

Articulated Thoughts During Simulated Situations: A Paradigm for Studying Cognition in Emotion and Behavior¹

Gerald C. Davison²

University of Southern California

Clive Robins and Marcia K. Johnson

State University of New York at Stony Brook

In spite of the widespread belief of cognitive behavior therapists and researchers that irrational thinking underlies much human psychological suffering, there is little if any direct evidence bearing on the assumption that people think in particular ways when confronted with stressful situations. A paradigm is proposed that appears capable of providing information about people's articulated thoughts as they occur in highly structured, experimenter-controlled situations. The results from an initial experiment indicate the utility of the paradigm in collecting data on how people think under both stressful and neutral conditions. The paradigm seems to offer great flexibility in examining thought processes under a wide range of conditions of interest to psychopathologists and cognitive researchers.

Cognitive approaches to therapeutic behavior change constitute one of the liveliest areas of clinical research in recent years (Bandura, 1977; Beck, 1967; Davison, 1966; Ellis, 1962; Goldfried & Davison, 1976; London, 1964; Mahoney, 1974; Meichenbaum, 1975; Mischel, 1968). The central assumption is that emotional reactions to environmental events are

¹This research was supported by NIMH grant MH 24237 and NSF grant BNS-7813054. Frank Durso served as one of the raters in the content analysis of the data, and helpful suggestions came from Donna Andrews and Marvin R. Goldfried.

²Address all correspondence to Gerald C. Davison, Department of Psychology, University of Southern California, Los Angeles, California 90089-1061.

primarily a function of how events are construed. This contention is often bolstered by referring to the pronouncements of philosophers and wise men of ancient Greece, Rome, and other early civilizations, such as Marcus Aurelius's "If thou art pained by an external thing, it is not this thing which disturbs thee but thine own judgment about it. And it is in thy power to wipe out this judgment now," or Epictetus's "[People] are disturbed not by things but by the views they take of them." The Buddha is reported to have made the even more radical statement, "All that we are is the result of what we have thought; it is founded on our thoughts, it is made up of our thoughts."

A therapeutic extension of these ideas is providing clients with particular ways of talking to themselves in order to ameliorate problems ranging from anxiety and depression to impulsivity and withdrawal. There are two basic ideas here: (1) Problems result from the way events are thought about, and (2) people can be helped by therapies directed at changing the things they say to themselves. A common misconception is that evidence for one of these ideas is necessarily evidence for the other. Our interest in the research reported here is the first question, namely, whether psychological problems arise from the way people construe their world.

EVIDENCE FROM THERAPEUTIC SITUATIONS

Typically in cognitive therapy, the investigator begins by explaining to subjects/clients what the theory is, to wit, that irrational "self-talk" underlies their problems. Clients are then encouraged to isolate such negative self-talk as a prelude to the therapeutic intervention. Consider the following excerpt from a widely cited investigation by Meichenbaum, evaluating the efficacy of a particular form of cognitive behavior modification for test anxiety. The therapy began with an "insight" phase, in which subjects were told that test anxiety "is the result of thoughts and verbalizations which are emitted both prior to and during the test situation. The subjects were informed that one of the goals of the therapy was for each subject to become aware of the self-verbalizations and self-instructions which he emitted in evaluative situations. . . . Over the course of eight sessions, this group discussed the following points: (a) the specific self-verbalizations group members had emitted in the pretreatment test situation . . . (b) the often irrational self-defeating and self-fulfilling aspects of such statements" (Meichenbaum, 1972, p. 374).

It appears that the primary purpose of this "insight" phase was to rationalize the intervention to the subjects and thereby enlist their cooperation in a treatment that required their active, sustained involvement.

Clinicians, regardless of their orientation, teach their clients, either directly or indirectly, the particular theory of psychopathology and of therapy that the clinician believes guides his/her therapeutic ministrations. Ellis (1977) is noted for his special forcefulness in persuading his clients to view their problems within his theoretical perspective: "Look for your 'shoulds,' look for your 'musts'" (p. 8); "irrational beliefs . . . stem . . . from your demanding, commanding statements about what *should* and *must* happen so that you can *absolutely* and *necessarily* get what you desire" (p. 9). "Once we [put the thousands of irrational ideas people have into a few general categories] . . . and then actively *look for* these categories, we can fairly quickly find them, show them to disturbed individuals, and also teach them how to give them up" (p. 5). "Quite quickly, once you get adept at looking for these demands, you will start to find them" (p. 9).

Not all therapists are as direct as Ellis and other cognitive behavior therapists in telling clients what their problems "really" are, but it is generally acknowledged that clients are influenced, often very subtly, to construe their problems in a fashion consistent with the preferred theoretical (and ethical) orientations of their therapists (e.g., Davison, 1976, 1978; Halleck, 1971; Rosenthal, 1955).

There is an additional problem with evidence from therapeutic interventions, namely, that the effectiveness of a therapy does not in itself tell us how the problem originated or is maintained (Davison, 1968; Rimland, 1964). If one has demonstrated that teaching a client to talk more "rationally" to himself alleviates a particular difficulty, one cannot logically conclude either that the problem developed from irrational self-talk or that the current problem is maintained by such self-talk, or even that there was any such self-talk in the first place. It is possible, for example, that neurotic anxiety is a result of a repression of infantile conflicts and that rational-emotive therapy is an effective way to lift this repression; or that neurotic anxiety arises from fortuitous classical conditioning, and that particular kinds of self-talk expose the person to the conditioned stimuli, thereby allowing the fear to extinguish. All one has demonstrated (and this would be no mean feat) is that one way to alleviate unhappiness is to have clients talk to themselves in a particular way.

OTHER EVIDENCE ON COGNITION INFLUENCING EMOTION

What techniques are available for exploring the relationship between certain types of thoughts and certain emotional responses to situations? In some studies (e.g., Velten, 1968; Rimm & Litvak, 1969; Russell & Brandsma, 1974), subjects are told to think or say to themselves state-

ments provided by the experimenter. Instructing the subject to say or think something emotional has been shown to affect physiological responses and mood ratings. While these findings support the general idea that what people are thinking affects how they currently feel, these studies are not directed at determining the thoughts people typically or spontaneously have under normal circumstances.

A widely used methodology to uncover typical thought patterns is to have subjects fill out one or more questionnaires and then to correlate their scores with, or look at differences on the scales as a function of, a psychiatric diagnosis. For example, Newmark, Frerking, Cook, and Newmark (1973) administered the Irrational Beliefs Test (Jones, 1968) to people diagnosed as neurotic, personality disorder, or normal, and found that many more of the neurotics than of the other two groups endorsed irrational beliefs. This could be seen as a confirmation of Ellis's principal assumption. Consider, however, the task requirements facing a subject who is taking the Jones Irrational Beliefs Test. To respond to an item tapping "It is essential that one be loved or approved by virtually everyone in his community," a subject has to reflect back over a host of situations and then draw a general conclusion about his or her characteristic ways of thinking. The experimenter assesses the subjects' abstracted or generalized views, not necessarily what they actually think in any particular type of situation.

There have also been attempts to distinguish certain subject-groups on the basis of their thoughts or self-statements in more carefully defined circumstances. Schwartz and Gottman (1976) presented written hypothetical situations to subjects and asked them to role-play assertive refusal responses. Low assertive subjects endorsed more negative and fewer positive statements about themselves than did high assertive subjects. Similarly, in a study of cardiac catheterization patients, Kendall, Williams, Pechacek, Graham, Shisslak, and Herzof (1979) found that poor adjustment ratings correlated with endorsements of negative self-statements on an inventory.

In a natural environment study by Vasta and Brockner (1979), subjects were asked to monitor all positive and negative self-evaluative statements for 3 hours per day for 8 days. Frequency of critical self-statements correlated negatively with self-esteem. Subjects high in self-esteem reported significantly more positive than negative self-statements, whereas subjects low in self-esteem reported positive and negative statements with equal frequencies.

Studies like those of Newmark et al., Kendall et al., Schwartz and Gottman, and Vasta and Brockner are useful, but they place severe restrictions on what thoughts subjects can furnish; subjects in these and related experiments are asked to endorse experimenter-defined categories

of thinking or even specific thoughts. Consequently, the experimenter might miss some thoughts by leaving important categories out of the questionnaire. More important, the experimenter might actually prompt the subject to report certain thoughts because providing categories or exemplars may be leading or suggestive. Given how little we know at this time about cognition and psychological disorder, it seems preferable to construct a paradigm that gives subjects greater leeway in the kind of data they will provide.

More open-ended responding has been used in a few studies. Sutton-Simon and Goldfried (1979) had acrophobics write down what their thoughts and feelings would be in a variety of heights situations. Fear of heights correlated with negative self-statements. Craighead, Kimball, and Rehak (1979) found that subjects high on the importance of social approval scale of the Irrational Beliefs Test made more negative self-statements while visualizing social rejection scenes than did subjects low on the scale. Goldfried and Sobocinski (1975) had earlier reported a similar finding.

Cacioppo, Glass, and Merluzzi (1979) told male subjects who were either high or low in social anxiety that they would be interacting with a female student in a conversation about college life. The experimenter then left and returned after 3 minutes to ask the subject to write down any thoughts he had had during the waiting period about himself, the situation, or the other person. Both the subject and the experimenter then rated the thoughts as positive, negative, or neutral. Subjects high on social anxiety generated more negative self-statements and rated the impending discussion and themselves more negatively than did less socially anxious subjects. This last paradigm has an advantage of being more realistic than the hypothetical situations described on questionnaires. However, it would be difficult, and prohibitively expensive, to arrange longer and more complex situations, especially those actually involving other people. In addition, any method in which there is a substantial interval between the thoughts or feelings and when they are reported runs the risk that important reactions will be forgotten or distorted during recall. Finally, many of these studies ask the subjects to write down their thoughts and feelings. Speaking is probably closer to thought and, in any event, generally less difficult to do than committing thought to paper. There is a good chance that subjects will attempt to express more complicated and/or subtle feelings and thoughts with verbal report than with written report, and that, in the same amount of time, more can be reported. (For a general review of theory and research on "thinking aloud," see Kendall & Hollon, 1981.)

It is only recently that an adequate methodology for the assessment of spontaneous thoughts in the natural environment has been developed. Hurlburt (1979, 1980) devised and Klinger (1978) independently used a

small, portable, random-interval tone generator that cues the subject to record in a small booklet his/her current thoughts, feelings, or behavior at intervals throughout the day. Klinger also had subjects rate content on a Thought Sampling Questionnaire containing a number of variables including duration, vividness, controllability, and degree of trust in one's own memory of the thought. The reported thoughts of Hurlburt's subjects were later analyzed for content by the experimenter (see also Hurlburt & Sippelle, 1978). The technique of in vivo thought-sampling seems to have enormous potential for providing on-the-spot assessment of a person's concerns. It should allow a wide range of responses and be free of the distortions that might be introduced by retrospective reports, or by asking the subject to describe his or her views in general terms or endorse statements provided by the experimenter. There are, however, drawbacks to this paradigm as well. It sacrifices control over, or direct knowledge about, what subjects are actually responding to. In addition, the interval sampled must necessarily be brief because subjects are going about their normal business while the sampling occurs. Most important, perhaps, it is likely that categories of cognition of great theoretical interest will be missed. For example, if one is interested in whether a given individual catastrophizes when criticized, the moments when this theoretically important event can be observed by random selection of time intervals are probably limited.

A PROPOSED EXPERIMENTAL PARADIGM

In summary, it seemed to us that a desirable paradigm for exploring cognition during complex events would have a number of features. It would permit open-ended verbal responding that reflects, as much as possible, ongoing thought processes rather than retrospective reporting; subjects should be constrained as little as possible in what they report, and they should be enabled to make their reports with ease and a minimum of delay. The experimenter should also be able to specify and manipulate the situations to which subjects are responding, while at the same time being able to present events that are sufficiently realistic and complex. Next, situations unlikely to be bothersome to subjects should be presented along with situations likely to produce anxiety; hitherto, direct comparisons have not been made. Finally, the procedure should not be prohibitively expensive in time or money. The present paper describes a paradigm that we believe comes close to meeting these requirements.

An audio recording of a conversation is created to simulate a complex event. Subjects are asked to pretend that the event is actually happening and that they are part of the situation. They are also told that

we are interested in the kinds of thoughts and feelings they are having as the situation occurs. A brief (15-25 seconds) segment of the tape is played, followed by a 30-second silence, during which subjects say what they are thinking and/or feeling. Another segment is played, followed by the subject's report, and so on. These verbalizations are tape-recorded for later analysis. Thus, the basic data obtained are articulated thoughts during simulated situations (ATSS, for short).

THE PRESENT STUDY

Reported here is an exploration of this technique. Of primary interest was whether subjects could "get into" the situations reasonably well and report their thoughts under our laboratory conditions. Additionally, we wanted to determine the interrater reliability that could be expected in analyzing the verbal reports obtained with this technique.

To assess the potential of the ATSS paradigm for yielding different responses depending on the situation, three tapes were constructed. Two of the tapes (Experimental) were designed to be stressful and directly relevant to the subject. In the third (Control), the conversation was not directly relevant to the subject. In one experimental tape, a teaching assistant was evaluating the subject's work, and in the other experimental tape, the subject was overhearing a conversation between two people about himself. Both the experimental tapes were largely critical of the subject. The control tape had the subject listening to a conversation about a professor whom he did not know.

These three situations were chosen because they varied in the amount of social-evaluative stress they should produce (Experimental vs. Control) and, within the experimental tapes, whether the criticism was directed primarily at the subject's academic performance or at his social skills. The tapes differed, of course, in many other respects as well, but these situations seemed to be a reasonable first sample of the types of events one might want to investigate people's reactions to. All three tapes were relevant to the population from which we were drawing subjects—undergraduates at a large university. All subjects received the control tape and then one of the two experimental tapes.

Subjects were also divided afterwards into groups depending on their scores on the Fear of Negative Evaluation Questionnaire (Watson & Friend, 1969). This afforded an opportunity to look at subject differences with the present paradigm. It seemed reasonable to expect that a person judged to be fearful of negative evaluation might respond differently to a situation in which he or she is criticized than would someone low on that

measure. Such a difference would not necessarily be expected in responses to a control, low-stress situation.

METHOD

Subjects

The subjects were 33 male undergraduates enrolled in an introductory psychology course at the State University of New York at Stony Brook. They were not preselected on any measures and participated in order to receive extra credit toward course requirements.

Materials

Three tapes were constructed. Each consisted of seven segments of 15-25 seconds' duration. Each segment was followed by a 30-second silence, during which subjects verbalized their thoughts. The tapes used were as follows.

Teaching Assistant (TA) Tape. A male speaker acted the part of a teaching assistant for an undergraduate course. The subject was asked to pretend he had come to see the TA in hopes of a grade increase on a paper or exam that the subject had submitted to the TA for regrading. Segments 2, 3, 5, and 7 contained statements designed to be critical evaluations of the student's work, attitude, or behavior, such as "Perhaps you did not study as hard as you should have." Segments 1 and 4 contained statements that were positive or complimentary about the student, such as "I very much liked your suggestions for future research in this area," and segment 6 was neutral.

Overheard Conversation (OC) Tape. There were two speakers, one male and one female. The subject was asked to pretend that he was over-hearing two acquaintances talking about him and the impression he had made at a recent social gathering. The acquaintances were in an adjoining room, unaware of the subject's presence. Segments 1, 2, 3, and 5 contained critical statements, such as "He just blurts out the first thing that comes to his mind" and "he looked really out of place." Segments 6 and 7 were more complimentary, with statements such as "He'll go out of his way to help you if he can, he's very generous," and segment 4 was somewhat supportive.

Control Tape. There were two speakers, one male and one female. The subject was asked to imagine that he was sitting in a cafeteria and

overhearing two people whom he did not know discuss a class he had never taken. Segments 1, 2, 3, and 5 contained statements that were critical of the class and/or the professor, such as “He doesn’t seem to care very much about what he’s teaching.” Segments 6 and 7 were more positive, e.g., “He is well respected internationally in his field.” Segment 4 was somewhat in defense of the teacher.

Procedure

Undergraduates were recruited for an experiment on “the kinds of things you think.” They were told that our interests were the thoughts and feelings people have in a variety of social situations. The male experimenter and the subject both sat at a table on which were placed two tape recorders. One of these was connected to a set of loudspeakers, and the other was equipped with a microphone. Subjects were instructed to listen to the tape of a simulated situation and to imagine as clearly as they could that they were a part of the event as it unfolded. They were also told to tune into their thoughts and feelings and, each time the tape stopped, to say these thoughts aloud. They were encouraged to be as frank and complete in their comments as possible. Subjects were aware that they were being recorded, speaking directly into a microphone during the 30-second silences.

All subjects were presented first with a description of the situational context for the Control tape and then with the Control tape itself. They were then presented with the appropriate situational context for one of the Experimental tapes. Seventeen subjects then received the TA tape, and 16 subjects the OC tape. Finally, all subjects completed the FNE questionnaire, which contains 30 true-false items describing reactions to situations of real or potential negative evaluation of others. Examples of these items are “I brood about the opinions my friends have about me” and “If I know someone is judging me, it has little effect on me.” Scores may range from a low of 0 to a maximum of 30. Watson and Friend (1969) have presented evidence for the acceptable reliability and validity of this scale. Subjects were divided at the median into High (FNE \geq 13) and Low (FNE $<$ 12) groups.³

³In ongoing studies an improved laboratory arrangement makes it possible for the experimenter and all taping equipment to be in a control room separate from the subject. The social interactions portrayed on our stimulus tapes are now played by actors from the university school of drama. Finally, subjects fill out questionnaires such as the FNE and articulate their thoughts in different sessions.

In summary, the study involved a mixed design, with one within-subject manipulated factor of Control versus Experimental tape, one between-subjects manipulated factor of TA versus OC experimental tape, and one between-subjects correlational factor of high versus low FNE score.

RESULTS

Dependent Variables

The data of interest are the verbalized responses of subjects to the taped situations. As one might expect, these responses were not always in the form of complete sentences. Also, a given sentence or phrase could contain a number of ideas. In order to enable us to score responses, all subjects' taped responses were first transcribed to written form. The task then was (a) to abstract from the flow of comments a sequence of discrete "idea-units" and (b) to categorize these units in some meaningful and useful way.

Idea Units. The principle for defining an idea unit was that it should be as small as possible while not distorting what was assumed to be the intention of the speaker. The following transcribed excerpt is illustrative; idea units are bracketed.

[Of course I realize what I'm saying.] [If I didn't realize what I was saying I wouldn't be saying it.] [It's their problem.] [Alright, maybe it makes me boring, talking] [but I'm only trying to express my point of view] [and it shows that I'm well-rounded] [and I'm able to socialize well.]

Clearly inference is necessary, and for that reason it is important to describe our assessment of the reliability of this effort. Two graduate students in psychology served as raters; one of them had also served as experimenter but, when performing this task and the categorizing to be elaborated below, was blind as to the score of a given subject on the FNE scale. The other also did not know any particular subject's FNE score and, further, had no connection with this project. As material for training, the raters compared unit divisions on a sample of 14 segments from pilot data. After achieving what appeared to be reasonable agreement on how to designate an idea unit, both raters independently divided the data from the 33 subjects into idea units. Recall that each subject had listened to two tapes: the Control tape and either the TA or the OC tape. Each tape was divided into seven segments; each of the segments was followed by a 30-second pause during which subject verbalized his reactions. Thus, each subject provided 14 verbalizations, each of which was composed of one or more idea units.

To assess reliability, 1 response segment of each subject was randomly selected from his total of 14. Reliability was then measured in three ways. First, the positions of unit end-boundaries were noted for each rater. The number of such boundaries on which both raters agreed was divided by the larger number of boundaries employed by either of the raters on that segment. This yielded a mean proportion agreement of .83 across the 33 segments ($SD = .16$, range = .60 to 1.00). Second, when the divisor in the above computation was replaced by the number of boundaries employed by Rater 1 (who later divided the complete data set into idea units), mean proportion agreement rose to .96 ($SD = .09$, range = .67 to 1.00). So, of the boundaries employed by Rater 1, Rater 2 employed 96% but also used additional boundaries that reduced the first reliability measure to 83%. Third, the Pearson product-moment correlation between raters for number of units per segment was .88.

We concluded that we had a reliable method for separating the flow of verbalizations into units. The work done by Rater 1 formed the basis for the subsequent coding.

Categorization Scheme. It is a truism that one always loses information when one categorizes. Obviously the code one employs should suit one's purpose. We wanted our scheme to be as exhaustive as possible in describing data that might be expected from our stimulus situations. Our principal theoretical interest guided the nature of our categories. In the most general sense, we were operating in Ellis's framework, which holds that maladaptive affect results from certain cognitions such as anticipating negative outcome, giving up, putting oneself down when criticized, or making demands on oneself or one's environment ("musturbation"). We supplemented what we had gleaned from the literature with our own intuitions and impressions formed in our clinical work and from listening to some of the subjects' tapes. We also tried to make the several categories overlap as little as possible. The result was a coding scheme of 25 categories, shown in Table I.

In the present study, there is a particular subset of these categories that is not applicable to responses to all of the taped situations. These are statements directed toward "the referent," which are likely only in the Control tape, in which the speakers on the tape are discussing a third party, a professor, and are not likely to be found in response to either of the Experimental tapes, in which the speakers are referring directly to the subject and never to a third party (i.e., the referent here is "Self").

It should also be noted that these categories are not mutually exclusive; it is possible for a given idea unit to be categorized in multiple ways. A fairly common example of this is that a "Critical Evaluation of Speaker" may be expressed as a "Should" ("He shouldn't have suggested I come if he didn't have enough time").

Table I. Categories of Articulated Thoughts

Critical evaluation of self:	A statement that is critical, pejorative, or negative about the self.
Critical evaluation of referent:	A statement that is critical, pejorative, or negative about the referent of the taped speakers' remarks when the object of the criticism is not the self.
Critical evaluation of speaker:	A statement that is critical, pejorative, or negative about the speaker(s) on the tape.
Critical evaluation of other:	A statement that is critical, pejorative, or negative about some third party other than the direct referent of the speaker's remarks.
Desire to harm:	A statement expressing the wish or intention to physically or psychologically harm the speaker.
Disagree with speaker:	A simple statement of disagreement with the speaker's view, involving no obvious evaluation of the speaker.
Positive evaluation of self:	A statement that is positive or complimentary about the self and that does not specifically serve to defend the self.
Positive evaluation of referent:	A statement that is positive or complimentary about the referent of the speaker's comments and that does not specifically serve to defend the third party.
Positive evaluation of speaker:	A statement that is positive or complimentary about the speaker on the tape and that does not specifically serve to defend the speaker.
Positive evaluation of other:	A statement that is positive or complimentary about a third party other than the direct referent of the speaker's remarks and that does not specifically serve to defend the third party.
Defense of self:	A statement that specifically serves to defend the self against statements of the speaker or previous statements of the self. These may be an explanation for behavior or simply an assertion of its correctness.
Defense of referent:	A statement that specifically serves to defend the referent against statements of the speaker or previous statements of the self. These may be an explanation for behavior or simply an assertion of its correctness.
Defense of speaker:	A statement that specifically serves to defend the speaker against statements of the speaker or previous statements of the self. This may be an explanation for behavior or simply an assertion of its correctness.
Defense of other:	A statement that specifically serves to defend some third party, other than the direct referent of the speaker, against statements of the speaker or previous statements of the self. These may be an explanation for behavior or simply an assertion of its correctness.
Agree with speaker:	A simple statement of agreement with the speaker's view, with no overt evaluation of the speaker or his behavior.
Empathy with speaker:	A statement that expresses understanding of the speaker's perceptions and feelings or that expresses a familiarity with the speaker's situation.
Positive anticipation:	A statement expressing anticipation of an event that would result in a positive outcome for the self.
Negative anticipation:	A statement expressing anticipation of an event that would result in a negative outcome for the self.
Resignation:	A statement that expresses perceived hopelessness or an intention not to attempt to have influence in a situation.
Problem solving:	A statement that specifies and/or evaluates a possible course of action or solution to a problem. These solutions or courses of action must be generated by the subject and not by the taped speaker(s).
Positive feeling:	A statement that expresses a positive affective state of the subject.
Negative feeling:	A statement that expresses a negative affective state of the subject.
Should:	A statement about a situation, individual, or set of individuals that implies a moral obligation to be or to behave in a particular way, to feel a given way, or to hold a given attitude.
Describe self:	A statement that describes typical or common behaviors, thoughts, attitudes, or feelings of the subject and that is not appropriately placed in any of the other categories.
Other:	Any statement that cannot be placed into one or more of the above categories.

As with the idea units, the reliability of the categories was evaluated. Each rater employed the coding scheme to score each of the aforementioned 33 sample segments, one from each of the 33 subjects. These segments had been divided by Rater 1 into 159 idea units.

Both raters employed an average of 1.12 categories per idea unit, with a range of 1 to 3. For any given unit, then, there could be full, partial, or no agreement on classification. There was at least partial agreement on 131 of the 159 units (82%) and total agreement on all classifications in 121 of the 159 (76%). Reliability was computed also by dividing the number of agreed classifications of a particular unit by the larger number of classifications used by either of the two raters for that unit. Across all units, the average of these scores was 78%. This figure was unchanged if the divisor in each case was replaced by the number of classifications used by Rater 1. Thus, of the classifications made by Rater 1, 78% were used also by Rater 2, and of all units, 76% were classified identically by the two raters. Of the units on which there was no agreement (18% of units), 76% were instances in which one rater used the category "other" where the other rater used some substantive category. There was thus total disagreement as a consequence of each rater using a different substantive category in only 4.32% (i.e., $24\% \times 18\%$) of all units, and these disagreements did not appear to be localized in any particular category. Thirty-six percent of all units were classified as "other" by Rater 1, and 35% by Rater 2. Following this assessment of reliability, Rater 1 went on to categorize the entire data set.

In summary, there was very high agreement between raters about how to divide the reported thoughts into idea units, and the raters agreed, for the most part, that each unit could be described adequately by a single, particular coding category. In addition, it is worth noting that about two-thirds of the idea units generated by subjects fell into substantive coding categories, suggesting that the coding scheme captured the majority of subjects' statements and that the tapes were successful in eliciting responses which, a priori, are of theoretical interest.

Overview of Statistical Analyses

All analyses were performed both on the number of idea units in each category and on the percentage of the total ideas generated by each subject that fell in each category. The results for each of these measures were largely the same, which could be anticipated because the total number of statements (summed across all categories) did not differ significantly as a function of FNE or stimulus tape. We shall report differences obtained with the percentage results, which are shown in Table II, noting those places in which the raw scores and percentage scores did not both exceed the .05 level of significance.

Two sets of analyses of variance were performed for each of the 25 categories of articulated thoughts. The first was a three-factor analysis of variance with a within-subjects factor of Condition (Control or Experimental) and two between-subjects factors: Group (TA or Overheard Conversation) and FNE (High or Low). The second set of analyses used the Control and Experimental conditions separately, each containing only between-subjects factors. A one-factor analysis of variance (FNE: High or Low) was performed on the Control condition data, and the Experimental condition data were analyzed by a 2×2 analysis of variance (Group: TA or Overheard Conversation, and FNE: High or Low). This second analysis served to clarify the location of any significant differences due to FNE found in the first analysis. Also, the Group factor in the first analysis compared those subjects receiving each of the Experimental tapes, but compared them across both Experimental and Control conditions combined. The Group factor in the second analysis compared the two Experimental tapes directly.

The results reported below for the effect of Condition (Control tape or Experimental tape) and its interactions with other variables are taken from the three-factor analysis of variance. The results for the effect of FNE are taken from the one-factor analysis on the Control data and from the two-factor analysis on the Experimental data, while those for the effect of Group and of $FNE \times Group$ are from the two-factor analysis. For clarity, the results will be presented grouped under the headings of Situation effects and Subject effects.

Situation Effects

Condition. This factor reflects whether or not the subject was personally a target of evaluation that was largely critical. It produced by far the greatest number of significant effects, many of which were also the largest effects.

The Experimental condition (when subject responded to either the TA or the OC stress tape) produced a significantly higher percentage of responses than did the Control condition in each of the following categories: Critical Evaluation of Speaker, $F(1, 29) = 26.283$, $MS_e = 236.987$; Desire to Harm, $F(1, 29) = 4.283$, $MS_e = 2.246$; Positive Evaluation of Self, $F(1, 29) = 28.629$, $MS_e = 8.321$; Defense of Self, $F(1, 29) = 62.669$, $MS_e = 99.393$; Negative Anticipation, $F(1, 29) = 7.075$, $MS_e = 4.164$; Positive Anticipation, $F(1, 29) = 4.620$, $MS_e = 3.251$; Resignation, $F(1, 29) = 8.384$, $MS_e = 11.878$; Positive Feeling, $F(1, 29) = 8.080$, $MS_e = 16.921$. The raw scores indicated more Should responses in the Experimental than in the Control condition, $F(1, 29) = 4.222$, $MS_e = 259.507$. The

Table II. Percent Group Means of Categories of Articulated Thoughts

	Control		TA		Overheard conversation	
	Low FNE	High FNE	Low FNE	High FNE	Low FNE	High FNE
Critical evaluation of self	.383	.000	.000	.506	3.059	3.860
Critical evaluation of referent	1.637	8.792	.000	.000	.000	.000
Critical evaluation of speaker	8.953	8.860	29.190	40.566	17.744	25.925
Critical evaluation of other	16.540	13.431	4.558	4.783	1.963	.765
Desire to harm	.000	.000	.000	.264	1.568	1.210
Disagree	.433	.592	.358	.773	.291	.228
Positive evaluation of self	.283	.000	1.488	2.432	5.867	5.998
Positive evaluation of referent	2.747	4.357	.000	.000	.000	.000
Positive evaluation of speaker	1.087	2.359	.245	3.202	.463	3.103
Positive evaluation of other	5.109	3.037	.794	.602	.463	.909
Defense of self	.357	.137	12.316	12.464	25.973	28.034
Defense of referent	5.851	5.229	.000	.000	.000	.000
Defense of speaker	.272	.551	1.935	1.173	.926	.410
Defense of other	1.098	.831	3.125	.411	.000	.000
Agree	1.394	2.056	.735	.748	.826	.765
Empathy	7.253	5.355	.000	.000	.000	.284
Negative anticipation	.000	.000	3.885	1.467	.000	.000
Positive anticipation	.000	.000	1.710	2.111	.000	.000
Resignation	.179	.878	8.939	2.358	.000	.686
Problem solving	8.182	11.235	9.168	7.146	8.679	4.404
Positive feeling	.128	.490	2.520	.359	4.810	5.136
Negative feeling	1.106	1.171	2.643	4.532	1.278	3.336
Should	3.969	3.518	5.638	5.672	3.644	4.881
Describe self	2.653	1.253	.915	.169	5.243	4.791
Other	34.021	29.787	15.968	16.110	23.236	12.935

higher incidence of critical evaluation of self and negative feeling in the Experimental condition was significant at the .10 level.

The Control condition produced a significantly higher percentage of responses than did the Experimental condition in each of the following Categories: Critical Evaluation of Referent, $F(1, 29) = 16.637$, $MS_e = 27.080$; Critical Evaluation of Other, $F(1, 29) = 22.010$, $MS_e = 196.574$; Positive Evaluation of Referent, $F(1, 29) = 10.788$, $MS_e = 18.729$; Positive Evaluation of Other, $F(1, 29) = 9.904$, $MS_e = 19.430$; Defense of Referent, $F(1, 29) = 11.805$, $MS_e = 42.373$; Empathy with Speaker, $F(1, 29) = 37.495$, $MS_e = 17.421$; and Other, $F(1, 29) = 18.746$, $MS_e = 194.966$.

Group. This factor reflects whether the Experimental tape involved evaluation of the subject in a social (the Overheard Conversation tape) or in an academic achievement context (the Teaching Assistant tape).

The TA tape produced a significantly greater percentage of responses than did the Overheard Conversation in the following categories: Negative Anticipation, $F(1, 29) = 7.075$, $MS_e = 8.327$; Positive Anticipation, $F(1, 29) = 4.620$, $MS_e = 6.501$; and Resignation, $F(1, 29) = 8.927$, $MS_e = 25.941$.

The Overheard Conversation produced a significantly greater percentage of responses than did the TA tape in the following categories: Critical Evaluation of Self, $F(1, 29) = 4.259$, $MS_e = 19.865$, raw scores $p < .08$; Positive Evaluation of Self, $F(1, 29) = 7.582$, $MS_e = 17.122$; and Defense of Self, $F(1, 29) = 8.883$, $MS_e = 197.868$. There were also more Descriptions of Self in the OC condition, $p < .10$.

Condition \times Group. A significant Condition \times Group interaction was found on Positive Evaluation of Self, $F(1, 29) = 6.731$, $MS_e = 8.321$, raw scores $p < .06$; Defense of Self, $F(1, 29) = 9.440$, $MS_e = 99.393$; Negative Anticipation, $F(1, 29) = 7.075$, $MS_e = 4.164$; Positive Anticipation, $F(1, 29) = 4.620$, $MS_e = 3.251$; and Resignation, $F(1, 29) = 10.043$, $MS_e = 11.868$. From the cell means, these interactions can be seen to reflect the following: Positive Evaluation of Self and Defense of Self occurred very little in response to the Control tape, to a greater degree to the TA tape, but most of all to the Overheard Conversation. Positive Anticipation, Negative Anticipation, and Resignation all occurred almost entirely to the TA tapes, with little or no thoughts of these kinds observed with the other two tapes.

Subject Effects

For the Control tape, High FNE subjects produced significantly more Critical Evaluation of Referent responses than did Low FNE subjects,

$F(1, 31) = 8.296$, $MS_e = 50.872$. These responses occurred only to the Control tape for all subjects, giving rise to a significant FNE \times Condition interaction, $F(1, 29) = 7.856$, $MS_e = 27.080$. Low FNE subjects made more Resignation responses than did High FNE subjects to the TA tape but not to either the Overheard Conversation or the Control tapes. This gave rise to a significant FNE \times Condition effect, $F(1, 29) = 4.644$, $MS_e = 11.878$, and to a significant effect of FNE \times Group, $F(1, 29) = 4.188$, $MS_e = 25.941$, raw scores $p < .07$. In addition, there was a tendency for the High FNE compared to the Low FNE subjects to report more Positive Evaluation of the Speaker in the Experimental Conditions, $p < .10$.

DISCUSSION

The results provide clear evidence that the cognitions reported by our subjects differed in several respects according to the nature of the situation. Both the Experimental and Control situations tended to elicit evaluative remarks, but they differed in the particular types of evaluative comments made. In reaction to the Experimental (TA or OC) tapes, where the subject is the evaluative target, there was more Critical Evaluation of the Speaker and Desire to Harm the Speaker, along with more Positive Evaluation of the Self and Defense of the Self. The Control situation, in which a third party was the evaluative target, elicited more of both positive and negative evaluative statements about the referent (the professor), as well as about others in general, rather than about the speaker or the self. Thus, while the speakers on all three tapes were criticizing *someone*, it was in the two Experimental tapes—where the subject himself was being judged—that we found higher frequencies of responses related to the subject and to the speaker; in contrast, in the Control tape—where the object of the speakers' evaluations is a third party, a professor—the subjects responded more often with evaluative comments about the professor. The factor of *personal involvement* in the respective situations seems to have influenced the way subjects were thinking.

The overall valence of the subjects' remarks also varied as a function of the target of the speakers' criticisms. When focus was on a third party (the professor on the Control tape), there was a relatively high number of both positive and negative evaluative thoughts, whereas the speaker who criticizes the subject (either TA or OC tape) is the recipient of almost entirely negative evaluative remarks. The subject himself was the recipient of almost entirely positive self-evaluations, perhaps as a defense against the criticism heaped on him.

The Experimental tapes differed from each other as well. This gave rise to interaction effects on several variables, in which a main effect of Condition was due to only one of the Experimental tapes differing from the Control tape. The TA tape, which would appear to be a situation in which an immediate decision is to be made affecting the subject (whether or not a higher grade will be awarded), elicited more of both Positive and Negative Anticipation and more Resignation than did the Overheard Conversation. The latter, in turn, elicited more self-referent statements, in the response classes of Critical Evaluation of Self, Positive Evaluation of Self, and Defense of Self. These findings make sense in that the TA tape oriented the listener to be concerned about an outcome about which he could be sanguine, pessimistic, or resigned, whereas the Overheard Conversation tape oriented him to be concerned more with how he evaluates himself as a person. Another way to look at these differences is that the TA tape set the occasion for goal-oriented anticipation (Am I going to get a better grade out of this discussion? There's no point in continuing this—his mind is made up), whereas the Overheard Conversation tape set the subject reflecting on whether he was as miserable a fellow as he was sometimes made out to be (I don't care what they think of the clothes I was wearing. I thought I looked pretty good that evening).

Our paradigm, then, appears to have succeeded in pulling statements from subjects about their thoughts and feelings that can be related to what they were exposed to. Our data, however, cannot tell us more analytically which of the many features of each tape subjects were reacting most strongly to. For example, one of the many factors not controlled for was importance of the discussion over the long haul. In the TA tape, the course for which the grade is at issue might have been construed by a given subject as critical in his admission to medical school, or it might have been a course for which the grade had little enduring importance. The TA in the tape might have been imagined to be a man with whom the subject anticipated further interactions because the course was part of the subject's major, or to be a man whom the subject was confident he could avoid. It has been demonstrated that subjects' thoughts—or at least what subjects are able or willing to verbalize in our experimental situation—can be reliably coded and do vary in an apparently sensible way with the situation. Future research can examine the effects on thinking of more systematic variations of parameters believed to be important in the generation of interpersonal stress.

The effects correlated with subject differences are more meager than the situational effects. This finding is in accord with the general conclusions from research on person-situation interactions (Magnusson & Endler, 1976) and might gladden the hearts of our situationalist colleagues. It may be that subjects—at least when drawn from nonclinical populations (cf. Wachtel,

1977)—can be understood adequately by knowing the situation they are in rather than “how much” of a trait they possess. Employing subjects drawn from clinical populations might reveal more significant differences; such research is currently under way.

The results we did obtain, however, lead to some interesting speculations. In the control condition, High FNE subjects were more critical of the referent than were Low FNE subjects, i.e., they were more inclined to endorse and elaborate the criticisms of the professor made by the speakers. In the Experimental condition, High FNE subjects reported more Positive Evaluation of the Speaker. These two findings, taken together, suggest that those subjects who are the most fearful of the evaluation of others are also those who are the most ready to evaluate others positively or negatively. They may have a generalized set for evaluation and see the world in evaluative terms more than do other people. The high incidence of Critical Evaluation of Speaker by High FNE subjects in the Experimental condition, though not significant, is congruent with this conjecture.

The range of subject effects, in any case, is not as great as we had expected. If there are many alternative explanations for a positive outcome in an experiment, the explanations for inconclusive results are just as numerous. It may be useful to mention a few. Perhaps the general tendency of a person to be fearful of a certain class of situations is *not* a function of what he tells himself in the situations as they unfold on a moment-to-moment basis; rather, the emotional reaction may occur for entirely different reasons (classical conditioning, for instance) but be *talked about* by the individual in self-statement terms when, on a later occasion, he is asked to do so (as is the case in clinical settings and in most of the available research). As we suggested earlier, there is an easily understood tendency for subjects to talk about their behavior in ways consistent with the experimenter’s or therapist’s explicit or implicit biases. Moreover, lay people would seem to seek causes of their behavior in what they are thinking more than they would explain their reactions in classical conditioning terms!

Nevertheless, it is quite possible that a larger number of real differences in self-statements as a function of FNE do exist and that these were not obtained because of limitations in the present study. One obvious limitation was the small *N* and the median split procedure, which hardly provide the most sensitive contrast between types of people. Also, in an effort to lend realism to the Experimental tapes, we constructed them to be less than totally negative. Thus, the tapes might not have been stressful enough to produce larger differences in responses from High and Low FNE subjects. Moreover, some other categorization of the ideas expressed might have revealed differences that our scheme missed. In addition, one could attempt paralinguistic analyses of the data, such as dysfluencies and

tone of voice. The use of videotape could also add to the analyses of verbal data; gestures and other body movements can either underscore or deny what a person says.

A fundamental question concerns the relationship of what people can and do report to what they actually think and feel. Cognitive-behavioral people like Ellis and Beck seem to assume that this relationship is simple and direct, as if thought *is* covert talking. This equation of cognition with self-talk is part of the legacy of mediational behaviorism, which can be traced back to metatheoretical progenitors like Mowrer, Dollard, and Miller (Mowrer, 1939; Dollard & Miller, 1950). The relationship, however, is undoubtedly much more complex. Some thoughts occur in a preverbal form and some in nonverbal modes such as imagery. The subject is faced with the task of translating these thoughts and feelings into words; this verbalizing process likely affects the content or form of the thoughts and feelings (Pope, 1977). Indeed, some ideational affect perhaps cannot even be approximated by language, or the language available to a particular subject. Of course, the likelihood that not all thoughts can be expressed does not mean that none can. Thus, our paradigm explores *articulated* thoughts, those that the subject can and/or will put in some verbal form. While this is certainly not all of cognition, it would seem to represent a substantial and clinically important part of it.

It should be made explicit that our paradigm cannot, in its present form, allow statements of cause and effect regarding thoughts and emotion. The most one can determine is whether there are *relationships* between what people think to themselves and the situations they find themselves in. The determination of such relationships, of course, is a necessary condition for addressing cause-effect issues; while correlation does not imply causation, the reverse certainly holds. An initial step, therefore, in studying whether people's cognitions affect how they feel and behave is to ascertain whether the two classes of phenomena covary in a manner that makes theoretical and common sense. The ATSS paradigm proposed here promises to be able to provide such data.

Implications and Future Directions

In the most general terms, our paradigm provides a direct and flexible way to assess the cognitions of different kinds of people under a wide variety of circumstances. On the assumption that the way people construe (talk to themselves about) the world is an important determinant of how they feel and behave—an assumption shared by all but the most radical of behaviorists—any method that permits access to this phenom-

enological space can be useful, not only in furthering our understanding of psychopathology but in tailoring treatment interventions to client problems.

A major advantage of using imagery and other symbolic presentations in the consulting room is the greater potential it provides in working with clients whose problems are not amenable to hands-on study and manipulation. Wolpe's (1958) systematic desensitization is a premier example of this, based as it is on Mary Cover Jones's pioneering clinical work in the 1920s (Jones, 1924) but altered via imagery to permit the therapist to deal with abstract problems like fear of criticism, as well as with concrete problems, like fear of elevators, without having to work with clients *in vivo*. In a similar way, our paradigm might enable the clinician or clinical researcher to create carefully designed representations of actual events in the consulting room or laboratory for the kind of direct, close-up study exemplified in the experiment reported here. Assuming that the kind of role-play inherent in our procedures does not do fatal violence to the external validity of the findings, a considerable advantage to employing the strategy described here is the enormous range of problems whose cognitive aspects one may want to tap into. Indeed, the limitations of the paradigm would appear to be only those imposed by investigator ingenuity and by the willingness and ability of a client to do the requisite role-playing, to introspect, and then report. Most clinicians, including behavior therapists, rely heavily on self-report already; there seems much to be gained by having that self-report more systematically collected and tied as closely as possible to those stressors the client's self-talk is assumed to be controlled by in the natural environment.

Finally, the paradigm holds promise as a way to study the active ingredients of therapy. An effective therapy, whether it be psychological or biological, may improve a person's functioning in any number of ways. It is conceivable, for example, that an intervention like rational-emotive therapy changes not what people tell themselves but how they feel about the very same self-statements. Thus, at least some variations of rational-emotive therapy might leave the self-talk intact, effecting a beneficial impact through an "unhooking" of the undesirable affect from the negative self-talk, an eventuality available for discovery via our paradigm. The effects of psychoactive drugs might also be investigated: Does a tranquilizer calm down an anxious person by changing the way he or she construes things, or by making him less upset about the way he construes things? Further, is drug-free maintenance of the therapeutic change influenced by one or the other of these two modes of action? Perhaps those drugs that an anxious person can be weaned from without losing therapy gains differ from chemicals that must continue to be taken in that they allow or encourage a change in self-talk, something that may

“belong to” a person more than does a drug-produced separation of cognition from affect. A fine-grain analysis of articulated thoughts might address such questions of possible importance for understanding the processes by which different therapeutic interventions exert their effects.

REFERENCES

- Bandura, A. Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 1977, *84*, 191-215.
- Beck, A. T. *Depression: Clinical, experimental and theoretical aspects*. Philadelphia: University of Pennsylvania Press, 1967.
- Cacioppo, J. T., Glass, C. R., & Merluzzi, T. V. Self-statements and self-evaluations: A cognitive-response analysis of heterosocial anxiety. *Cognitive Therapy and Research*, 1979, *3*, 249-262.
- Craighead, W. E., Kimball, W. H., & Rehak, P. J. Mood changes, physiological responses and self-statements during social rejection imagery. *Journal of Consulting and Clinical Psychology*, 1979, *47*, 385-396.
- Davison, G. C. Differential relaxation and cognitive restructuring in therapy with a “paranoid schizophrenic” or “paranoid state.” *Proceedings of the 74th Annual Convention of the American Psychological Association*. Washington, D.C.: American Psychological Association, 1966.
- Davison, G. C. Systematic desensitization as a counterconditioning process. *Journal of Abnormal Psychology*, 1968, *73*, 91-99.
- Davison, G. C. Homosexuality: The ethical challenge. *Journal of Consulting and Clinical Psychology*, 1976, *44*, 157-162.
- Davison, G. C. Not can but ought: The treatment of homosexuality. *Journal of Consulting and Clinical Psychology*, 1978, *46*, 170-172.
- Dollard, J., & Miller, N. E. *Personality and psychotherapy*. New York: McGraw-Hill, 1950.
- Ellis, A. *Reason and emotion in psychotherapy*. Secaucus, New Jersey: Lyle Stuart, 1962.
- Ellis, A. The basic clinical theory of rational-emotive therapy. In A. Ellis & R. Grieger (Eds.), *Handbook of rational-emotive therapy*. New York: Springer, 1977.
- Goldfried, M. R., & Davison, G. C. *Clinical behavior therapy*. New York: Holt, Rinehart & Winston, 1976.
- Goldfried, M. R., & Sobocinski, D. Effects of irrational beliefs on emotional arousal. *Journal of Consulting and Clinical Psychology*, 1975, *43*, 504-510.
- Halleck, S. L. *The politics of therapy*. New York: Science House, 1971.
- Hurlburt, R. T. Random sampling of cognitions and behavior. *Journal of Research in Personality*, 1979, *13*, 103-111.
- Hurlburt, R. T. Validation and correlation of thought sampling with retrospective measures. *Cognitive Therapy and Research*, 1980, *4*, 235-238.
- Hurlburt, R. T., & Sippelle, C. N. Random sampling of cognitions in alleviating anxiety attacks. *Cognitive Therapy and Research*, 1978, *2*, 165-170.
- Jones, M. C. A laboratory study of fear: The case of Peter. *Pedagogical Seminary*, 1924, *31*, 308-315.
- Jones, R. G. *A factored measure of Ellis's irrational belief system with personality and maladjustment correlates*. Unpublished doctoral dissertation, Texas Technological College, 1968.
- Kendall, P. C., & Hollon, S. D. Assessing self-referent speech: Methods in the measurement of self-statements. In P. C. Kendall & S. D. Hollon (Eds.), *Assessment strategies for cognitive-behavioral interventions*. New York: Academic Press, 1981.

- Kendall, P. C., Williams, L., Pechacek, T. F., Graham, L. E., Shisslak, C., & Herzof, N. Cognitive-behavioral and patient education interventions in cardiac catheterization procedures: The Palo Alto medical psychology project. *Journal of Consulting and Clinical Psychology*, 1979, 47, 48-59.
- Klinger, E. Modes of normal conscious flow. In K. S. Pope & J. L. Singer (Eds.), *The stream of consciousness: Scientific investigations into the flow of human experience*. New York: Plenum, 1978.
- London, P. *The modes and morals of psychotherapy*. New York: Holt, Rinehart & Winston, 1964.
- Magnusson, D., & Endler, N. S. (Eds.). *Personality at the crossroads: Current issues in interactional psychology*. Hillsdale, New Jersey: Erlbaum, 1976.
- Mahoney, M. J. *Cognition and behavior modification*. Cambridge: Ballinger, 1974.
- Meichenbaum, D. Cognitive modification of test-anxious college students. *Journal of Consulting and Clinical Psychology*, 1972, 39, 370-380.
- Meichenbaum, D. Self-instructional methods: In F. Kanfer & A. Goldstein (Eds.), *Helping people change*. New York: Pergamon, 1975.
- Mischel, W. *Personality and assessment*. New York: Wiley, 1968.
- Mowrer, O. H. A stimulus-response analysis of anxiety and its role as a reinforcing agent. *Psychological Review*, 1939, 46, 553-565.
- Newmark, C. S., Frerking, R. A., Cook, L., & Newmark, L. Endorsement of Ellis' irrational beliefs as a function of psychopathology. *Journal of Clinical Psychology*, 1973, 29, 300-302.
- Pope, K. S. *The stream of consciousness*. Unpublished doctoral dissertation, Yale University, 1977.
- Rimland, B. *Infantile autism*. New York: Appleton-Century-Crofts, 1964.
- Rimm, D., & Litvak, S. Self-verbalization and emotional arousal. *Journal of Abnormal Psychology*, 1969, 74, 181-187.
- Rosenthal, D. Changes in some moral values following psychotherapy. *Journal of Consulting Psychology*, 1955, 19, 431-436.
- Russell, P., & Brandsma, J. A theoretical and empirical integration of the rational-emotive and classical conditioning theories. *Journal of Consulting and Clinical Psychology*, 1974, 42, 389-397.
- Schwartz, R., & Gottman, J. Toward a task analysis of assertive behavior. *Journal of Consulting and Clinical Psychology*, 1976, 44, 910-920.
- Sutton-Simon, K., & Goldfried, M. R. Faulty thinking patterns in two types of anxiety. *Cognitive Therapy and Research*, 1979, 3, 193-203.
- Vasta, R., & Brockner, J. Self-esteem and self-evaluative covert statements. *Journal of Consulting and Clinical Psychology*, 1979, 47, 776-777.
- Velten, E. A. A laboratory task for induction of mood states. *Behaviour Research and Therapy*, 1968, 6, 473-482.
- Wachtel, P. *Psychoanalysis and behavior therapy: Toward an integration*. New York: Basic Books, 1977.
- Watson, D., & Friend, R. Measurement of social-evaluative anxiety. *Journal of Consulting and Clinical Psychology*, 1969, 33, 448-457.
- Wolpe, J. *Psychotherapy by reciprocal inhibition*. Stanford: Stanford University Press, 1958.