

The Categorization of Thought Disorder

Howard Berenbaum^{1,2} and Deanna Barch¹

Accepted June 5, 1995

We describe the development of a new system for categorizing thought disorder. In the development phase (Study 1), we examined the degree to which speech samples and definitions of thought disorder subtypes taken from: (1) the Scale for the Assessment of Thought, Language, and Communication (TLC); (2) the Thought Disorder Index (TDI); and (3) the Assessment of Bizarre-Idiosyncratic Thinking (BIT), reflected disturbances in form versus disturbances in content. Ratings were provided by naive judges, experienced clinicians, and linguistic experts. The results contributed to the development of a new system dividing thought disorder into disturbances in (1) fluency, (2) discourse coherence, (3) content, and (4) social convention. In the validation phase (Study 2), 21 schizophrenic and 19 manic subjects were interviewed, interpreted proverbs, and responded to Rorschach cards. Subjects' speech was rated using the TLC, TDI, and BIT. We also measured hallucinations, delusions, and digit span performance. The results of Study 2 provided evidence supporting the validity of our new categorization system.

Disturbances in thought and speech have long been a central focus for researchers interested in understanding the etiology of schizophrenia. As

We wish to thank the following individuals for serving as our clinical and linguistic experts in Experiment 1: Sari Gilman Aronson, Gary Dell, Cynthia Fisher, Susan Garnsey, Georgia Green, William Kohen, Tom Lamberth, Pat McGuire, and Don Thomas. We are also grateful to Melanie Lotocky and Christine Krugler for providing us with the linguistic ratings of the speech samples described in Experiment 1. We wish to thank Suzanne Drungle, Bill Fiala, Bridget Hegeman, Susan Nordquist, and Neal Stolar for their help collecting the data in Experiment 2. Finally, we are very grateful for the kind cooperation of the staff, particularly John Gergen, and the patients of Covenant Medical Center in Urbana, Illinois.

¹ University of Illinois at Urbana-Champaign.

² Address all correspondence to Howard Berenbaum, Department of Psychology, University of Illinois, 603 East Daniel Street, Champaign, Illinois 61820.

pointed out by Maher (1972), "Psychopathologists have tended to regard the phenomena of schizophrenic language as reflections of a more basic disturbance of thought" (p. 3). Although a wealth of research has been conducted on thought disorder, confusion abounds concerning its definition (Andreasen, 1982). This difficulty defining thought disorder is not surprising when one considers that both of the following samples of schizophrenic speech have been presented as examples of thought disorder: (1) "After John Black has recovered in special neutral form of life the honest bring back to doctor's agents must take John Black out through making up design meaning straight neutral underworld shadow tunnel" (Chaika, 1990, p. 24); and (2) "When I first looked at it, it looked like a bat flying away, then I looked at it again, it looked like a bat coming toward me" (Solovay, Shenton, & Holzman, 1986, p. 494). Several psychopathologists (e.g., Andreasen, 1982; Cutting & Murphy, 1988) have pointed out that "thought disorder" is not a unitary construct, but rather encompasses several different components.

The question of how to conceptualize and measure thought disorder is analogous to the question of how to conceptualize and measure intelligence. Like thought disorder, intelligence is comprised of several different components. Just as the different facets of intelligence tend to covary, so do the different facets of thought disorder. Almost any reliable measure of intelligence can distinguish between individuals with profound developmental disabilities and control subjects. Similarly, almost any reliable measure of thought disorder can distinguish between schizophrenic and nonpsychiatric individuals. Even though different intellectual skills tend to covary, and almost all intelligence tests can distinguish between extreme groups, there can be little doubt that conceptualizing intelligence as being comprised of several different skills has enabled researchers to advance our understanding of these skills. For example, it is unlikely that the neural mechanisms underlying visuospatial skills and verbal comprehension would have become clear if researchers had never measured anything more specific than general intelligence. In our view, it is as important to distinguish between the different facets of thought disorder as it is to distinguish between the different facets of intelligence. We believe that a complete understanding of the varying phenomena labeled as thought disorder will probably never be achieved if we do not move beyond examining measures of global thought disorder.

Unfortunately, even if one agrees that thought disorder is an amalgamation of several different deficits, it is not entirely clear what the different components are or how they should be defined. An early distinction concerning the different facets of thought disorder was made by Schilder (1920/1951), who distinguished between disturbances in content and distur-

bances in form. Taylor (1981) defined disorder of content in a rather straightforward way, stating that "what the patient is talking about is thought content" (p. 48). A somewhat more elaborate definition was provided by Grebb and Cancro (1989), who wrote, "Disorders of content reflect ideas, beliefs, and interpretations of stimuli" (p. 761). Disordered content has sometimes been defined much more narrowly. For example, Fish (1964) limited his definition of disorders of content to delusions and over-valued ideas.

Disturbances in form have also been defined in a variety of different ways. According to Taylor (1981), "The form of speech is characterized by its rate, pressure, rhythm, idiosyncrasy of word usage, tightness of associational linkage and forms of associational linkage" (p. 48). According to Fish (1964), disorders in the form of thought refer to an inability to think abstractly. Examples of what Fish called formal thought disorder are "metonyms and personal idioms," "interpenetration of themes," and "overinclusion." Fish (1964) considered several clinical features such as flight of ideas and incoherence to be examples of disorders in the stream of thought. These disturbances are often considered problems in form by other psychopathologists.

Although psychopathologists have not agreed on how to define thought disorder, they have succeeded in developing instruments that describe specific subtypes of thought disorder. The three most commonly used instruments are the Scale for the Assessment of Thought, Language, and Communication (Andreasen, 1979a, 1986), the Thought Disorder Index (Johnston & Holzman, 1979; Solovay *et al.*, 1986), and the Assessment of Bizarre-Idiosyncratic Thinking (Marengo, Harrow, Lanin-Kettering, & Wilson, 1985). Throughout the remainder of the paper these three instruments are referred to using the acronyms TLC, TDI, and BIT, respectively.

These instruments have been quite valuable for researchers who wish to study thought disorder. However, their development has not solved the problem of different psychopathologists having different definitions of thought disorder and its components. What is missing from the field is some direction concerning how to conceptualize thought disorder at a level superordinate to specific subtypes, yet more specific than referring to "thought disorder." Distinguishing between different facets of thought disorder has the potential to improve our ability to develop testable models of the phenomena that comprise thought disorder. In addition, delineating the various components of thought disorder would enable researchers to compare studies that used different instruments to measure thought disorder (cf. Berenbaum, 1991; Shenton, Holzman, Solovay, & Coleman, 1991). Finally, understanding which dimensions of thought disorder are measured by the different

thought disorder subtypes would aid the interpretation of results from studies in which the thought disorder subtypes varied in their relationships with other variables (e.g., Solovay *et al.*, 1987).

The goal of this project was to explore approaches to conceptualizing and measuring the different facets of thought disorder. We wished to develop and begin validating a system of categorizing thought disorder that could help us understand how thought disorder is measured by different instruments. Many psychopathologists have described different dimensions of thought disorder (e.g., Andreasen, 1982; Cutting & Murphy, 1988), and some psychopathologists have even described categorization systems in which they indicated which specific categories were measured by which specific thought disorder subtypes. For example, Solovay *et al.* (1987) described several different ways of categorizing the thought disorder subtypes measured by the TDI. However, psychopathologists have yet to indicate which specific thought disorder subtypes, from each of several instruments reflect which specific categories of thought disorder. For example, a categorization system has not yet been developed that indicates whether a thought disorder subtype measured by one instrument (e.g., "idiosyncratic symbolism" as measured by the TDI) belongs to the same thought disorder category (e.g., disturbance in content of speech) as a thought disorder subtype measured by a different instrument (e.g., "derailment" as measured by the TLC). We believe that such a categorization system would make it easier to compare the results of research projects that employ different thought disorder rating instruments, and would also be valuable for the development and testing of new theoretical models of thought disorder.

STUDY 1: DEVELOPMENT PHASE

The goal of Study 1 was to explore the following issues: (1) Are the dimensions of form and content of speech sufficient to describe all of the different thought disorder subtypes described in the TLC, TDI, and BIT? and (2) If the dimensions of form and content are not sufficient, how else might we conceptualize the variety of subtypes that comprise thought disorder? Ultimately, the results of Study 1 were used to guide our development of a new thought disorder categorization system that was then evaluated in Study 2.

Method

Judges

There were three groups of judges who rated the degree to which speech samples and definitions of thought disorder subtypes reflected form

versus content: (1) a "naive" group composed of 16 undergraduates who received a small stipend for participating in the research project; (2) a "linguist" group composed of three psycholinguists and one linguist who were all university professors in either a Department of Psychology or a Department of Linguistics; and (3) a "clinician" group composed of four psychologists and one psychiatrist. All five clinicians had extensive experience with psychotic individuals. Their number of years of post-Ph.D./M.D. clinical experience ranged from 3 to 19, with a mean of 10.8 ($SD = 6.0$). In addition to the three groups of judges described above, two graduate students in psycholinguistics conducted linguistic analyses of the speech samples. The nature of the speech samples and the linguistic analyses are described below. None of the judges or raters were familiar with the TLC, TDI, or BIT, and none were aware of the goals of this project.

Materials

We wanted the judges to rate definitions and speech samples representing each of the thought disorder subtypes. To accomplish this, we used definitions and speech samples from the instruction manuals for the TDI (Solovay *et al.*, 1986), the TLC (Andreasen, 1986), and the BIT (Marengo *et al.*, 1985). A list of the thought disorder subtypes taken from each of the instruments is provided in Appendix A (Table A1). Because the goal of this project was to explore disturbances in what people say, we did not examine those thought disorder subtypes that measure other related disturbances. We did not examine the BIT "behavior" category because it is rated on the basis of the individual's behavior rather than on the basis of what is said. We did not examine the TLC categories of "poverty of speech" and "pressure of speech" because they are rated primarily on the basis of how much is said rather than on the basis of what is said.

Nineteen different TLC categories were examined. Small alterations of the original TLC categories were used in this study following the modifications reported by Berenbaum, Oltmanns, and Gottesman (1985). The altered TLC categories were "tangentiality," "derailment," and "loss of goal." In addition, two new categories were defined as "nonsequitur response" and "verbigeration." None of the other changes described by Berenbaum *et al.* (1985), such as dropping infrequently used categories, were made in this study. Ten different categories from the BIT were used. The TDI manual describes 23 different main categories. Six of the main categories include between two and five subcategories. For four of these six main categories we used only a single definition and speech sample. Because we thought the subcategories of "peculiar verbalizations" ("peculiar expression," "stilted inappropriate expression," and "idiosyncratic word us-

age") as well as the subcategories of "queer responses" ("queer expression," "queer imagery," and "queer word usage") might differ in the degree to which they measured form versus content of speech, we used separate definitions and speech samples for each of these subcategories. Thus, we examined 27 different TDI categories. The total number of categories examined from all scales was 56.

The speech samples taken from the three different instruments were selected so that they would be similar in terms of their lengths and the proportion of samples that included both a question and a response. A one-way analysis of variance indicated that the speech samples from the three instruments did not differ significantly in length [$F(2,53) = 2.34, p > .1$]. A chi-square test indicated that the speech samples from the different instruments did not differ significantly in the proportion of samples containing questions and responses [$\chi^2(2, N = 56) = .04, p > .9$].

Typewritten copies of the speech samples and subtype definitions were mounted on index cards. Each speech sample and definition was mounted on a separate card. Thus, there were two sets of 56 cards, one set for the speech samples and one set for the definitions. The subtype names were not placed on any of the cards, which were identified by random code numbers.

Procedure

The first three sets of judges were asked to rate each definition and speech sample on the degree to which it reflected a disturbance in form versus content of speech. The ratings were made using a 5-point scale anchored by 1 = "definitely a disturbance in form" and 5 = "definitely a disturbance in content." This rating scale reflected our initial hypothesis that the different subtypes reflected either disturbances in form or disturbances in content. We treated this as an ordinal scale rather than as a categorical scale because the closer two scores were to each other the more they were assumed to resemble each other; for example, a thought disorder subtype that was assigned a score of "5" was assumed to be more similar to a subtype assigned a score of "4" than to a subtype assigned a score of "1." All judges rated all of the speech samples before rating the definitions. The index cards were shuffled so that the different judges rated the cards in different random orders.

The naive judges and the clinicians were asked to make their ratings using the definitions of disturbances in form and content shown at the top of Table 1. These definitions were guided by those provided by Grebb and Cancro (1989) and Taylor (1981). The linguists were asked to make their

¹ Our hypothesis that there would be differences among the different subcategories was supported by the data, as can be seen in Appendix A (Table A1).

Table I. Definitions of Form and Content of Speech

Investigators' definitions

Form

Form of speech refers to how a person is talking. The form of speech is evaluated by examining the kinds of words people use (are they real or made up) and whether the connections between words, phrases, and sentences are meaningful and understandable.

Content

Content of speech refers to what a person is talking about. When what a person is saying seems peculiar, illogical, not consistent with reality, or of idiosyncratic significance, this indicates a disturbance in the content of speech.

Linguists' definitions

Form

- Linguist 1: Ungrammaticality, neologisms, sound-related slips of the tongue
 Linguist 2: Deviant in phonology (mispronounced or rhyming sound), lexical status (new words) and syntax (ungrammaticality)
 Linguist 3: Word finding or encoding difficulties, syntax (ungrammaticality)
 Linguist 4: Ungrammaticality, sound-related word substitutions, items involving some kind of repetition or the intrusions of related words

Content

- Linguist 1: Violations of pragmatic maxims (having too much or too little detail, using idiosyncratic terminology, jumping from one topic to another without relating them), disturbance of ideas
 Linguist 2: Deviant semantics (illogical) or pragmatics (inappropriate or untrue)
 Linguist 3: Something wrong with what is said, not the way it was said
 Linguist 4: Anything not rated as form or no problem

ratings using their own definitions of disturbances in form and content of speech. They were also asked to provide us with brief descriptions of their definitions, which are presented at the bottom of Table I.

Three of the four linguists identified a number of thought disorder definitions and speech samples which they felt did not indicate any kind of disturbance. As a result, the ratings from the linguist group provided us with the following two sets of scores for each definition and for each speech sample: (1) the mean form versus content rating of those judges who considered the speech sample or definition to be problematic and (2) whether or not the speech sample or definition was considered to be nonproblematic by one or more members of the linguist group.

The fourth set of judges, the psycholinguistics graduate students, was asked to indicate which speech samples contained one or more linguistic errors. The categories and criteria for linguistic errors were developed by

the judges and authors prior to rating the speech samples. The categories were designed to include common linguistic errors such as violations of grammar, unclear references, speech errors, and made-up words.

Interrater Reliability

The mean of each set of judges' ratings was used in the data analyses presented below. Therefore, interrater reliability of the form versus content ratings was measured using intraclass correlations (Shrout & Fleiss, 1979), with judges treated as random effects, and the mean of the judges' ratings treated as the unit of reliability. The intraclass correlations across the three sets of judges ranged from .74 to .80 for the speech samples and from .79 to .84 for the definitions.

In rating which speech samples contained a linguistic error, the two psycholinguistics graduate students agreed with each other 96% of the time. Interrater reliability measured using kappa was .89. Disagreements between these two judges were resolved by consensus.

Results and Discussion

We began by examining the relationships between the first three sets of judges' scores. As can be seen in Table II, the correlations between the different sets of judges' ratings were moderately positive, and all were statistically significant. The mean scores assigned by each set of judges to each thought disorder subtype are presented in Appendix A (Table A1). Appendix A also indicates which thought disorder subtypes were judged to be non-disturbed by at least one of the linguists.

As one would guess based on the correlations presented in Table II, and as can be seen in Appendix A, the scores assigned by the three groups of judges displayed a modest degree of similarity (Table A1). However, only a subset of the subtypes that tended to be considered disturbances in form

Table II. Correlations Between Ratings Made By Different Groups of Judges

	Naive	Clinicians	Linguists
Naive	—	.51*	.69*
Clinicians	.71*	—	.45*
Linguists	.71*	.51*	—

Note. The scores used to compute these correlations were the average rating made by each set of judges for each thought disorder subtype. Correlations above the diagonal are for speech samples and correlations below the diagonal are for definitions.

* $p < .01$, one-tailed.

by the clinicians and naive judges were also judged that way by the linguist group. Thus, it appears that the linguists defined form of speech more narrowly than we had defined it for the clinicians and naive judges.

We originally expected two dimensions of thought disorder, disturbed form and content, to be sufficient to describe the wide assortment of thought disorder subtypes. We hoped the data we collected would allow us to say which thought disorder subtypes reflected disturbances in form and which reflected disturbances in content. However, as can be seen from the results presented in Appendix A (Table AI), the thought disorder subtypes could not be dichotomized as neatly as we had expected. As a result, we developed a new categorization system composed of the following four categories: (1) disturbances in fluency, (2) disturbances in discourse coherence, (3) disturbances in content, and (4) disturbances in social convention. The subtypes from the TLC, TDI, and BIT that we feel measure each of the four categories of disturbance are presented in Table III. The rationale used for assigning subtypes to the four thought disorder categories is presented below. There were several thought disorder subtypes that we did not feel confident placing in any of the four aforementioned categories; these subtypes are listed in the "not categorized" column in Table III. The mean form versus content ratings of each group of judges for each category of disturbance are presented in Table IV. Also presented in Table IV is the percentage of thought disorder subtypes in each category that were judged to be nondisturbed by at least one of the linguists.

We placed in the fluency category those subtypes which we believe reflect a disturbance in the ability to produce independent, grammatical speech with understandable words. These thought disorder subtypes are generally considered to be severe, and occur rather infrequently. As can be seen in Appendix A (Table AI) and Table IV, the thought disorder subtypes placed in the fluency category tended to be considered disturbances in form by all three groups of judges. This group of thought disorder subtypes was also the least likely to be considered non-disturbed by the linguists.

We placed in the discourse coherence category those subtypes which we believe reflect a disturbance in the sequential flow of information from the speaker to the listener. Compared to the subtypes placed in the fluency category, the subtypes in the discourse coherence category are generally considered to be less severe, and occur more frequently. These thought disorder subtypes tended to be considered disturbances in content by the clinicians and naive judges, and to be considered disturbances in content by the linguist group. This group of thought disorder subtypes was more likely to be considered nondisturbed by linguists than were the subtypes in the fluency category but less likely to be considered nondisturbed than were the subtypes in the social convention and content categories.

Table III. Proposed Thought Disorder Categories

Scale	Category				Not categorized
	Fluency	Discourse coherence	Social convention	Content	
TLC	Paraphasia-phonemic Paraphasia-semantic Neologisms Word approximations Verbigeration Echolalia Incoherence	Nonsequitur response Tangential response Derailment Distractable speech Loss of goal Clanging	Poverty of content Circumstantiality Stilted speech	Self-reference Perseveration Illogicality	
TDI	Word finding difficulties Neologisms Idiosyncratic word usage	Looseness Clangs	Flippant response Vagueness Stilted inappropriate expression	Fluidity Perseveration Confusion Idiosyncratic symbolism Inappropriate distance Familiar combinations Playful confabulation Incongruous combinations	Relationship verbalization Fragmentation Peculiar expression Queer expression Queer word misuse Absurd response Incoherence
BIT	Strange verbalizations Confused or disorganized ideas	Lack of relationship between the subject's statement and the question asked Lack of shared communication	Overelaborated response	Contamination Contamination Queer imagery Autistic logic Coherent but odd ideas Peculiar or idiosyncratic reasoning or logic Deviant with respect to social convention Intermingling Attending to part rather than whole	

Table IV. Form Versus Content and Nondisturbed Ratings of Speech Samples and Definitions

	Category			
	Fluency	Discourse coherence	Social convention	Content
Naïve				
Definition rating (M)	2.4	2.6	2.8	3.6
Speech sample rating (M)	2.4	3.1	2.5	3.6
Clinician				
Definition rating (M)	2.1	1.7	3.1	3.8
Speech sample rating (M)	2.1	2.4	3.0	3.5
Linguist				
Definition rating (M)	2.1	3.8	3.8	4.5
Speech sample rating (M)	2.1	4.0	2.4	4.1
Judged nondisturbed (%) ^a	17	40	71	75

Note: Rating scale: 1 = definitely form, 5 = definitely content.

^aPercentage of thought disorder subtypes for which at least one linguist judged either the speech sample or the definition to be nondisturbed.

We placed in the content category those subtypes which we believe reflect disturbances in what the individual is talking about, rather than a disturbance in how the idea is being presented. These thought disorder subtypes were those that were most consistently considered disturbances in content by all three groups of judges. The majority of subtypes in the content category was considered nondisturbed by at least one of the linguists.

We placed in the social convention category those subtypes which we believe reflect a violation of social conventions concerning economy and style in conveying information. The form versus content scores of subtypes in the social convention category tended to fall between those in the discourse coherence and those in the content categories. The majority of subtypes in the social convention category was considered nondisturbed by at least one of the linguists.

Having reorganized the thought disorder subtypes from the TLC, TDI, and BIT into the four categories described above, we proceeded to examine whether the speech samples from the four categories would differ in the linguistic ratings they received. The majority (67%) of the speech samples in the fluency category contained linguistic errors, compared to 20% of the discourse coherence, 14% of the social convention, and 5% of the content samples. A chi-square test indicated that the categories differed significantly in the proportion of speech samples containing linguistic errors [$\chi^2(3, N = 49) = 16.15, p < .005$]. Fisher's exact tests showed that the fluency category

contained more linguistic errors than any of the other categories ($p < .05$), none of which differed significantly from each other.

The results of Study 1 indicated that the dimensions of form and content of speech are not sufficient to describe all of the different thought disorder subtypes contained in the TLC, TDI, and BIT. Consequently, prompted by the data we collected, along with our own sense of how and why the different thought disorder subtypes resembled one another, we developed a new system of categorizing thought disorder. In our new system, thought disorder is divided into disturbances in (1) fluency, (2) discourse coherence, (3) social convention, and (4) content. Finally, linguistic analyses of speech samples representing the different thought disorder subtypes provided some preliminary validation for the system.

STUDY 2: VALIDATION PHASE

Although the data we collected in Study 1 helped guide our development of the new thought disorder categorization system, we readily admit that it was also influenced by our own views of language and thought disorder. We consider this quite reasonable considering that Study 1 was conducted in the context of discovery. Consequently, there was clearly a need to determine whether our thought disorder categorization system would be supported by a new, independent set of data based on speech samples collected from psychiatric patients. The goal of Study 2 was to examine the validity of the thought disorder categorization system developed in Study 1. In particular, we wished to examine whether (1) the relationships among the different thought disorder subtypes varied as a function of the thought disorder categories to which they had been assigned and (2) whether the different thought disorder categories differed in their associations with a variety of psychopathological phenomena. If pairs of thought disorder subtypes assigned to the same category were more strongly associated with each other than were pairs of subtypes assigned to different categories, it would suggest that the new categorization system is succeeding in making meaningful distinctions among the thought disorder subtypes. Similarly, if the different thought disorder categories were differentially associated with other phenomena, it would provide evidence for the potential usefulness of the new categorization system.

Method

Subjects

Subjects were 19 individuals with bipolar disorder (all of whom were hospitalized because of manic episodes) and 21 individuals with schizo-

phrenia. Psychiatric diagnoses were made using DSM-III-R criteria following a structured clinical interview and a review of clinical records. Additional sociodemographic and clinical data, along with information concerning medication status, is provided in Table V. Chi-square tests indicated that significantly more schizophrenic than manic subjects were psychotic [$\chi^2(2, N = 40) = 7.98, p < .065$] and that significantly more manic than schizophrenic subjects had symptoms that are typically associated with mania [$\chi^2(2, N = 40) = 8.98, p < .005$]. However, psychotic symptoms were not uncommon in the manic subjects, and symptoms associated with mania such as racing thoughts were not uncommon in the schizophrenic subjects. All subjects were hospitalized at the time of their participation in the study.

Procedure

Clinical Assessment. Each subject was interviewed by one or the other investigator (HB or DB) using the mood and psychotic disorders sections of the Structured Clinical Interview for DSM-III-R (SCID; Spitzer, Williams,

Gibbon, & First, 1992). Both interviewers have had extensive experience using a variety of different structured clinical interviews and clinical rating scales. The interviewers also used the Brief Psychiatric Rating Scale (BPRS; Lukoff, Nuechterlein, & Ventura, 1986) to rate each subject's clinical state. For the purpose of the present study, we focused on two symptoms, hallucinations and delusions (which are measured by the "Hallucinatory Behavior" and "Unusual Thought Content" scales of the BPRS, respectively).⁴ We examined hallucinations because it has been proposed that disturbances in language production are associated with auditory hallucinations (Hoffman, 1986).⁵ We examined delusions because of the possibility that the disturbed thinking frequently presumed to underlie delusions is related to thought disorder as conceptualized and measured by the TLC, TDI, and BIT. The clinical assessment and accompanying ratings were completed prior to the speech elicitation and measurement of attentional processes described below.

Speech Elicitation and Measurement of Attentional Processes. To rate thought disorder using the Scales for the Assessment of Thought, Language and Communication (TLC; Andreasen, 1979a, 1986), subjects were interviewed using a semistructured interview that contained open-ended questions regarding interests and daily activities. To rate thought disorder using the Assessment of Bizarre-Idiosyncratic Thinking (BIT; Marengo *et al.*, 1985), subjects were administered the Gorham Proverbs Test (Gorham, 1956). To rate thought disorder using the Thought Disorder Index (TDI; Johnston & Holzman, 1979; Solovay *et al.*, 1986), subjects were administered part of the Rorschach test. Following the procedure used by Hurt, Holzman, and Davis (1983), subjects were administered one achromatic card, one chromatic card, and one red-and-black card. Subjects were audiotaped throughout the speech elicitation portion of the experiment.

To measure short term memory and selective attention, subjects were administered the Digit Span Distraction Task (Oltmanns & Neale, 1975). In this test, subjects perform a digit span task under two conditions: (1) while ignoring digits presented by a male voice in between the target digits presented in a female voice (distraction condition) and (2) in the absence of any distraction (nondistraction condition). Scores for performance in the distraction and nondistraction conditions were the percentages of correctly remembered digits in each condition. We examined attentional processes because at least some forms of thought disorder have been found to be associated with disturbances in attentional processes (e.g., Harvey, Earle-

Table V. Sociodemographic Characteristics, Clinical Data, and Medication Status

Variable	Schizophrenic	Manic
Age		
M	35.3	32.6
SD	11.6	14.7
Education		
M	13.4	13.1
SD	3.6	2.0
Race		
% white	95	68
Gender		
% male	67	53
Symptoms (1)		
Psychotic	95	58
Elevated	52	95
Medications taken (%)		
Neuroleptics	100	80
Lithium	38	68
Anticonvulsants	19	32
Antiparkinsonian	48	42

(1) Psychotic symptoms were judged to be present if the subject received a rating of at least three (definitely present) on either of the following BPRS items: (1) hallucinations or (2) unusual thought content.

(2) Symptoms of elevation were judged to be present if the subject received a rating of at least three (definitely present) on any of the following BPRS items: (1) grandiosity, (2) tension, or (3) excitement.

⁴ Data concerning hallucinations and delusions were unavailable for one of the manic subjects.

⁵ Of the 16 subjects in this study judged to have hallucinations, only 1 did not have auditory hallucinations.

Boyer, & Levinson, 1988). The speech elicitation and attentional measurement portions of the experiment were carried out by members of the research team who were blind to subjects' diagnostic and clinical status.

Thought Disorder Ratings. All thought disorder ratings were made using transcriptions of the audiotaped speech samples. TLC, BIT, and TDI ratings were made by three, three, and four undergraduate research assistants, respectively. Each of the three instruments was rated by a different set of raters. The research assistants making the thought disorder ratings were blind to the clinical and attentional status of the subjects whose speech they were rating. The raters were trained by one of the authors (D.B.) using transcripts that were not part of this study. To prevent rater drift, the raters reviewed practice transcripts with DB on a weekly basis. Some thought disorder subtypes were rated as not present (Appendix B, Table AII). The interrater reliabilities of the remaining subtypes,⁶ measured using intraclass correlations with judges treated as random effects and the mean of the raters treated as the unit of reliability, are presented in Table VI. Those subtypes which were rated infrequently, such as Idiosyncratic Word Usage, tended to have low intraclass correlations; however, the raters usually agreed on whether these subtypes were present. For example, the average pairwise percentage agreement was 77% for Idiosyncratic Word Usage and 95% for derailment. Raters assigned categorical scores to the TLC poverty of speech and pressure of speech thought disorder subtypes. For the remaining thought disorder subtypes, each rater rated the number of times each subject exhibited each thought disorder subtype. Thought disorder subtype scores were computed by averaging the scores given by each of the raters. To control for verbal productivity, TDI scores were corrected for the total number of Rorschach responses following the procedure used by Solovay *et al.* (1987). Similarly, TLC scores were corrected following the procedures used by Berenbaum *et al.* (1985).

Results and Discussion

We began by examining whether thought disorder subtypes that had been assigned to the same categories were more highly correlated with each other than were thought disorder subtypes that had been assigned to different

Table VI. Interrater Reliabilities of Thought Disorder Ratings

Subtype	Instrument	Category	Reliability
Neologisms	TLC	FLU	.77
Word approximations	TLC	FLU	.35
Nonsequiturs	TLC	DC	.78
Tangential responses	TLC	DC	.78
Derailment	TLC	DC	.49
Loss of goal	TLC	DC	.63
Poverty of speech	TLC	N/A	.88
Pressure of speech	TLC	N/A	.53
Word finding difficulties	TDI	FLU	1.0
Idiosyncratic word usage	TDI	FLU	.09
Looseness	TDI	DC	.50
Flippant response	TDI	SC	.95
Vagueness	TDI	SC	.50
Stilted inappropriate expression	TDI	SC	.38
Perseveration	TDI	CON	.67
Incongruous combinations	TDI	CON	.87
Idiosyncratic symbolism	TDI	CON	.67
Inappropriate distance	TDI	CON	.61
Confusion	TDI	CON	.83
Confabulation	TDI	CON	.45
Relationship verbalization	TDI	N/A	.89
Absurd response	TDI	N/A	.87
Peculiar expression	TDI	N/A	.39
Strange verbalizations	BIT	FLU	.83
Confused and disorganized ideas	BIT	FLU	.78
Lack of relationship between the subject's statement and the question asked	BIT	DC	.81
Lack of shared communication	BIT	DC	.89
Artificial language	BIT	SC	.79
Overelaboration	BIT	SC	.80
Irrelevant wandering	BIT	SC	.50
Coherent but odd ideas	BIT	CON	.87
Peculiar reasoning or logic	BIT	CON	.60
Intermingling	BIT	CON	.73
Attending to part rather than whole	BIT	CON	.92

Note. FLU, fluency; DC, discourse coherence; SC, social convention; CON, content; N/A, not applicable (not considered to be a disturbance in fluency, discourse coherence, social convention, or content)

⁶ The original "strange verbalization" category from the BIT contained four subcategories, one of which was "artificial language." The definition of "artificial language" provided by the BIT was almost identical to the definition of "stilted inappropriate expression" provided by the TDI, and of "stilted speech" provided by the TLC. Therefore, we examined "artificial speech" as a separate subtype and placed it in the Social Convention category. The "strange verbalization" subtype used in the Fluency category included the remaining three subtypes.

categories.⁷ For example, the correlation between confabulation and inappropriate distance was expected to be higher than the correlation between

⁷ Due to the limited sample size, we were unable to address this issue using more powerful multivariate techniques such as confirmatory factor analysis.

confabulation and vagueness because both confabulation and inappropriate distance were assigned to the content category, whereas vagueness was assigned to the social convention category. We conducted a two-way analysis of variance (ANOVA) using instrument combination (same instrument versus different instruments) and category combination (same category versus different categories) as independent variables and Fisher-transformed correlations between pairs of thought disorder subtypes as the dependent variable. Thus, this ANOVA was conducted using 378 "cases," with each "case" being the Fisher transformed correlation (across all subjects) between a pair of thought disorder subtypes. This ANOVA revealed a significant main effect for category combination [$F(1,374) = 4.48, p < .05$].⁸ As can be seen in Table VII, thought disorder subtypes that had been assigned to the same categories were more strongly correlated with each other than were thought disorder subtypes that had been assigned to different categories. Neither the main effect for instrument nor the interaction between category and instrument was statistically significant.

We then conducted an additional analysis to verify that the statistically significant effect for category combination obtained in the ANOVA described above was not merely an artifact of the data analytic strategy or the large number of degrees of freedom. For each subject, we computed two scores: (1) the average absolute difference between all thought disorder subtype z scores that belonged to the same category and (2) the average absolute difference between all thought disorder subtype z scores that belonged to different categories. We used z scores (which were computed using the means and standard deviations of the total sample) rather than the raw scores because the different instruments each used different rating scales. Our rationale for conducting this analysis was that if, as we hypothesized, thought disorder subtypes that belonged to the same category resembled each other more than did thought disorder subtypes that belonged to different categories, the mean absolute within-category difference would be smaller than the mean absolute between-category difference. This hypothesis was tested us-

⁸ The results did not change when the original correlations were used instead of the Fisher transformed z values.

Table VII. Average Correlations Between Pairs of Thought Disorder Subtypes

Instrument	Category	
	Same	Different
Same	23	14
Different	16	11

ing a paired-samples t test in which each subject was a "case." Consistent with the results of the ANOVA described above, there was a statistically significant effect for category combination [$t(39) = 1.83, p < .05$], with within-category differences being smaller than between-category differences.

Next we examined whether the different thought disorder categories were all similarly associated with hallucinations, delusions, and digit span performance. To examine this issue we began by computing scores for each of the four categories. Category scores were computed by summing the z scores of the thought disorder subtypes that had been assigned to each category. However, we included only those subtypes which either (1) had correlations greater than .30 with the sum of the remaining subtypes in the category or (2) were more strongly correlated with the sum of the remaining subtypes in the category than they were, on average, with the scale scores corresponding to the other categories. Almost all subtypes were retained using these criteria; the only subtypes not retained were neologisms (TLC), loss of goal (TLC), overelaborated response (BIT), and confusion (TDI). For the purpose of comparison with the category scores, we also computed total instrument scores for the TLC, TDI, and BIT by summing the individual thought disorder subtype z scores. The internal consistencies of the category and instrument scores, calculated using Cronbach's alpha, were as follows: fluency, .61; discourse coherence, .72; social convention, .47; content, .71; TLC, .38; TDI, .53; and BIT, .86.

Based on previous research (e.g., Harvey *et al.*, 1988), we expected higher levels of thought disorder to be associated with the presence of hallucinations and delusions, and with poorer performance on the attentional task. As can be seen in Table VIII, the patterns of correlations with delusions

Table VIII. Correlations Between Thought Disorder Category and Instrument Scores and Hallucinations, Delusions, and Digit Span Scores

	Hallucinations	Delusions	Distraction digit span	Non distraction digit span
Fluency	-.03	.47**	-.31*	-.42**
Discourse coherence	.03	.26	-.31*	-.31*
Social convention	-.23	-.02	.32	.21
Content	.01	.03	-.10	-.21
TDI total	-.14	.08	-.02	-.19
BIT total	-.07	.21	-.15	-.23
TLC total	.10	.17	-.36*	-.28*

* $p < .05$, one-tailed.

** $p < .01$, one-tailed.

and digit span scores varied across the four thought disorder categories. Poor digit span scores were significantly associated with greater fluency and discourse coherence disturbance.⁹ The severity of delusions was significantly associated with greater degrees of disturbance in fluency, and there was a trend in the same direction ($p < .06$) for disturbances in discourse coherence. In contrast, disturbances in content and social convention were not significantly associated with either digit span scores or delusions. The TLC, TDI, and BIT total scores were generally less strongly associated with delusions and digit span scores than were disturbances in fluency and discourse coherence.¹⁰

Next we tested the differences between correlations reported in Table VIII using the procedure described by Meng, Rosenthal, and Rubin (1992). The correlation between delusions and disturbances in fluency differed significantly from the correlations between delusions and disturbances in social convention ($z = 2.25, p < .05$, two-tailed) and disturbances in content ($z = 2.42, p < .01$, two-tailed). The correlation between nondistracted digit span scores and disturbances in social convention differed significantly from the correlations between nondistracted digit span scores and disturbances in fluency ($z = 2.82, p < .01$, two-tailed) and disturbances in discourse coherence ($z = 2.34, p < .01$, two-tailed). Similarly, the correlation between distraction digit span scores and disturbances in social convention differed significantly from the correlations between distraction digit span scores and disturbances in fluency ($z = 2.85, p < .01$, two-tailed) and disturbances in discourse coherence ($z = 2.88, p < .01$, two-tailed). Although the correlation between distraction digit span scores and disturbances in content was smaller than the correlations between distraction digit span and disturbances in both discourse coherence and fluency, neither difference was statistically significant using two-tailed tests.

For the most part, the patterns of correlations described above were similar for schizophrenic and manic subjects.¹¹ However, there were two differences that were statistically significant. Among schizophrenic subjects, greater disturbances in fluency were associated with higher hallucination ratings ($r = .37, p < .05$, one-tailed), whereas among manic subjects the

correlation was in the opposite direction ($r = -.41$); the difference between these correlations was statistically significant ($z = 2.37, p < .01$, two-tailed). Thus, among schizophrenic subjects, both hallucinations and delusions were associated with disturbances in fluency ($r = .37$ and $r = .62$, respectively) whereas among manic subjects, delusions but not hallucinations were associated with disturbances in fluency ($r = .68$ and $r = -.41$, respectively).

The second significant difference between schizophrenic and manic subjects concerns the correlation between distraction condition digit span scores and disturbances in content. Among schizophrenic subjects, more severe disturbances in content were associated with worse distraction condition digit span scores ($r = -.55, p < .01$, one-tailed), whereas among manic subjects the correlation was in the opposite direction ($r = .14$); the difference between these correlations was statistically significant ($z = 2.2, p < .05$, two-tailed). Disturbances in content were also more strongly associated with nondistracted digit span scores among schizophrenic subjects ($r = -.52, p < .01$, one-tailed) than among manic subjects ($r = -.03$) though this difference was not statistically significant.

The results of Study 2 indicated that the thought disorder subtypes that had been assigned to the same categories were more highly correlated with each other than were the thought disorder subtypes that had been assigned to different categories. The results also indicated that the different thought disorder categories were differentially associated with delusions and attentional processes. Thus, the results of Study 2 provide evidence supporting the validity of the thought disorder categorization system that had been developed in Study 1.

GENERAL DISCUSSION

The goal of this project was to explore alternative approaches to conceptualizing and measuring thought disorder. The results of Study 1 indicated that the thought disorder subtypes measured by the TLC, TDI, and BIT could not be neatly dichotomized into disturbances in either form or content. Consequently, we developed a new system for categorizing thought disorder subtypes. We proposed conceptualizing thought disorder as being comprised of disturbances in (1) fluency, (2) discourse coherence, (3) content, and (4) social convention. The results of Study 2 provided evidence supporting the validity of our new categorization system.

Earlier in this paper, we argued that psychopathologists will not be able to develop a thorough understanding of the different phenomena labeled as thought disorder if they do not move beyond examining measures of global thought disorder. This argument is supported by the results of this project.

⁹ None of the thought disorder category or instrument scores were significantly associated with increased distractibility as measured by the distractibility index from the Digit Span Distraction task (nondistracted performance - distraction performance).

¹⁰ The results were almost identical when using the TDI weighted total score suggested by Solovay *et al.* (1986) and the BIT total score suggested by Marengo *et al.* (1985).

¹¹ With the exception of the TDI total scores, on which the manic subjects had significantly higher scores than the schizophrenic subjects [$t(38) = 2.21, p < .05$, two-tailed], the two diagnostic groups did not differ significantly on any of the thought disorder category or instrument scores.

In Study 2 we found that the different thought disorder categories were differentially associated with several other aspects of psychopathology. For example, disturbances in fluency, but not in social convention, were associated with the presence of delusions and attentional impairment. Thus, distinguishing among the different facets of thought disorder provided critical information about the relationships among specific symptoms. Had only global thought disorder measures been used, such as the total scores from the TLC, TDI, and BIT, these relationships would have remained hidden.

The results of this study also provide information relevant to understanding the etiology of the various facets of thought disorder. Disturbances in fluency and discourse coherence, but not in social convention or content, were related to attentional deficits among both schizophrenic and manic patients. Previous research (e.g., Harvey, Docherty, Serper, & Rasmussen, 1990; Harvey *et al.*, 1988; Pandurangi, Sax, Pelonero, & Goldberg, 1994; Strauss, Buchanan, & Hale, 1993) has found that attentional processes and thought disorder are related. This study furthers our understanding of this relationship by clarifying which aspects of thought disorder are related to attentional deficits. Again, had only global measures of thought disorder been used, information about such specific relationships would not have been clear. By delineating the various facets of thought disorder, we can now make explicit and testable predictions about the associations between particular aspects of thought disorder and potential etiological mechanisms such as attentional deficits.

In our view, psychopathologists must develop testable models and specific predictions, such as those we are about to describe, to make significant progress in understanding thought disorder. We believe that the thought disorder categorization system we have developed will contribute to this effort. We hypothesize that the proximal causes of disturbances in fluency and discourse coherence are disturbances in the language production system. The language production system carries out functions such as discourse planning, monitoring, editing, and grammatical and phonological encoding and relies on attentional resources in order to perform many of these functions (Levelt, 1989). It is our prediction that the proximal causes of disturbances in content and social convention do not reside in the language production system but, instead, reflect problems such as intrinsic thinking disturbances and impaired perspective (e.g., Harrow, Lanin-Kettering, & Miller, 1989). To account for the covariation between the different facets of thought disorder, we hypothesize that they all share more distal antecedents. For example, Cohen and Servan-Schreiber (1992) have suggested that a disturbance in memory for context, stemming from a frontal lobe dopaminergic dysfunction, is related to thought disorder in schizophrenia. A disturbance in contextual memory could disrupt language production by

preventing an individual from using context to guide discourse planning and/or editing, thereby causing disturbances in fluency and discourse coherence. It is conceivable that the same contextual memory disturbance could lead to disturbances in content and social convention, not via the language production system, but through other means, such as by limiting the individual's ability to determine what a speaker in a particular context should do.

The results of this project suggest several avenues for future research. It will be important for future studies to examine more closely which specific thought disorder subtypes belong to which categories of thought disorder. It will also be important to determine whether some thought disorder subtypes are more representative of certain categories than are other thought disorder subtypes. For example, are derailments more indicative of a disturbance in discourse coherence than are tangential responses? If the different subtypes vary in the degree to which they reflect different thought disorder categories, it may be useful for future research to focus on those subtypes that are most prototypic of the different categories. A related issue is whether some thought disorder subtypes reflect more than a single type of disturbance. For example, might loss of goal be indicative of both a disturbance in discourse coherence and a disturbance in social convention? To address these issues, future studies will have to have much larger sample sizes than we had in Study 2 of this project. Another issue that needs to be addressed in future research concerns the categorization of the thought disorder subtypes that we did not feel sufficiently confident to categorize. Which, if any, of those thought disorder subtypes represent disturbances in fluency, discourse coherence, social convention, or content? The most important direction for future research is to explore the antecedents and correlates of the different categories of disturbance. Such research will need to determine which antecedents and correlates are common to all facets of thought disorder, and which are associated with only a subset of thought disorder categories. In conducting such research, it will be important to explore potential differences between individuals with different psychiatric disturbances. For example, we found that among schizophrenic but not among manic subjects, hallucinations were associated with disturbances in fluency, and attentional deficits were associated with disturbances in content. Finding such group differences may reveal different pathways that can lead from disturbances in basic mechanisms, such as attentional processes, to overt signs of psychopathology, such as hallucinations and disturbed speech.

Although we believe that the system we have proposed for categorizing thought disorder subtypes is a potentially useful one, we recognize that it is just one of many possible categorization schemes. Other psychopathologists have proposed alternative ways to categorize thought disorder subtypes (e.g.,

Andreasen, 1979b; Solovay *et al.*, 1987), though none have described systems that can be used with multiple thought disorder measurement instruments. We do not wish to argue that our system is the only one that may prove to be useful. Just as there is more than one meaningful way of categorizing the different skills that comprise intelligence, there is undoubtedly more than a single meaningful way of categorizing thought disorder. We expect our system to be particularly valuable for psychopathologists who wish to take a psycholinguistic based approach to understanding thought disorder. Ultimately, the usefulness of our system will be judged on its ability to guide theory construction and to predict and explain empirical findings.

Table A1. Mean Form Versus Content Rating for Speech Samples and Definitions

Item	Speech samples				Definitions							
	Naive		Clinicians		Linguists		Naive		Clinicians		Linguists	
TLC												
Nonsequitur	3.7	1.4	4.5*	2.7	1.2	4.8						
Paraphasia-phonemic	2.5	2.2	1.0	2.1	2.4	1.0						
Distractions speech	2.3	1.6	4.5	2.9	2.4	4.5						
Tangentiality	3.3	1.8	4.5*	2.3	2.6	4.0						
Derailment	3.9	3.0	4.0*	3.2	1.4	4.3						
Stilted speech	2.8	3.2	2.0*	1.9	3.4	3.5						
Self-reference	3.9	3.4	5.0*	3.4	4.2	4.3						
Paraphasia-semantic	2.7	3.2	1.8	2.7	3.0	1.8						
Incoherence	1.9	1.4	1.0	1.8	1.0	1.5						
Illogicality	4.3	3.4	4.8	3.5	3.6	4.8						
Circumstantiality	2.1	2.6	1.3*	3.1	2.8	4.3						
Poverty of content	2.4	2.6	3.0*	3.5	4.2	3.0						
Neologisms	1.6	1.2	1.8	1.8	1.0	1.5						
Loss of goal	2.6	3.0	3.3*	2.4	2.4	4.3						
Clanging	3.7	3.4	4.5	2.0	1.2	2.0*						
Verberation	2.8	2.0	3.3	2.7	2.2	1.8						
Perseveration	3.1	3.2	3.3	2.9	2.8	3.8						
Echolalia	2.6	2.4	3.3	2.4	1.6	2.8						
Word approximations	1.7	1.4	1.3	2.6	2.6	2.0						
BIT												
Coherent but odd ideas	3.1	2.8	4.0	3.7	4.4	5.0						
Deviant with respect to social convention	3.5	4.4	5.0*	3.5	4.0	4.8						
Peculiar or idiosyncratic reasoning or logic	3.5	3.0	3.3	3.5	3.8	4.1						
Lack of relationship between the subject's statement and the question asked												
Intermingling	3.3	1.8	4.0	2.7	1.4	4.5						
Attending to part rather than whole	4.1	3.0	4.0*	3.7	3.8	4.1						
Confused or disorganized ideas	3.9	4.4	5.0*	3.7	4.0	4.0						
Overelaborated response	3.2	1.0	2.5	3.4	2.6	4.5						
Lack of shared communication	2.8	2.4	4.0*	3.2	2.4	4.5						
Strange verbalizations	2.1	1.0	1.3	2.0	1.4	3.8						
	1.6	1.6	2.7	1.9	2.6	2.0						

Table A1 continued. Mean Form Versus Content Rating for Speech Samples and Definitions

Item	Speech samples			Definitions		
	Naive	Chicretians	Linguists	Naive	Chicretians	Linguists
11)						
Stilted inappropriate expression	2.8	3.2	3.0*	1.9	2.8	2.3*
Idiosyncratic word usage	2.7	3.6	2.0*	2.7	2.0	2.3*
Fluidity	3.4	3.8	3.0*	3.9	3.0	5.0*
Perseveration	3.5	3.6	4.0*	3.1	4.2	4.8
Incongruous combinations	3.4	4.6	5.0*	3.7	4.4	5.0
Relationship verbalization	2.3	3.8	3.5*	3.1	3.4	3.5
Idiosyncratic symbolism	3.3	3.6	4.0*	3.5	3.6	4.3*
Inappropriate distance	3.5	3.8	4.3	3.5	3.6	4.5
Flippant response	2.4	3.0	5.0*	3.0	3.4	5.0
Fabulized combinations	3.8	4.0	5.0*	4.1	4.0	4.5
Playful confabulation	3.5	3.4	4.5*	3.0	3.6	4.5
Fragmentation	2.7	3.0	2.0*	2.9	2.8	3.3
Vagueness	2.3	3.8	2.3	3.2	2.8	3.7*
Peculiar expression	3.5	3.2	2.1*	2.3	1.8	4.0
Queer expression	2.1	2.2	2.8	3.3	3.6	3.3*
Queer word misusage	4.5	5.0	4.7*	2.0	2.4	1.7*
Confusion	3.1	1.6	4.0*	3.5	3.6	4.3*
Foreshness	2.7	2.6	4.3	3.3	2.0	4.5
Queer imagery	4.3	3.4	4.3*	3.5	3.2	4.7*
Absurd response	2.5	4.0	3.0*	3.5	3.2	4.5
Confabulation	3.7	4.0	4.5*	3.9	4.2	4.5
Autistic logic	3.1	3.6	5.0*	3.9	4.0	4.8
Contamination	3.8	3.0	4.7*	3.9	3.6	4.7*
Neologisms	1.9	1.0	2.3	1.7	1.0	2.3
Incoherence	3.6	2.0	1.8	3.0	2.2	3.7*
Word finding difficulties	2.9	3.6	3.5*	2.5	2.6	1.8
Clangs	3.0	4.8	4.0*	3.0	1.4	1.5*

* Considered nonproblematic by one or more linguists

APPENDIX B

Table AII. Thought Disorder Subtypes Rated as Not Present in Experiment 2

TLC	Distractible speech
	Stilted speech
	Self-reference
	Incoherence
	Illogicality
	Circumstantiality
	Poverty of content
	Clanging
	Verbigeration
	Perseveration
	Echolalia
BIT	Deviant with respect to social convention
TDI	Fluidity
	Neologisms
	Clangs
	Playful confabulation
	Queer imagery
	Autistic logic
	Contamination

REFERENCES

- Andreasen, N. C. (1979a). Thought, language, and communication disorders. I. Clinical assessment, definition of terms, and evaluation of their reliability. *Archives of General Psychiatry*, *36*, 1315-1321.
- Andreasen, N. C. (1979b). Thought, language, and communication disorders. II. Diagnostic significance. *Archives of General Psychiatry*, *36*, 1325-1330.
- Andreasen, N. C. (1982). Should the term 'thought disorder' be revised? *Comprehensive Psychiatry*, *23*, 291-299.
- Andreasen, N. C. (1986). The scale for assessment of thought, language and communication (TLC). *Schizophrenia Bulletin*, *12*, 473-482.
- Berenbaum, H. (1991). Distinguishing between content and form of speech. *Archives of General Psychiatry*, *48*, 280.
- Berenbaum, H., & Oltmanns, T. F. (1983). *Instruction manual for the assessment of formal thought disorder*. Unpublished manuscript.
- Berenbaum, H., Oltmanns, T. F., & Gottesman, I. I. (1985). Formal thought disorder in schizophrenics and their twins. *Journal of Abnormal Psychology*, *94*, 3-16.
- Chaika, E. (1990). *Understanding psychotic speech: Beyond Freud and Chomsky*. Springfield, IL: Charles C. Thomas.
- Cohen, J. D., & Servan-Schreiber, D. (1992). A parallel distributed processing approach to biology and behavior in schizophrenia. *Psychological Review*, *99*, 45-77.
- Cutting, J., & Murphy, D. (1988). Schizophrenic thought disorder: A psychological and organic interpretation. *British Journal of Psychiatry*, *152*, 310-319.

- Fish, F. (1964). *An outline of psychiatry for students and practitioners*. Bristol: John Wright & Sons.
- Gorham, D. R. (1956). A proverbs test for clinical and experimental use. *Psychological Reports, 1*, 1-12.
- Grebb, J. A., & Cancro, R. (1989). Schizophrenia: Clinical features. In H. I. Kaplan & B. J. Sadock (Eds.), *Comprehensive textbook of psychiatry*, Vol. 1, 5th ed. Baltimore, MD: Williams and Wilkins.
- Harvey, P. D., Earle-Boyer, E. A., & Levinson, J. C. (1988). Cognitive deficits and thought disorder: A retest study. *Schizophrenia Bulletin, 14*, 57-65.
- Harvey, P. D., Docherty, N. M., Serper, M. R., & Rasmussen, M. (1990). Cognitive deficits and thought disorder. II. An 8-month follow-up study. *Schizophrenia Bulletin, 16*, 147-156.
- Harrow, M., Lanin-Kettering, I., & Miller, J. G. (1989). Impaired perspective and thought pathology in schizophrenic and psychotic disorders. *Schizophrenia Bulletin, 15*, 605-623.
- Hoffman, R. E. (1986). Verbal hallucinations and language production processes in schizophrenia. *Behavioral and Brain Sciences, 9*, 503-548.
- Hurt, S. W., Holzman, P. S., & Davis, J. M. (1983). Thought disorder: The measurement of its changes. *Archives of General Psychiatry, 40*, 1281-1285.
- Johnston, M. H., & Holzman, P. S. (1979). *Assessing schizophrenic thinking*. San Francisco, CA: Jossey-Bass.
- Lanin-Kettering, I., & Harrow, M. (1985). The thought behind the words: A view of schizophrenic speech and thinking disorders. *Schizophrenia Bulletin, 11*, 1-7.
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge: MIT Press.
- Lukoff, D., Nuechterlein, K. H., & Ventura, J. (1986). Manual for expanded Brief Psychiatric Ratings Scale (BPRS). *Schizophrenia Bulletin, 12*, 594-602.
- Maher, B. (1972). The language of schizophrenia: A review and interpretation. *British Journal of Psychiatry, 120*, 4-17.
- Marengo, J., Harrow, M., Lanin-Kettering, I., & Wilson, A. A. (1985). A manual for assessing aspects of bizarre-schizophrenic thinking. In M. Harrow & D. Quinlan (Eds.), *Disordered thinking and schizophrenic psychopathology* (pp. 394-411). New York: Gardner Press.
- Meng, X. L., Rosenthal, R., & Rubin, D. B. (1992). Comparing correlated correlation coefficients. *Psychological Bulletin, 111*, 172-175.
- Oltmanns, T. F., & Neale, J. M. (1978). Schizophrenic performance when distractors are present: Attentional deficit or differential task difficulty. *Journal of Abnormal Psychology, 84*, 205-209.
- Pandurangi, A. K., Sax, K. W., Pelonero, A. L., & Goldberg, S. C. (1994). Sustained attention and positive formal thought disorder in schizophrenia. *Schizophrenia Research, 13*, 109-116.
- Rochester, S. R., & Martin, J. (1979). *Crazy talk: A study of the discourse of schizophrenic speakers*. New York: Plenum.
- Schilder, P. (1951). On the development of thoughts. In D. Rapaport (Ed., Trans.), *Organization and pathology of thought* (pp. 497-518). New York: Columbia University Press.
- Shenton, M. E., Holzman, P. S., Solovay, M. R., & Coleman, M. (1991). Distinguishing between content and form of speech: Reply to Berenbaum. *Archives of General Psychiatry, 48*, 281.
- Shenton, M. E., Solovay, M. R., Holzman, P. S., Coleman, M., & Gale, H. J. (1989). Thought disorder in the relatives of psychotic patients. *Archives of General Psychiatry, 46*, 897-901.
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*, 420-428.
- Solovay, M. R., Shenton, M. E., Gasperetti, C., Coleman, M., Kestenbaum, J., Carpenter, T., & Holzman, P. S. (1986). Scoring manual for the thought disorder index. *Schizophrenia Bulletin, 12*, 483-496.
- Solovay, M. R., Shenton, M. E., & Holzman, P. S. (1987). Comparative studies of thought disorders. I. Mania and schizophrenia. *Archives of General Psychiatry, 44*, 13-20.
- Spitzer, R. L., Williams, J. B. W., Gibbon, M., & First, M. B. (1992). The Structured Clinical Interview for DSM-III-R (SCID). I. History, rationale, and description. *Archives of General Psychiatry, 49*, 624-629.
- Strauss, M. E., Buchanan, R. W., & Hale, J. (1993). Relations between attentional deficits and clinical symptoms in schizophrenic outpatients. *Psychiatry Research, 47*, 205-213.
- Taylor, M. A. (1981). *The Neuropsychiatric Mental Status Examination*. New York: Spectrum.