes toward posthumous organ donation and commitment to donate. *Health Psychology, 5*, 565-580.
- Pellegrino, J. W. (1971). A general measure of organization in free recall for variable unit size and internal sequential consistency. *Behavior Research Methods and Instrumentation, 3*, 241-246.
- Petty, R. E., Wells, G. L., & Brock, T. C. (1976). Distraction can enhance or reduce yielding to propaganda: Through disruption versus

- effort justification. *Journal of Personality and Social Psychology*, 34, 874–884.
- Rappold, V. A., & Hashtroudi, S. (1991). Does organization improve priming? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17, 103–114.
- Redelmeier, D. A., & Kahneman, D. (in press). Patients' memories of painful medical treatments: Real-time and retrospective evaluations of two minimally invasive procedures. *Pain*.
- Roediger, H. L., III (1990). Implicit memory: Retention without remembering. *American Psychologist*, 45, 1043–1056.
- Roenker, D. L., Thompson, C. P., & Brown, S. C. (1971). Comparison of measures for the estimation of clustering in free recall. *Psychological Bulletin*, 76, 45–48.
- Rudy, T. E., Merluzzi, T. V., & Henahan, P. T. (1982). Construal of complex assertion situations: A multidimensional scaling analysis. *Journal of Consulting and Clinical Psychology*, 50, 125–137.
- Schacter, D. L. (1985). Priming of old and new knowledge in amnesic patients and normal subjects. *Annals of the New York Academy of Sciences*, 444, 41–53.
- Schacter, D. L. (1987). Implicit expression of memory in organic amnesia: Learning of new facts and associations. *Human Neurobiology*, 6, 107–118.
- Schwartz, R. M. (1986). The internal dialogue: On the asymmetry between positive and negative coping thoughts. *Cognitive Therapy and Research*, 10, 591–605.
- Schwartz, R. M., & Garamoni, G. L. (1986). A structural model of positive and negative states of mind: Asymmetry in the internal dialogue. In P. C. Kendall (Ed.), *Advances in cognitive-behavioral research and therapy* (Vol. 5, pp. 1–62). New York: Academic Press.
- Schwartz, R. M., & Garamoni, G. L. (1989). Cognitive balance and psychopathology: Evaluation of an information processing model of positive and negative states of mind. *Clinical Psychology Review*, 9, 271–294.
- Seamon, J. G., & Chumbley, J. I. (1977). Retrieval processes for serial order information. *Memory and Cognition*, 5, 709–715.
- Sedikides, C. (1992). Mood as a determinant of attentional focus. *Cognition and Emotion*, 6, 129–148.
- Segal, Z. V., Hood, J. E., Shaw, B. F., & Higgins, E. T. (1988). A structural analysis of the self-schema construct in major depression. *Cognitive Therapy and Research*, 12, 471–485.
- Segal, Z. V., & Vella, D. D. (1990). Self-schema in major depression: Replication and extension of a priming methodology [Special issue]. *Cognitive Therapy and Research*, 14, 161–176.
- Shimamura, A. P., & Squire, L. R. (1984). Paired-associate learning and priming effects in amnesia: A neuropsychological study. *Journal of Experimental Psychology: General*, 113, 556–570.
- Shuell, T. J. (1969). Clustering and organization in free recall. *Psychological Bulletin*, 72, 353–374.
- Skowronski, J., & Welbourne, J. (1997). Conditional probability may be a flawed measure of associative strength. *Social Cognition*, 15, 1–12.
- Smith, E. R., & Branscombe, N. R. (1988). Category accessibility as implicit memory. *Journal of Experimental Social Psychology*, 24, 490–504.
- Stone, A. A., & Shiffman, S. (1994). Ecological momentary assessment (EMA) in behavioral medicine. *Annals of Behavioral Medicine*, 16, 199–202.
- Swanson, J. L., & Tokar, D. M. (1991). College students' perceptions of barriers to career development. *Journal of Vocational Behavior*, 38, 92–106.
- Tarico, V. S., Van Velzen, D. R., & Altmaier, E. M. (1986). Comparison of thought-listing rating methods. *Journal of Counseling Psychology*, 33, 81–83.
- Thurstone, L. L. (1931). The measurement of attitudes. *Journal of Abnormal Psychology*, 26, 249–269.
- Trafimow, D., & Fishbein, M. (1995). Do people really distinguish between behavioral and normative beliefs? *British Journal of Social Psychology*, 34, 257–266.
- Trafimow, D., Triandis, H. C., & Goto, S. H. (1991). Some tests of the distinction between the private self and the collective self. *Journal of Personality and Social Psychology*, 60, 649–655.
- Van Lange, P. A. M., & Rusbult, C. E. (1995). My relationship is better than—and not as bad as—yours is: The perception of superiority in close relationships. *Personality and Social Psychology Bulletin*, 21, 32–44.
- Velton, E. (1968). A laboratory task for induction of mood states. *Behavioral Research and Therapy*, 6, 473–482.
- Wilson, T. D., & Brekke, N. (1994). Mental contamination and mental correction: Unwanted influences on judgments and evaluations. *Psychological Bulletin*, 116, 117–142.

Received January 3, 1997

Revision received February 26, 1997

Accepted March 4, 1997 ■