The Articulated Thoughts in Simulated Situations Paradigm: A Think-Aloud Approach to Cognitive Assessment

Gerald C. Davison, Sandra G. Navarre, and Ralph S. Vogel

What’s on your mind? What were you thinking when you encountered that situation you’ve been dreading? How did you arrive at that conclusion? These are but a few of the questions that cognitive and clinical psychologists have put to subjects and patients for many years and in many different contexts. The answers to such queries inform psychologists’ speculations about the role of thought in emotion and behavior.

In general, cognitive scientists have developed performance measures to infer cognitive processes. Experimental situations are created to make inferences about underlying cognitive processes. In contrast, clinical psychologists have tended to rely on clients’ introspective and retrospective self-reports in interviews or on questionnaires.

The core assumption in cognitive-behavioral theory, research, and practice is that affective and behavioral responses are mediated by thought processes, both conscious and unconscious. Operating within this paradigm, researchers have focused in recent years on the development of cognitive assessment methods, but with little consensus in approach to the measurement of cognition. As a result, vastly different modes of cognitive assessment have proliferated, without adequate attention to validity issues. Indeed, measures of the same constructs often do not correlate highly.

One mode of cognitive assessment, the think-aloud approach, is viewed as of particular value in accessing the products as well as the processes of cognition. Because think-aloud methods assess cognitions concurrently with their occurrence, they may be better suited to tapping actual thought content than other modes are. Standard think-aloud methods have subjects verbal-
ize cognitions while performing some task, and responses are recorded for subsequent evaluation. For example, early problem-solving research had subjects verbalize their thoughts while solving complex problems, and later research had subjects think aloud while they tried to perceive correctly an ambiguous picture that was slowly being brought into focus.

Despite the advantages of standard think-aloud methods, they have been criticized on several grounds: They are potentially reactive, because they impose dual demands of speaking and task performance, with the possible consequence that the assessment will have an undue influence on the behavior itself; subjects may report only a portion of cognitions experienced; and cognitions of high relevance to but low frequency in a particular dysfunction may be missed.

Our own efforts over the past 15 years have entailed the application of a paradigm dubbed Articulated Thoughts in Simulated Situations (ATSS). We believe that this think-aloud paradigm can facilitate valid cognitive assessment while benefiting from the advantages of think-aloud methodology and minimizing the potential problems.

THE ATSS PROCEDURE

The following exemplifies instructions given to subjects for the ATSS procedure:

In this study we are interested in the kinds of thoughts people have when they are in certain situations. Often, when people are going about their daily affairs, interacting with others and so forth, they have a kind of internal monologue going through their heads, a constant stream of thoughts or feelings that reflect their reactions to something that is happening.

What we’d like you to do is to play a part in a couple of situations we have taped. Your part will involve listening to situations and tuning in to what is running through your mind, and then saying these thoughts out loud. The tapes are divided into seven segments. At the end of each segment, there will be a tone, followed by a pause of 30 s, during which time we would like you to say out loud whatever is going through your mind. Say as much as you can until you hear another tone. Of course, there are no right or wrong answers, so please just say whatever comes to mind, without judging whether it seems appropriate or not. The more you can tell us, the better.

Try to imagine as clearly as you can that it is really you in the situation right now. Note that your task is not to speak back to any one of the voices on the tape, as though you were having a conversation with one of them. Rather, you should tune in to your own thoughts and say them out loud. The microphones in front of you will enable us to tape your comments.

Because thinking aloud in this manner is an unusual task, subjects are given a practice tape to orient them to the procedure, as well as to allow screening for compliance with the task (virtually no subjects are noncompliant). A typical experimenter prompt is "Rather than say what you would think if you were in the situation, try to believe that you actually are in the situation and then say whatever is going through your mind." After completion of the practice tape, further prompts are very seldom given to subjects. Each stimulus tape runs from 2 to 3 min; each segment is from 15 to 20 s in duration.

Among the many stimulus scenarios we have employed, the "Social Criticism" tape was designed to elicit cognitions related to social anxiety. This tape begins by saying that the subject is at a social function and overhearing two acquaintances talking about him or her; as the situation unfolds, these two acquaintances repeatedly berate the subject's behavior, manner of dress, and personality. The "Garage" tape was designed to elicit angry, hostile, or aggressive cognitions. This tape involves being made to wait for one's car, which was supposed to be ready, and overhearing somewhat fumbling and insensitive personnel discussing not only that it is not ready, but also that a needed part is not available and that repairs will be costly. The "Unknown Professor" tape has the subject overhear two strangers discussing, with both supportive and nonsupportive statements, a professor unknown to the subject. This tape represents a neutral, or control, condition.

ADVANTAGEOUS FEATURES OF ATSS

In comparison with other cognitive assessment approaches, the ATSS paradigm has a number of desirable features that information processing theory and research suggest can facilitate the accurate assessment of people's thoughts. These features concern the extent to which subjects are constrained in the data they provide, the degree to which subjects' reports are retrospective, the control of situations in which cognitions of interest are likely to occur, and the flexibility of the approach in studying a broad range of cognitions.

Unstructured Response Format

Unlike the most prevalent approach to cognitive assessment, paper-and-pencil endorsement methods, the ATSS method is relatively unstructured. That is, with endorsement methods, subjects select from limited options that alternative which most closely represents their thinking; in contrast, with ATSS, subjects are asked to report all cognitions (i.e., open-ended responding) rather than being limited to experimenter-selected options that may not be representative of their
actual cognitions. The use of unstructured responses increases the potential for obtaining a rich sample of actual cognitions of interest. Constraints on the data are imposed later via strategies to analyze the content, according to experimenter interest, and a virtually unlimited number of different coding schemes can be used on the same data set.

**On-Line Assessment**

ATSS utilizes a near-concurrent approach to cognitive assessment. Division of the ATSS tapes into short segments facilitates retention of small chunks of information in short-term memory during each 30-second think-aloud period. Thinking aloud that immediately follows each brief segment taps cognitions as close to “on-line” as possible, without interfering with the listening task (and thus minimizing reactivity). Reduction of the subject’s reliance on long-term memory and on retrospective reporting of generalized patterns of thinking over time minimizes censoring and distortions (e.g., self-presentation bias, confabulation), a common problem with paper-and-pencil measures.

**Situation Specificity**

ATSS audiotapes present specific, concrete stressors designed to elicit situationally relevant cognitions of interest. Thus, the ATSS paradigm provides for a high degree of experimental control over and situational specificity in the assessment of cognitions. The experimenter can associate cognitions of interest with specific and complex situations, make comparisons of cognitive categories across individuals or groups, and evoke cognitions in situations that are of theoretical concern but perhaps low frequency (avoiding the “hit or miss” problem of other think-aloud methods). A particular benefit is that situations that might be impractical or unethical to study in vivo can be simulated with ATSS.

**Flexibility of the Paradigm**

Assessment of cognitions is limited only by the experimenter’s creativity in designing provocative stimulus material and appropriate categories for analyzing content, and by the ability to involve subjects in ethically acceptable role-playing. Also, nonstressful ATSS tapes allow for comparisons between provocative and control situations, a feature lacking in most other cognitive assessment methods.

**EMPIRICAL FINDINGS**

The several experiments described in this section provide evidence for several forms of validity for the ATSS method. Face validity means that the items of an assessment measure are readily recognizable as representative of the construct being measured. For example, if one were measuring anger via observation of overt behavior, most people would agree that extreme anger is evident in a person threatening another person with bodily harm while exclaiming his hatred of the other. Concurrent validity for an assessment procedure is demonstrated when scores on it correlate with those of another assessment made at the same time and believed to be a valid measure of the same variable, as when a newly developed anxiety questionnaire and an established one taken by a group of subjects during the same testing session correlate highly with each other. If the new questionnaire were taken months before the established one, however, such a correlation would demonstrate the predictive validity of that new measure. A rigorous test of construct validity involves both concurrent and predictive strategies but is more interesting and informative because the correlation—often between two very different-looking measures, like a questionnaire and an observation of overt behavior—makes sense only within the framework of a hypothesis or theory.

The initial experiment with ATSS demonstrated its ability to elicit thoughts that could be coded reliably (a major concern with qualitative data like ours), were of theoretical interest, and could discriminate stressful versus neutral conditions. For example, when subjects listened to the Social Criticism tape, they articulated more thoughts critical of the person speaking, more desire to harm the speaker, more defense of self, more resignation, and more “should” responses than they did when thinking aloud to the Unknown Professor control tape. These findings support the construct validity of the method.

In an effort to study Ellis’s cognitive theory of emotion, the next experiment employed a less detailed coding system. We found a greater number of irrational beliefs in the stressful than in the neutral situations. Subjects also rated stressful situations as more anxiety-provoking than neutral situations; this finding suggested good face validity for ATSS. In addition, subjects who scored high on fear of negative evaluation (FNE) and a questionnaire measure of irrational beliefs rated the stressful ATSS situations as more anxiety-provoking than did low scorers on these trait measures, providing additional construct validation for ATSS. Self-reported anxiety correlated significantly with irrational thinking in both stressful and neutral ATSS situations, providing evidence of concurrent validity.

A group effect would, we then thought, more likely be found in a comparison of a clinical sample with a control group than in a comparison of unselected students scoring high or low on FNE. Our next exper-
iment, as well as studies conducted independently of our lab,' provided this comparison plus additional construct validation for ATSS. In our experiment, articulated thoughts of socially anxious subjects (referred from the campus counseling center) were found to be more irrational than those of control subjects. As in previous ATSS studies, both groups demonstrated more irrational beliefs in the stressful than in the neutral situations.

A concern with any think-aloud procedure is that subjects will not or cannot report everything that is on their minds. (Endorsement methods, of course, often do not permit subjects to report all that they might.) We therefore investigated subjects’ articulations when asked to think aloud twice to the same tape. Specifically, we looked at whether what subjects articulated depended on whether they were asked to say different things at the two times or just to repeat the procedure. It was found that these two demand instructions had less effect when articulated thoughts were coded functionally (viz., irrational thinking) than when they were coded in terms of surface similarity of statements at Time 1 and Time 2. This study emphasized the importance of the coding strategy selected and provided evidence that the fact that subjects do not verbalize everything on their minds probably has very little effect on functional (or, as some cognitive psychologists would say, schematic) analyses.

Do ATSS data relate meaningfully to overt behavior? This important question was addressed in two experiments. The first compared ATSS assessment of self-efficacy (i.e., confidence that one can perform well) with a behavioral assessment in speech-anxious subjects. We found that indices of anxiety coded from behavioral samples of speech-giving in an academic class correlated with self-efficacy coded in articulated thoughts as sampled in a separate lab situation weeks later. Specifically, self-efficacy in stressful (but not in supportive) ATSS situations correlated inversely with behaviorally rated speech anxiety and with questionnaire-based trait speech anxiety. This multimodal assessment of speech anxiety provided additional construct validation for ATSS, relating the method for the first time to actual behavior. In a subsequent study on the prediction of smoking relapse, articulated thoughts reflecting positive states of mind predicted longer periods of abstinence than did less positive states of mind. Such findings add to the predictive and construct validity of ATSS.

In the first ATSS study employing subjects with serious psychopathology, we found that depressed patients exhibited greater cognitive bias than nondepressed patients (who had other forms of psychopathology) in response to a simulated situation that depicted a negative occurrence (viz., that it rained on the day of a planned barbecue). This finding furnished both additional construct validation for ATSS and confirmation of an aspect of Aaron Beck’s cognitive theory of depression.

The ATSS method was also used to explore cognitive factors within the context of a large clinical trial on borderline hypertension and its amelioration. Type A individuals were found to articulate fewer self-supportive statements in a social-evaluative situation than Type B individuals. Findings from this part of the study also demonstrated that ATSS does correlate with analogous paper-and-pencil measures (though we have never argued that such correlations are necessary to establish the validity and utility of our paradigm). For example, ATSS measures of angry-hostile-aggressive (AHA) cognitions correlated positively with scores on the trait portion of the Spielberger State Trait Anger Scale (STAS-T). Also, ATSS measures of self-deprecating and self-supporting thoughts correlated with FNE positively and negatively, respectively.

The principal part of this clinical trial entailed a comparison of the standard nondrug intervention for hypertension (i.e., dietary and exercise regimens) with a treatment that also included intensive relaxation training. As previous behavioral medicine research had shown, heart rate (HR) and blood pressure (BP) were reduced more in patients who received relaxation therapy than in patients who did not. An important finding for our present purposes is that patients in the relaxation group had greater reductions in AHA cognitions, as measured by the ATSS, and these changes correlated positively with reductions in HR and BP; standard paper-and-pencil measures of anger and hostility (STAS-T and the Cook-Medley Ho Scale) did not pick up the changes. The greater sensitivity of ATSS to cognitive changes permitted the first demonstration of a link between reduction in angry thinking and diminution in blood pressure, a correlation widely assumed to exist.

ATSS has also been used to study children’s cognitions. In one study, in an independent lab, 8- to 11-year-old boys and their mothers articulated their thoughts to simulated situations that depicted parental conflict. Several findings shed light on the cognitive aspects of being in a high- or low-conflict home. For example, boys from low-conflict homes were more optimistic than boys whose families were physically or verbally aggressive. In a recently completed study in our lab, the cognitive aspects of Dodge’s theory of childhood aggression were examined in elementary-aged children. ATSS detected predicted differences between aggressive and nonaggressive boys in information processing deficits (hostile attribution biases and anticipated consequences for actions) in a tape in which the intentions of the provocateur were ambig-
uous but not when the aggressive intentions were obvious.

Ongoing ATSS research is examining the effects of induced mood on articulated thoughts, the test-retest reliability of ATSS measures of anxiety and AHA cognitions over a 1-month span, and cognitive factors in the coping styles of individuals who are caring for relatives with dementia.

CONCLUSIONS

The ATSS paradigm is a cognitive assessment approach that has been used to study cognitive factors in a number of clinically relevant contexts, including depression, hypertension, speech and social anxiety, smoking relapse, family conflict, and childhood aggression. ATSS is suitable with children, both younger and older adults, depressed patients, and populations with limited reading and writing skills. It may be especially useful when relatively little is known about the cognitive terrain one is interested in. ATSS was found in one of our studies to be superior to standard paper-and-pencil questionnaires in detecting cognitive changes from therapeutic intervention, suggesting its utility in process research. A variety of ATSS coding strategies have been reliably established, and the several studies described in this review furnish evidence for its face, concurrent, predictive, and construct validity.

ATSS is, though, far from a finished product; more research is needed regarding such issues as its psychometric properties, whether simulated situations evoke the same cognitive processes as real-life events, and whether and how it relates to other cognitive assessment techniques. Nonetheless, the paradigm is versatile and adaptable, limited only by ethical constraints and by the creativity of the experimenter interested in what is on people’s minds as they negotiate their way through life.

Acknowledgments—We wish to thank our many colleagues who have co-authored or otherwise contributed to the articles described in this manuscript. We appreciate also the input from two anonymous reviewers. Moreover, our labor-intensive research would not have been possible without the diligence and skills of the many coders and experimenters-most of them undergraduate students at the University of Southern California-who applied the several content-analytic schemes utilized in these efforts.

Notes

5. Davison, Robins, and Johnson, note 4.