Evaluating Bizarre-Idiosyncratic Thinking: A Comprehensive Index of Positive Thought Disorder

by Joanne T. Marengo, Martin Harrow, Ilene Lanin-Kettering, and Arnold Wilson

Abstract

A summary of our method of assessing positive thought disorder, or bizarre-idiosyncratic thinking, from two short verbal tests is presented. This measure provides for standardized thought disorder assessments of: (1) the overall presence and severity of thought disorder, and (2) the type of disordered thinking shown. A definition and examples of bizarre-idiosyncratic thinking are provided, along with information on the reliability and validity of the scoring system. A method of establishing subject groups based on the severity of positive thought disorder also is presented—ranging from no thought disorder, to abnormal thinking, to severe formal thought disorder. This measure has been used to assess longitudinal changes in thought disorder symptoms over time and to evaluate relationships between thought disorder, other aspects of psychopathology (such as delusions), and adjustment in other areas of functioning across diagnostic groups.

Investigators have attempted to assess disordered thinking from a number of different viewpoints (Buss and Lang 1965; Maher, McKean, and McLaughlin 1966; Gottschalk and Gleser 1969; Payne 1970; Maher 1972; Chapman and Chapman 1973; Andreasen 1979a, 1979b; Chapman 1979; Johnston and Holzman 1979). Much of our own approach to studying positive thought disorder has revolved around the construct of bizarre-idiosyncratic speech and thinking (Harrow and Quinlan 1985; Marengo and Harrow 1985). A comprehensive assessment of bizarre-idiosyncratic thinking involves an evaluation of most of the major types of phenomena usually included under the term “formal positive thought disorder” (Fish 1962; Andreasen and Olsen 1982).

Since the concept of positive thought disorder deals with a potentially important symptom of disturbance, we have focused on this feature as one major aspect of a long-term, multidisciplinary research project studying the course of psychosis and adjustment in schizophrenia (Chicago Followup Study, Michael Reese Hospital and Medical Center, Chicago, Illinois). In this context, our studies of positive thought disorder have been primarily directed toward answering the following research questions: (1) Is thought disorder characteristic only of schizophrenia, or do other psychotic and nonpsychotic populations also demonstrate this behavior? (2) What changes in positive thought disorder occur over the course of the schizophrenic disorder? (3) Is the degree of disordered speech and thinking during posthospital or postacute stages of schizophrenia an index of the severity of disturbance? (4) What is the longitudinal relationship between disturbed language and thought and other aspects of psychiatric adjustment (e.g., work and social functioning, delusions and hallucinations)?

This manual represents a portion of a larger and more detailed manual by the same authors: The assessment of bizarre-idiosyncratic thinking: A manual for scoring responses to verbal tests. In: Harrow, M., and Quinlan, D., eds. Disordered Thinking and Schizophrenic Psychopathology. New York: Gardner Press, Copyright ©1985. This shorter and slightly modified version of the manual has been adapted for the current purpose with the kind permission of Gardner Press.

Reprint requests should be sent to Dr. J.T. Marengo, Dept. of Psychiatry, Michael Reese Hospital and Medical Center, 29th & Ellis Ave., Chicago, IL 60616.
The thought disorder measure presented here focuses on assessing the frequency and severity of positive thought disorder. While more qualitative typologies exist, it is assumed that measuring the degree of severity of thought disorder, which relates to hypotheses regarding a continuum of disturbance, will provide fruitful results.

The following system outlines the scoring of bizarre-idiiosyncratic verbalizations from two short verbal tests. The complete scoring manual and further examples can be found in Harrow and Quinlan (1985). This system represents an extension of prior research efforts in the area. Some of the material from the current procedure incorporates criteria developed by D. Adler that have been used successfully in previous research (Harrow, Tucker, and Adler 1972; Adler and Harrow 1973, 1974; Harrow and Quinlan 1977, 1985). However, more specific criteria are now included for evaluating bizarre-idiiosyncratic responses, and characteristic examples of such responses are offered as an aid to assessment. Analyses of disturbed speech and thinking, drawn from a variety of patient populations in various settings and contexts, have provided a framework for the current work.

**Definition of Bizarre-Idiosyncratic Thinking**

We have defined bizarre-idiiosyncratic speech and behavior as that which is: (1) unique to the particular subject; (2) deviant with respect to conventional social norms; and (3) frequently hard to understand, or to empathize with, in the context from which the response arose. While these three features are central to the concept, other less frequent characteristics are verbalizations that: (4) may appear confused, contradictory, or illogical; (5) may involve sudden or unexpected contrasts; and (6) are usually inappropriate or unresourceful in relation to the task at hand.

The first two of the above features are directly linked to social norms, though they are viewed from different vantage points: that of the individual and that of society. The last three characteristics listed in the definition may also appear as features of bizarre-idiiosyncratic thinking, but they are features that are present some of the time, instead of being invariant. It is surprisingly easy to spot bizarre-idiiosyncratic statements, since the manner of presentation or the ideas themselves impress the rater (including a naive rater) as unusual, strange, odd, or inappropriate.

The above definition provides us with a broad conceptualization of what are usually considered idiiosyncratic statements. It is important to emphasize that bizarre or idiiosyncratic qualities may be found in abstract and concrete as well as correct and incorrect responses, but they are analyzed and scored as a dimension separate from these other aspects of response behavior.

To illustrate, suppose the examiner presents the proverb:

Strike while the iron is hot.

And the subject responds with:

It could mean (pause) Hercules! (Could you say more?) I saw the movie Hercules. (Yes ...) and it means don't iron over your hands and don't strike anybody before you cast the first stone.

Although a response may be scored for positive thought disorder without meeting all of the above criteria, this particular response meets many of the criteria for bizarre-idiiosyncratic thinking noted above. It represents an explicit example of disordered thinking. The evaluator first wonders, “Hercules? What is the relation of the response to the original proverb?” The response deviates in an idiiosyncratic way from consensual norms for answering this question. Upon reflection, one sees some of the possible reasons for the subject’s interpretation in the suggestion of striking and violent action, with this leading to a response involving Hercules. In both this and other components of the response, it is difficult for the scorer to empathize with or understand the context from which the answer arose. In addition, the idea of ironing one’s hands is unexpected with respect to what has preceded it, and unusual on its own account. The entire statement appears confused, and the line of thinking that has led to the response has resulted in an idiiosyncratic answer to the question.

**Procedure**

The procedure includes: (1) assessing subjects with two short verbal tests; and (2) assigning an overall score for the severity of thought disorder to each of the subject’s responses. This procedure is outlined below. In addition, scoring criteria have been developed to characterize the subtypes of disordered thinking found in each response. Five categories and 11 subcategories of bizarre-idiiosyncratic thinking have been defined and will be summarized in this article. More complete scoring criteria for these categories can be found in Harrow and Quinlan (1985). A protocol representing a range of overall and category scores for bizarre-idiiosyncratic thinking is provided in Appendix I. This scored protocol may be used as a basis for practice and scoring comparisons.
Instruments and Administration of Tests. During the course of pilot work, it became clear that test materials vary in their potential for eliciting idiosyncratic verbal behavior. Even for those subjects whose speech would be labeled as odd by many, some types of verbal materials and tests do not readily elicit idiosyncratic responses. We have found that more open-ended procedures better lend themselves to eliciting this response behavior in those subjects with a potential for thought-disordered speech. Other tests prove cumbersome or unreliable. However, two verbal tests, the Gorham Proverbs Test (Gorham 1956) and the Comprehension Subtest of the Wechsler Adult Intelligence Scale (WAIS) (Wechsler 1955, 1981), are relatively short and easy to administer and were found to be satisfactory tools for evoking idiosyncratic verbal responses in those individuals with a potential for disordered thought.

The Proverbs Test. The Gorham Proverbs Test (clinical, free-answer set) consists of three parallel forms, each containing 12 proverbs. In our research, the multiple forms have been useful in the collection of longitudinal data over time from the same subject.

For purposes of standardizing subject assessment, directions for the test are read as follows:

I am going to read you some sayings. For example, the saying, “Large oaks from little acorns grow” could mean that great things may have small beginnings. Now, please tell me what each saying means instead of just telling me more about it. Try to answer every one.

WAIS Comprehension Subtest. The Comprehension Subtest of the WAIS (Wechsler 1955) or WAIS-R (Wechsler 1981) consists of questions relating to social comprehension and judgment (e.g., Why does land in the city cost more than land in the country?). Twelve of these questions provide the materials for our evaluation of bizarre-idsyncratic thinking (questions 3–14).

The Comprehension subtest is introduced to the subject as instructed in the WAIS manual. Encouragement toward responding is sometimes needed in this test, as well as in the proverbs test, and is appropriate to use in the manner standardized by the Wechsler manual.

Recording the data. It is essential that the verbalizations of the subject, as well as those of the examiner, be written down verbatim. The protocol also should include notations with reference to the behavior or affect stimulated by the test material or expressed during the course of testing. If it is impossible to record the verbalizations of an acutely disturbed patient, a transcribed tape recording may result in the most accurate test protocol. The examiner should inquire about all unclear or odd words or responses in a curious but nonintrusive manner, and these inquiries should be noted in parentheses. Since the object of these tasks is to obtain enough information to assess various aspects of speech and thinking that contribute to idiosyncratic responses, examiners must exercise judgment based on their own understanding of what has been said as the criteria for initiating inquiry.

Other verbal tests. Other comparable proverbs tests have been used in previous research (Benjamin 1944). Recently, we also have used a verbal test (Lanin-Berndt Communications Interview) with a series of standard questions designed to tap several important dimensions of verbal behavior (Berndt 1981; Lanin et al. 1981). Our efforts to derive a scoring system for tests of this general type led to the development of the current manual.

The Rorschach Test and free-verbalization situations have also been useful in our assessments of bizarre-idsyncratic thought (Quinlan et al. 1972; Quinlan and Harrow 1974; Reilly et al. 1975). Both of these techniques show many valuable features. However, the Rorschach Test takes longer to administer and to transcribe. The free-verbalization technique, while sensitive to eliciting bizarre-idsyncratic behavior, also presented some problems in its standardization and transcription.

Scoring

Overall score for bizarre-idsyncratic responses. An overall score for bizarre-idsyncratic thinking is assigned to each response. It assesses each response for its fit with the current verbal context and what is generally considered appropriate and understandable in our society. The extent to which a response as a whole is bizarre, as well as the extent to which it meets the criteria of any or all of the specific types of bizarre-idsyncratic thinking, varies greatly. In the present scoring system, we evaluate degrees of thought disorder by assigning scores of 0, .5, 1, or 3, ranging from absent to severe thought disorder. Several examples of responses at each level of bizarre-idsyncratic speech are provided here. Other examples with the appropriate overall score for bizarre-idsyncratic thinking are presented in Harrow and Quinlan (1985).

Overall score values

0 = Idiosyncratic verbalizations are absent.
.5 = Mild cognitive slips. Verbalizations that contain some mildly
idiosyncratic material. The response is slightly odd, but in a social situation the verbalization is not unusual enough to draw considerable attention. Two examples of responses scored .5 are presented below:

Q: When the cat's away, the mice will play.
A: Yeah. On the earth, up at the top, in the middle, XYZ. The end, the beginning of the end of the beginning.

Q: One swallow doesn't make a summer.
A: Boy, that's greedy as hell, man, that's real greedy. That's like pulling my actual backwards.

Summed overall scores for the proverbs and comprehension tests range from 0 to 36. In assigning the overall score for potential bizarre-idiosyncratic thinking to each response, the rater is essentially assessing how strange or deviant the response is in relation to more conventional answers. Even responses that one can understand or empathize with will at times be scored, since they may show odd features or deviate from social convention in an unusual or unexpected manner. However, responses that deviate from the conventional answer and that are also difficult to understand are assigned more severe thought disorder ratings.

As outlined above, an overall score of 0 is assigned when the response is not bizarre or idiosyncratic in any way. An overall score of .5 is assigned when the response is slightly off, or contains cognitive slips that are not grossly deviant. In a social situation, such responses would not really startle people or raise major questions. This score is meant to capture slight deviations, some of which are expected to be found in normal records as well. An overall score of 1 is assigned to a response that is clearly idiosyncratic or bizarre. An overall score of 3 is assigned only to extremely unusual or very bizarre statements.

Establishing Subject Groups Based on the Severity of Bizarre-Idiosyncratic Thinking

Continuum model. Responses reflecting thought disorder can be placed on a continuum extending from very severe bizarre-idiosyncratic thinking to normal, socially consensual thinking, with considerably heavier weightings for very severe bizarre responses, as opposed to mild cognitive slips. It also can be useful to have some system available to facilitate placing subjects into rough categories according to whether they show severe thought pathology, moderate levels of thought pathology, or no thought pathology. Accordingly, we assign subjects' total summed overall scores from the Proverbs and Comprehension tests to categories or levels reflecting various degrees of bizarre-idiosyncratic thinking. In this system, subjects' total overall scores have been categorized
into a continuum of positive thought disorder. The five specific categories included in the current system are: (1) no thought disorder; (2) minimal to mild bizarre-idsiosyncratic thinking; (3) signs of abnormal thinking or definite thought disorder; (4) severe thought disorder; and (5) very severe thought disorder. The conversion levels presented in table 1 are used.

Levels 1 and 2 of the Thought Disorder Scale are representative of little or no pathological thinking and lie well within the normal range. Level 3 (moderate level) reflects definite evidence of thought pathology or abnormal thinking. The number of patients at levels 4 and 5 provides an estimate of severely thought-disordered patients. The combined number of patients at levels 3, 4, and 5 gives the percentage with definite signs of positive thought disorder or abnormal thinking.

Composite index of bizarre-idsiosyncratic thinking. A composite index of bizarre-idsiosyncratic thinking also can be computed from both verbal tests, the Proverbs and the Comprehension tests, using the five levels of bizarre-idsiosyncratic thinking described above. In this composite index, each subject is assigned to the highest (most severe) level into which he or she fits on either of the two tests. This composite index classifies subjects according to whether they show any signs of bizarre-idsiosyncratic thinking, given opportunities on two separate tests to do so.

In our research we also obtain measures of thought disorder from the Object Sorting Test (Goldstein and Scheerer 1941; Goldstein 1944) and have categorized the scores from the Object Sorting Test into the same five levels of bizarre-idsiosyncratic thinking (Harrow et al. 1985). We have combined the category scores or levels from the Object Sorting Test with the parallel category scores or levels from the Proverbs and Comprehension tests to obtain our major composite index of bizarre-idsiosyncratic thinking (table 1).

This composite index is based on scores from the three separate measures of thought disorder. For this composite index, each patient is assigned to one of the five categories or levels ranging from no abnormal thinking (level 1) to very severe thought disorder (level 5) according to the most severe level attained on any of the three tests. This classification system categorizes patients according to whether they show any positive thought disorder, given three separate opportunities to do so. The composite score encompasses a broad sample of behavior and is more comprehensive than indexes that include only one or two tests.

Relation of this System of Evaluating Bizarre-Idsiosyncratic Thinking to Other Systems of Thought Disorder Assessment

Rapaport, Gill, and Schafer (1968) system. Note that in several ways the conceptual framework we employ to assess bizarre-idsiosyncratic thinking is different from that used by Rapaport in his research studying disordered thinking, although the two systems are comparable in several respects. Our conceptual framework is closely tied to implicit conceptual norms that people have acquired over time about what is appropriate and what is deviant in a particular response situation. The framework used by Rapaport is based, in part, on assessing pathological verbalizations by whether they involve too much distance or a loss of distance from the original stimuli, with such judgments being based on implicit social norms (Rapaport, Gill, and Schafer 1968). However, while there is some difference in the conceptual basis of the systems, there is much similarity in the actual practice of scoring disordered verbalizations. The system we use involves assessing consensually deviant responses and behavior along dimensions of thought disorder that Rapaport might have labeled as "too much distance" or a "loss of distance" from the original stimuli.

Johnston-Holzman Thought Disorder Index (1979). Our construct of bizarre-idsiosyncratic thinking and Johnston and Holzman's Thought Disorder Index also show similarity. In both systems, bizarre, strange, and deviant responses are assigned scores, and in both systems more severely deviant responses are given heavier or more pathological weighting. Johnston and Holzman's Thought Disorder Index is used with the WAIS Comprehension Subtest and with other WAIS subtests, as well as with the Rorschach Test.

Our Proverbs test measures of bizarre-idsiosyncratic thinking showed significant correlations ($r = .61$) with thought disorder as measured on Johnston and Holzman's (1979) Thought Disorder Index in a sample of young schizophrenic subjects studied during postacute phases of disturbance. We have found this system valuable. We tend, however, to favor the current verbal tests and system of assessment since they are less time consuming.

Andreasen's Scale for the Assessment of Thought, Language, and Communication (TLC) Andreasen's (1979a, 1979b) TLC Scale defines different language behaviors as subtypes of thought disorder. Ratings for the severity of thought disorder are included and are based on a standardized means of...
Table 1. Derivation of thought disorder groups and the Thought Disorder Scale from 3 psychological tests

<table>
<thead>
<tr>
<th>Thought disorder group</th>
<th>Thought Disorder Scale</th>
<th>Thought disorder continuum</th>
<th>Proverb or Comprehension Test total score</th>
<th>Object Sorting Test total score</th>
</tr>
</thead>
<tbody>
<tr>
<td>No abnormal thinking</td>
<td>1</td>
<td>Absent</td>
<td>0-.5</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Mild</td>
<td>1-2.5</td>
<td>1</td>
</tr>
<tr>
<td>Abnormal thinking</td>
<td>3</td>
<td>Definite</td>
<td>3-6.5</td>
<td>2</td>
</tr>
<tr>
<td>Severe thought disorder</td>
<td>4</td>
<td>Severe</td>
<td>7-11.5</td>
<td>3-6</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>Very severe</td>
<td>12-36</td>
<td>7-21</td>
</tr>
</tbody>
</table>

assessment. In general, the types of strange and bizarre thinking tapped in our measure of positive thought disorder are also assessed in Andreasen's composite index of positive thought disorder derived from the TLC. For instance, the various types of thought disorder assessed by Andreasen (e.g., tangentiality, incoherence, and loss of goal) are also scored for bizarre-idiosyncratic thinking when they appear in the verbal tests we administer to assess positive thought disorder.

RDC and DSM-III criteria of thought disorder. The key components of bizarre-idiosyncratic thinking in the current system also can be helpful in assessing whether a patient has the type of formal thought disorder listed in major diagnostic systems such as the Research Diagnostic Criteria (RDC) and DSM-III (Spitzer and Endicott 1968; American Psychiatric Association 1980). Four of the five types of pathological speech and thinking outlined in the RDC as constituting formal thought disorder (impaired understandability of speech, loosening of association or derailment, illogical thinking, and neologisms) are viewed in the present system as components of bizarre-idiosyncratic thinking.

Similarly, although DSM-III is a little less optimistic about the concept of formal thought disorder, three of the specific types of thought disorder which, in effect, are substituted in DSM-III for formal thought disorder (incoherence, marked loosening of associations, and markedly illogical thinking) are included in the present system. Note that these types of thought pathology are included as symptoms that can be used in DSM-III as partial criteria for a diagnosis of schizophrenia.

Reliability and Validity. Evaluations of both reliability and validity of the current system of assessment have been based on normal subjects and on outpatient and inpatient psychiatric groups. These assessments have taken place in a variety of settings and have included chronic and acute, psychotic and nonpsychotic, and medicated and unmedicated patient samples.

Reliability. Interrater reliability has been obtained for: (1) the overall scores of bizarre-idiosyncratic thinking using Proverbs and Comprehension tests; and (2) for the subcategories of bizarre-idiosyncratic thinking presented in the following section. Interrater reliability estimates always have been high for the overall scores and moderate to high for scoring the specific categories.

Four separate assessments of 20 patient records produced interrater reliabilities of $r = .93$, $r = .88$, $r = .67$, and $r = .91$ for total overall scores on the Proverbs and Comprehension tests. Interrater reliabilities for each category of idiosyncratic thinking were as follows:

Category: Range of $r$

I. Linguistic form and structure: .82-.99
II. Content of statement: Ideas expressed: .84-.99
III. Intermingling: .75-.85
IV. Relationship between response and question: .73-.99
V. Behavior: .47-.98

In addition to interrater reliability, internal consistencies for the Proverbs test using Cronbach's $\alpha$ produced a reliability of .85 for all possible combinations of item-by-item scores. Subjects were also tested on two parallel forms of the Proverbs test administered successively in the same test session.

Gorham's (1956) Proverbs Set 1 and Proverbs Set 3 were used. The correlation of subjects' overall scores for bizarre-idiosyncratic thinking between these two sets of proverbs was $r = .79$ (Harrow and Miller 1980).
The total overall scores of Proverbs and Comprehension tests for the same subjects have also been compared to determine whether a subject with more pathological scores on one of these tests is likely to have more pathological scores on the other test. In a sample of 18 unmedicated State hospital inpatients assessed with both tests during the first week of hospitalization, the reliability coefficient was $r = .74$ between Proverbs and Comprehension total scores. A second sample of 104 medicated patients (63 schizophrenic and 41 nonschizophrenic subjects) from a private and State hospital setting provided a reliability coefficient of $r = .53$ between Proverbs and Comprehension total scores. Significant correlations between the tests emerged for: (1) the total number of thought-disordered responses, $r = .57$; (2) the number of moderate and severe bizarre responses, $r = .55$; and (3) the frequency of severely bizarre responses, $r = .47$. This indicates that the bizarre-idiosyncratic thinking found is not peculiar to a specific test.

However, a significant difference emerged between the total scores derived from the two tests, $t = 2.69$, $df = 102$, $p < .01$. The Proverbs test elicited more severely bizarre responses than the Comprehension subtest when administered at the acute phase of patients' disorders. The Proverbs test appears to be more sensitive to eliciting thought-disordered responses in acute patients who have a tendency toward bizarre thinking (Marengo 1983). This appears to be a function of differences in test stimuli and demand characteristics. The Proverbs test presents less familiar material to the patient and demands more abstraction and more complex thinking. The subject cannot rely as much on previous experience and socially stereotyped responses when responding to many of the proverbs.

Validity. We have conducted a series of investigations which bear on the construct validity of this system for assessing positive thought disorder. A number of these studies are based on our longitudinal research assessing (1) the relationship between bizarre-idiosyncratic speech and other aspects of psychopathology, and (2) the level of bizarre-idiosyncratic thinking across different stages of patients' disorders. Early results on various patient and normal samples have supported the validity of the construct of bizarre-idiosyncratic thinking.

Our research on bizarre-idiosyncratic thinking with the Proverbs and Comprehension tests has indicated the following:

- Bizarre-idiosyncratic thinking was positively correlated with linguistic errors on a structured communication task (Lanin-Kettering 1983). Thought disorder as measured on the Proverbs Test showed a significant association with an overall index of contextual coordination, $r = .48$, $p < .001$, and was also correlated with several component measures, including the ability to maintain topic, $r = .29$, $p < .03$, and to produce sentence-to-sentence coherence, $r = .43$, $p < .003$. Assessment included a separate sample of psychiatric and normal subjects.

- Schizophrenic and nonschizophrenic patients ($n = 85$) with a disturbance of associative processes on a word association test also manifested significantly more severe idiosyncratic thinking in the Subtest of the WAIS (Silverstein, Harrow, and Marengo 1980).

- Correlations of the Proverbs and Comprehension test total scores with the Object Sorting Test measure of bizarre-idiosyncratic thinking were $r = .60$ and $r = .50$, respectively, in 50 psychiatric subjects assessed at the acute inpatient phase.

- We have assessed the relationships between earlier versions of our indexes of bizarre-idiosyncratic thinking and a measure of communication deviance employed by some theorists on schizophrenic families (Wild et al. 1965; Wild 1972), using the Object Sorting Test. In a study of 40 schizophrenic and nonschizophrenic patients, these relationships were significant. Wild and Lidz's measure of transactional thinking was significantly related to our earlier measures of bizarre-idiosyncratic thinking assessing WAIS Comprehension Subtest responses ($r = .45$), and to our similar measure assessing Object Sorting Test responses ($r = .67$) (Quinlan et al. 1978). These results suggest a significant relationship between measures of bizarre-idiosyncratic thinking or positive thought disorder and measures of communicative deviance from family studies.

In a series of studies using the composite index of positive thought disorder (including the three thought disorder measures), the system described above has proved useful.

- At the acute phase of disturbance, early schizophrenic patients showed significantly more severe bizarre-idiosyncratic thinking than other psychotic and nonpsychotic patient groups (with the exception of manic subjects), indicating the index's sensitivity to diagnostic factors (Marengo and Harrow 1985).

- Differences exist in the severity of bizarre-idiosyncratic thinking over different stages of psychiatric disturbance. A decline in the severity
of bizarre-idiosyncratic thinking is evident between acute-phase assessment and the stage of partial recovery (7 weeks after acute-phase assessment). This reduction is significantly associated with improvement in other aspects of patients’ clinical conditions (Harrow et al. 1982; Harrow, Marengo, and Lanin-Kettering 1983). 

- In a longitudinal study we found significant schizophrenic-nonschizophrenic differences in disordered thinking at the acute phase, and considerably smaller differences at 1.5-year followup, a postacute period (Harrow, Marengo, and McDonald 1986).

- We have found significant associations between the Thought Disorder Scale (based on the composite index) and other major psychopathology, such as delusions (Harrow, Silverstein, and Marengo 1983; Marengo and Harrow 1985; Harrow and Marengo, this issue; Harrow, Marengo, and McDonald 1986). In addition, we have found significant relationships between the Thought Disorder Index and the degree of dysfunction in current and subsequent overall adjustment shown by schizophrenic and other psychiatric groups (Harrow, Silverstein, and Marengo 1983; Marengo 1983; Harrow and Marengo, this issue; Harrow, Marengo, and McDonald 1986). In addition, we have found significant relationships between the Thought Disorder Index and the degree of dysfunction in current and subsequent overall adjustment shown by schizophrenic and other psychiatric groups (Harrow, Silverstein, and Marengo 1983; Marengo 1983; Harrow and Marengo, this issue; Harrow, Marengo, and McDonald 1986).

While we view the points of the thought disorder scale as lying on a continuum, the specific cutoff points that establish the thought disorder groups have proved valuable. When patients are categorized according to whether they have severe thought disorder, as opposed to signs of abnormal thinking, as opposed to little or no thought disorder on our composite index, significant differences among groups in various types of major psychopathology emerge (Harrow and Marengo, this issue). This has occurred in comparisons among schizophrenic patients with no or only mild thought disorder (Thought Disorder Scale levels 1 and 2), others showing signs of abnormal thinking (Thought Disorder Scale level 3), and those exhibiting severe or very severe thought disorder (Thought Disorder Scale levels 4 and 5). Thought disorder group differences have been found in concurrent assessments of work functioning, overall psychotic activity, and rehospitalization rates during posthospital phases of schizophrenia and major depressive disorders (Marengo 1983). These data indicate some degree of validity for the cutoff points established and listed above.

In other phases of our research, the Thought Disorder Scale and the thought disorder groups (using the composite index) were predictive of the future functioning of patients when thought disorder was assessed at an early, posthospital phase of disorder. The presence of severe thought disorder (Thought Disorder Scale level 4 or 5) is a good prognostic indicator of future sustained functioning impairments in both schizophrenic and nonschizophrenic disorders (Harrow and Marengo, this issue). By contrast, non-thought-disordered patients and those showing abnormal thinking (but not severe thought disorder) evidence a more benign or remitting subsequent course of psychosis than do severely thought-disordered patients over the longitudinal course of disturbance (Marengo 1983).

In these early studies, the results from the Thought Disorder Index have generally fit in with common clinical observations and theories about bizarre-idiosyncratic thinking and positive thought disorder.

Note that in our assessments of nonpatient control samples, some normal subjects have manifested signs of abnormal thinking (Marengo and Harrow 1985). Normal subjects have also demonstrated signs of severe positive thought disorder. From 10 to 15 percent of normal subjects showed level 4 or 5 on the Comprehension or Proverbs test when one of these tests was used alone, and 10 percent showed severely disordered thinking on the composite index (Marengo and Harrow 1985). The finding of some control subjects with severe levels of thought disorder should not be surprising, in view of various findings over the years of certain levels of psychopathology, and even severe psychopathology, in some “normals.”

We should also point out that thought-disordered individuals show a great deal of normal speech. This attests to the phasic nature of thought disorder, in some subjects, even during periods of acute upset. Our measures assess the behavioral potential for bizarre-idiosyncratic thinking in spontaneous, open-ended speech situations.

**Specific Categories and Subcategories of Bizarre-Idiosyncratic Thinking**

In addition to the overall scores for thought disorder severity, we have outlined a system for evaluations of each response focusing on criteria constructed to delineate the anomalies of positive thought disorder. Five categories and 11 subcategories of bizarre-idiosyncratic thinking provide the criteria for evaluating components of bizarre-idiosyncratic verbalizations in greater detail. These major categories and subcategories represent various types of abnormal ideas, behavior, and language. They also provide one way
of categorizing some of the different types of disordered thinking that can be seen in response to specific tasks, as well as in people's day-to-day behavior.

The presence of these components of disordered thinking can be independently studied in light of different theoretical predictions, developments at different points in the unfolding of, or recovery from, a particular disorder, and in understanding differences among clinical populations. With bizarre-idiosyncratic thinking as a more general construct, these categorical evaluations provide the opportunity to study particular kinds of thought disorder.

We have found these subtypes of bizarre-idiosyncratic verbalizations and behavior useful, and we do score them in our own research. We should emphasize, however, that scoring or attending to these specific criteria is not absolutely necessary for attaining the overall score for bizarre-idiosyncratic thinking. One can use the overall system of assessment of thought disorder severity on the basis of the criteria outlined in the previous pages without the detailed and specific subcategories noted in the following pages. If the following subcategories are used, the initial overall rating of bizarre-idiosyncratic thinking should be made first, before scoring these individual subcategories.

We have used five basic categories to study both the subject's linguistic form (i.e., the manner in which ideas are communicated) and the content of the responses (i.e., the ideas themselves). Detailed examples of specific responses and how they should be scored within the 5 categories and 11 subcategories of bizarre-idiosyncratic responses are provided in Harrow and Quinlan (1985). These categories are not exclusive, and many responses are scored for several types of component problems, particularly since problems in form and problems in content are often difficult to tease apart.

Five Categories of Bizarre-Idiosyncratic Thinking

I. Linguistic form and structure. Here, the structure of language within the response is under scrutiny. A problem in this area implies that it is difficult to understand the subject's statement owing to distortions in word use, grammatical form, or the linkage of words and phrases. A response also may be communicated poorly. Questions are raised about peculiarities in the individual's verbal style, the linguistic structure of the response, or gaps in communication that may interfere with the clear communication of meaning.

II. Content of the statement: The ideas expressed. Under primary consideration are the ideas presented within the response. This category pertains to peculiarities within a response such as idiosyncratic reasoning, asocial attitudes, and disorganized or confused ideas. The ideas or attitudes the subject presents are evaluated from the perspective of conventional attitudes, logical thought, or cogent explanations.

III. What is intermixed into the response. In this category, the focus is on whether the response moves away from a consensual answer owing to a shift to loosely associated ideas or the subject's personal preoccupations. The twofold emphasis of this category is on: (1) the mixing or blending into the response of personal material from the subject's past or current experience, and (2) the extensive elaboration of a theme or idea, which does not fit neatly into the structure of the response, making it appear somewhat unusual. The question raised is: Are there any personal associations or tangential ideas that the subject is expressing as part of his response that are idiosyncratic or bizarre?

IV. Relationship between question and response. Here, the focus of evaluation is on how the ideas the subject presents relate to the question asked. Is the subject able to address the task of interpretation, or answer a question? Is the subject inappropriately, or totally, focused on private, autistic associations or thoughts?

V. Behavior. Behavior is assessed using conventional norms for conduct in a testing situation, including physical, affective, and verbal behavior. Is the subject's behavior deviant either in its extreme expression or in its incongruity and impropriety within the testing context?

Suggested Step-by-Step Procedure for Evaluating the Individual Categories of Bizarre-Idiosyncratic Thinking.

The assignment of scores within these categories is determined using the criteria listed below. The extent to which a response meets each (or none) of the relevant category criteria is evaluated and the response is assigned a score of 0, .5, 1, or 3 for each category.

In assigning category and subcategory scores, it helps to:

- Consider how the subject expresses ideas. Is the response poorly structured, composed of vague elements, or obscured by idiosyncratic terminology? What is the subject's choice of words? Is there any unusual word usage or use of artificial, pedantic, or stilted language? Does the subject communicate ideas in a reasonable fashion, or are there unexplained
gaps or missing referents that make it very difficult to understand? If yes, then evaluate the response in terms of subcategory indicators and assign a score reflecting the degree of thought disorder in Category I.

• Consider next the ideas presented. Does the subject's statement make any sense at all? If the response makes no sense, score a 3 in Category II. Does the reasoning deviate from rules of logic or social convention? Does the statement involve an odd meaning or outlook? If yes, then evaluate the extent and assign an appropriate score in Category II and in its appropriate subcategories.

• Then evaluate intermingling in the response. At times a subject's answer shows intermingling of personal material from the very beginning of a response. At other times, after a smooth beginning (in which the initial thrust of the response is partially or completely correct), does the subject's response then wander onto tangential or irrelevant topics? Does the response reflect intermingled personal material? If yes, then evaluate the extent to which this additional material causes the response to appear odd and assign scores to the appropriate subcategories and to Category III.

• Now consider the relation of the responses to the question or proverb. Does the content of the response (at least initially) refer to the question asked (irrespective of whether it is correct or incorrect, concrete or abstract, and despite other aspects of bizarreness)? Is there any discernible association between the words or concepts of the question or proverb on the one hand, and the subject's statement on the other? Assign a score reflecting the distance of the response from the question or proverb in Category IV.

• Finally, consider whether the subject's physical, affective, or verbal behavior deviates from conventional norms during the test situation. If it does, then evaluate the degree and assign a score of 0, .5, 1, or 3 in Category V.

Each of the five categories is scored on the scoring sheet. The scoring of the above five categories is based primarily on the following 11 subcategories of disturbed language and thinking. These indicators have proved useful in orienting raters to the behavioral foundations of the five descriptive categories of idiosyncratic thinking. In addition, some of these subcategories may be related to specific mechanisms involved in bizarre-idiosyncratic thinking. Some bizarre-idiosyncratic responses may be influenced by, or in part be a consequence of, confusion-disorganization, a tendency to intermingle personal concerns when thinking about neutral material, a disorder in logical reasoning, or attending to part rather than all of a question. However, while these 11 subtypes of bizarre-idiosyncratic thinking may be important, they do not represent an exhaustive list of all possible types of bizarre behavior and ideas, or all the possible dimensions with which one can look at these phenomena. There are a vast number of ways responses can be strange and bizarre, and a vast number of ways that responses can deviate from social convention in an idiosyncratic manner.

Category II: Content of the Statement

3. Responses involving coherent but odd ideas.
4. Responses that are deviant with respect to social convention.
5. Peculiar or idiosyncratic reasoning or logic.

Category I: Linguistic Form and Structure

1. Strange verbalizations.
   a. Single words used in an unusual or peculiar manner (which are, in the rater's best judgment, not attributable to intellectual or cultural deficits).

b. Mild or moderate cognitive slippage in sentence structure, the expression of ideas, or the construction of new words (the new word is close in form to the correct word).

c. Neologisms (a new word with private meaning). Real neologisms (involving a private meaning) are very unusual, and are scored "3" to reflect their very severe idiosyncratic quality.

d. Artificial, pedantic, or stilted language, inappropriate to the level of discourse in the testing situation.

2. Lack of shared communication.
   a. Responses that are not explicitly stated.

   b. Small gaps in communication, in which words are not explained or referents are unclear.

   c. Larger gaps in communication, in which phrases are not explained. Elements of private language may be apparent with unshared or unexplained concepts or ideas.

   d. Disorganized or poor linkage between consecutive words, phrases, or sentences within the response.

Category II: Content of the Statement

3. Responses involving coherent but odd ideas.
4. Responses that are deviant with respect to social convention.
5. Peculiar or idiosyncratic reasoning or logic.

   a. Responses that are incorrect and illogical in terms of common knowledge about people, events, or the environment.

   b. Responses violating a logical paradigm, such as predicate logic.

   c. Self-contradictory responses or responses with confused logic.

   d. Responses with peculiar, autistic logic.
6. Confused or disorganized ideas.
   a. Combinations of words put together in a manner that only dimly makes sense.
   b. Grammatically correct sentences that do not hold a logical thought.

Category III: Intermixing
7. Overelaborated response.
   a. Irrelevant wandering within a partially correct or correct answer.
   b. Elaboration that is far too extensive, to the point where the original question is almost lost from sight.
8. Intermingling of personal concerns or associations into the response.

Category IV: Relationship Between Question and Response
9. Attending to a part of rather than the whole question: associations or interpretations of a word or phrase that suggest the subject's response is not based on the question as a conceptual whole, and also make the response appear strange or idiosyncratic.
10. Lack of a relationship between the subject's statement and the question asked—almost as if a different question is being asked.

Category V: Behavior
11. Strange behavior—including physical and affective behavior.

In scoring a response according to the list of major categories and subcategories outlined above, the scorer first analyzes the subject's response and scores it according to each of the 11 subcategories. The scorer then assigns scores on the five major categories. The presence of one type of subcategory of bizarre-idiosyncratic thinking in a response (e.g., lack of shared communication) does not exclude the simultaneous presence of another type or subcategory from that same response (e.g., intermingling or an overelaborated response).

The categories were constructed to focus on the distinct properties of a response that may contribute to an overall impression of unusual or odd verbal behavior. However, at times, although a category score is indicated (i.e., something unusual occurs in linguistic style, or in how a response is stated), the idiosyncrasy is not attributable to any specific subcategory or behavioral indicator. In such instances, a category score for bizarre-idiosyncratic thinking is still justified in that particular area, while the individual behavioral descriptors are left blank.

The categories and the subcategories that they comprise are not an exhaustive list of all possible aspects of idiosyncratic thought and language. The overall response score is a general barometer of bizarre-idiosyncratic thinking, and the categories and their foci are probes for the components of such verbal behavior. Any one response may be scored for overall bizarre-ness but may not fit neatly into any of the outlined categories or subcategories, with all of the categories and subcategories consequently rated as 0.

On the other hand, if a score of 1 is assigned for any of the five categories, an overall score of at least 1 is logically indicated, as the overall response or part of it is clearly bizarre or idiosyncratic. Although the overall score should be at least as great as that given in any individual category, the accumulation of category scores may, and often does, add up to more than the overall score. (The overall score for a response can only equal 0, .5, 1, or 3 and is not the summation of category scores.)

As we have indicated, many responses are scored in more than one category or subcategory. In a sample of hospitalized schizophrenic patients, for example, we found that of those responses scored for bizarre-idiiosyncratic language and thought, approximately 50 percent were scored in one subcategory, 30 percent in two subcategories, and 20 percent in three or more subcategories. This ratio, however, may differ with varying populations.

Conclusion
In this article, we have summarized our method of assessing positive thought disorder in two verbal tests. This has included information on some of our conceptualizations about bizarre-idiosyncratic thinking and positive thought disorder, as well as the reliability and validity of our scoring system. While we have conveyed the essential aspects of our scoring system here, more thorough and detailed scoring procedures for the categories of bizarre-idiiosyncratic thinking and additional scoring examples are presented in Harrow and Quinlan (1985).

References
American Psychiatric Association. DSM-III: Diagnostic and Statistical Manual of Mental Disorders. 3rd ed.


Silverstein, M.L.; Harrow, M.; and Marengo, J. "Thought Disorders: One or Many?" Unpublished manuscript, Michael Reese Medical Center, Chicago, IL, 1980.


**Acknowledgment**

This work was supported in part by Grant No. MH-26341 from the National Institute of Mental Health, and by research grants from the John D. and Catherine T. MacArthur Foundation and the Irving B. Harris Foundation.

**The Authors**

Joanne T. Marengo, Ph.D., is Senior Psychologist, Department of Psychiatry, Michael Reese Hospital and Medical Center, and Research Associate (Assistant Professor), Department of Psychiatry, The University of Chicago. Martin M. Harrow, Ph.D., is Director of Psychology, Michael Reese Hospital and Medical Center, Chicago, IL, and Professor, Department of Psychiatry and Behavioral Sciences, The University of Chicago. Ilene Lanin-Kettering, Ph.D., is Senior Account Executive, Kapuler Marketing Research, Inc., Arlington Heights, IL. Arnold Wilson, Ph.D., is Assistant Professor, Department of Psychology, New School for Social Research, New York, NY.
Appendix 1: Scored Protocol

1. Where there's a will, there's a way.
   Destiny is always truth.

2. Rome was not built in a day.
   You'd have to ask my brother that. He's an architect. In the Koran, my brother is Moses, so he's the best to answer that question. I'm studying the Romans in the Bible now. I'm studying their mental attitudes toward life. I'm in school now—University of Mozambique. It's through the mind I'm learning.¹

3. When the cat's away, the mice will play.
   It refers to supervision—children get aggressive and selfish when their parents aren't there. Behavior is mischievous when there is no supervision.

   Intimidation. (Q) A barking dog is trying to frighten or intimidate you.

5. A stream cannot rise higher than its source.
   Water doesn't run up hill.

6. Don't swap (trade) horses when crossing a stream.

7. The used key is always bright.
   I don't know what that means in frontal knowledge, but not in lateral.

8. Gold goes in at any gate except heaven's.
   Heaven has a gate of gold. You can't buy heaven.

9. One swallow (bird) doesn't make a summer.
   Teaches patience.

10. The wife is the key to the house.
    She's the protective image of the household. She's the leader behind closed doors.

11. Riches serve a wise man but command a fool.
    That's true. (Enc.) It's the difference between sitting yoga and traveling or putting on your clothes and using transportation. (Q) It's transcendental sound.

12. Don't cast pearls before swine (pigs).
    Swine are only intended to eat ground or soil but not pearls which are the epitome of eternity.

¹Subject has never attended a foreign university.
# Score sheet for bizarre-Idiosyncratic responses

<table>
<thead>
<tr>
<th>Response number</th>
<th>Overall score</th>
<th>I. Linguistic form or structure</th>
<th>II. Content of the statement, ideas expressed</th>
<th>III. Intermixing</th>
<th>IV. Relationship between response and question</th>
<th>V. Behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prov/Comp</td>
<td></td>
<td>Peculiar word form or use</td>
<td>Coherent but odd ideas</td>
<td>Deviant with respect to social convention</td>
<td>Confused Ideas</td>
<td>Overall category II</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>0</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>3</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>3</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>0</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
<td>.5</td>
<td>9</td>
<td>.5</td>
<td>.5</td>
<td>.5</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>11</td>
<td>1.5</td>
<td>11</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
</tr>
<tr>
<td>12</td>
<td>12</td>
<td>1</td>
<td>12</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>13</td>
<td>13</td>
<td>1</td>
<td>13</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>14</td>
<td>1</td>
<td>14</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

**Overall Sum:** 145 4.0 5.0 8.5 5.0 00 00 3.0 8.0 3.0 3.0 3.0 1.0 7.0 8.0 00

**Sum 1 & 3pt. rsp.:** 14 4 4 8 5 0 0 3 8 3 3 3 1 7 8 0

**# of 0's:** 05 10 8 7 9 12 12 11 8 11 11 11 11 9 8 12

**# of .5's:** 01 02 1 0 0 0 0 0 0 0 0 0 0 0 0 0

**# of 1's:** 02 1 1 2 2 0 0 0 2 0 0 0 1 1 2 0

**# of 3's:** 04 1 1 2 1 0 0 1 2 1 1 1 0 2 2 0

**Note:** Prov/Comp. = Proverbs test, Comprehension test; pt. rsp. = response point.