The Thought Disorder Index: A Reliability Study

Michael J. Coleman, J. Tyler Carpenter, Christine Waternaux, Deborah L. Levy, Martha E. Shenton, Jonathan Perry, David Medoff, Hazel Wong, Dara Monoach, Patricia Meyer, Carita O'Brian, Camille Valentino, Delbert Robinson, Melissa Smith, David Makowski, and Philip S. Holzman

This article reports interrater reliability among 4 teams of raters who independently evaluated thought disorder in 20 Rorschach protocols using the Thought Disorder Index (TDI). Intraclass correlation coefficients were calculated to assess the degree of association among the 4 teams for total thought disorder scores, severity levels, and qualitative thought disorder factors. Highly acceptable agreement was obtained. Spearman rank order correlation coefficients for these same variables were significant for all possible pairings of teams. A repeated-measures analysis of variance indicated that the absolute amount of thought disorder tagged by each team differed even though the teams' relative rankings of thought disorder among subjects was very similar. Such scoring differences reflect individual differences in threshold for detecting deviant thinking.

Thought disorder has long been regarded as a hallmark of schizophrenia. Kraepelin (1896/1919) described the disordered thought processes of schizophrenia as "derailments." Bleuler (1911/1950, 1924) wrote that the central feature of schizophrenic thought was "loosening of associations," which included associations by contiguity, bizarre and clang associations, as well as condensations of ideas and incoherent discourse. Subsequent investigations have confirmed the presence of these disorders of thinking in schizophrenic conditions but have also identified thought disorder in patients with affective psychoses, organic brain diseases, and borderline personality disorder (Andreasen & Powers, 1974; Cameron, 1944; Carlson & Goodwin, 1973; Clayton, Pitts, & Winokur, 1965; Edell, 1987; Gershon, Benson, & Frazier, 1974; Goldstein, 1944; Harrow & Quinlan, 1977; Johnston & Holzman, 1979; O'Connell, Cooper, Perry, & Hoke, 1989; Shenton, Solovay, & Holzman, 1987; Solovay, Shenton, & Holzman, 1987).

The study of thought disturbances in psychotic patients has spawned the development of objective scales for the measurement of thought disorder. Several such scales are in common use, including the Scale for the Assessment of Thought, Language, and Communication (TLC; Andreasen, 1979, 1986); the thought disorder assessment scale devised by Harrow and colleagues (Harrow & Quinlan, 1985; Marengo, Harrow, Lanin-Kettering, & Wilson, 1986); and the Thought Disorder Index (TDI; Johnston & Holzman, 1979; Solovay et al., 1986). The psychometric properties of these scales, including their interrater reliability, are a necessary requirement for evaluating these scales. This article reports a study of interrater reliability assessments for the TDI.

The TDI is a multidimensional system for tagging, classifying, and measuring disturbances in thinking. It provides both a qualitative and quantitative inventory of cognitive slippage. The TDI emerged from the work of Rapaport, Gill, and Schafer (1945/1968), who developed a set of scoring categories to identify thinking disturbances in psychotic and neurotic adults. These scoring categories were later refined by Watkins and Stauffacher (1952), who fashioned them into a quantitative scoring system, the "Delta Index." The Delta Index yielded severity-weighted scores that were summed and divided by the number of Rorschach responses to yield a quantitative index of thought disorder. Johnston and Holzman (1979) further revised this system, eliminating unreliable scoring categories and adding other categories to the system. Subsequent refinements have also been made (Solovay et al., 1986).

The TDI has been used in several different studies to (a) discriminate among schizophrenic, manic, and schizoaffective patients on the basis of qualitative differences in thought disorder (Holzman, Solovay, & Shenton, 1985; Shenton, Solovay, & Holzman, 1987; Solovay, Shenton, & Holzman, 1987); (b) detect thought disorder in patients with borderline and schizotypal personality disorders (Edell, 1987); (c) predict psychotic symptoms in patients with personality and affective disorders (O'Connell, Cooper, Perry, & Hoke, 1989); (d) measure neuroleptic-induced reduction of thought disorder in schizophrenic patients (Hurt, Holzman, & Davis, 1983; Spohn et al., 1986); (e) identify unique patterns of thought disorder in patients with unilateral right hemisphere cortical damage (Daniels et al., 1988); (f) detect the often subtle cognitive slippage that occurs in the unaffected first-degree relatives of schizophrenic, manic,

Michael J. Coleman, Harvard Medical School/McLean Hospital; J. Tyler Carpenter, Harvard University/McLean Hospital; Christine Waternaux and Deborah L. Levy, Harvard Medical School/McLean Hospital; Martha E. Shenton, Harvard Medical School/Brockton Veterans Affairs Hospital; Jonathan Perry, Harvard University/McLean Hospital; David Medoff and Hazel Wong, McLean Hospital; Dara Monoach and Patricia Meyer, Harvard University/McLean Hospital; Carita O'Brian, McLean Hospital; Camille Valentino, Delbert Robinson, Melissa Smith, and David Makowski, Hillside Hospital (a division of Long Island Jewish Medical Center); and Philip S. Holzman, Harvard University, Harvard Medical School, and McLean Hospital.

Correspondence concerning this article should be addressed to Michael J. Coleman, Psychology Laboratory, McLean Hospital, 115 Mill Street, Belmont, Massachusetts 02178.

and schizoaffective patients (Shenton, Solovay, Holzman, Coleman, & Gale, 1989); and (g) characterize the amounts and kinds of thought disorder shown by children who were hospitalized for psychotic and behavioral symptoms, children whose mothers had been hospitalized for psychotic symptoms, and normal children (Arboleda & Holzman, 1985). The TDI has been shown to be unrelated to race and socioeconomic status (Haimo & Holzman, 1979; Johnston & Holzman, 1979).

In principle, any verbal sample can be assessed using the TDI, although it is most commonly applied to responses to Rorschach cards (Rorschach, 1921/1942). The Rorschach Test is advantageous for two reasons. First, it seems to elicit more instances of thought disorder than such nonprojective tests as the Wechsler Adult Intelligence Scale (Johnston & Holzman, 1979; Spohn et al., 1986). Second, TDI scores based on the Rorschach Test are uncorrelated with IQ (Johnston & Holzman, 1979), whereas those based on the WAIS are significantly but modestly and negatively correlated with IQ (Gold & Hurt, 1990; Johnston & Holzman, 1979).

Previously, Johnston and Holzman (1979) reported interrater reliability among two raters for total TDI scores. Pearson r ranged from .82 to .93. Solovay, Shenton, and Holzman (1987), also using two judges in the same laboratory, reported interrater reliability for total TDI scores of .89, individual scoring categories of .81, severity levels of .79, and various factor scores ranging from .84 to .89. We report here a systematic study of interrater reliability on all features of TDI scoring, using four rating teams at three separate sites, who independently scored the same Rorschach responses using the TDI.

Method

Instrument

The TDI distinguishes 23 qualitative categories of thought disturbance, which are weighted along a continuum of severity (.25, .50, .75, and 1.0), with the .25 level representing very mild forms and the 1.00 level reflecting the most severe forms of thought disorder. The TDI categories and severity levels are listed in Table 1. The inventory in Table 1 makes no claim to completeness; it does, however, comprise most types of thought disorder encountered—with greater or lesser frequency—in the examination of psychotic patients. The TDI manual contains more complete descriptions of the categories as well as the psychometric characteristics of the validation studies (Johnston & Holzman, 1979; Solovay et al., 1986).

The total thought disorder score is computed as the sum of the frequency of each instance of thought disorder, multiplied by its category weight (e.g., .25, .50, .75, or 1.0), divided by the number of Rorschach responses to control for verbal productivity, and multiplied by 100 to express the value as a percentage:

$$\frac{(.25 \times a) + (.50 \times b) + (.75 \times c) + (1.0 \times d)}{R} \times 100$$

where a = the number of thought-disordered responses at the .25 level; b = the number of thought-disordered responses at the .50 level; c = the number of thought-disordered responses at the .75 level; d = the number of thought-disordered responses at the 1.0 level; and R = the total number of Rorschach responses.

The relatively infrequent occurrence of many individual scoring categories necessitates the grouping of thought disorder categories for pur-

Table 1
Thought Disorder Index Scoring Categories by Level of Severity

.25 level

1. Inappropriate distance
   a. Increase of distance
   b. Excessive qualification
   c. Concreteness
   d. Overspecificity
   e. Syncretistic response
2. Flippant
3. Vagueness
4. Peculiar verbalization
5. Word-finding difficulty
6. Clang
7. Perseveration
8. Incongruous combination
   a. Composite responses
   b. Arbitrary form–color
   c. Inappropriate activity

.50 level

8. d. Internal–external response
9. Relationship verbalization
10. Idiosyncratic symbolism
    a. Color symbolism
    b. Image symbolism
11. Queer response
    a. Queer expression
    b. Queer word usage
    c. Queer imagery
12. Confusion
13. Looseness
14. Fabulized combination
15. Playful confabulation
16. Fragmentation

.75 level

17. Fluidity
18. Absurd response
19. Confabulation
    a. Details in one area generalized to a larger area
    b. Extreme elaboration
20. Autistic logic

1.00 level

21. Contamination
22. Incoherence
23. Neologism

poses of analysis. In previous studies, several techniques were used to group thought disorder categories or to create factors. The first was a grouping based on an a priori classification of the 23 TDI categories. The second was an empirical classification based on a principal-components analysis with varimax rotation. The third was a post hoc grouping of categories based on both the a priori and the principal-components analysis results. These factors have demonstrated their usefulness in distinguishing diagnostic groups (Holzman, Solovay, & Shenton, 1985; Shenton, Solovay, & Holzman, 1987; Solovay, Shenton,

& Holzman, 1987). Unit weights have been used to compute the factor scores. The post hoc factors best differentiated diagnostic groups, and therefore we adopted the post hoc grouping for use here. The derivation of these factors and their application to diagnosis are described elsewhere (Shenton, Solovay, & Holzman, 1987; Solovay, Shenton, & Holzman, 1987).

Subjects

Twenty Rorschach protocols were randomly selected from a larger sample of protocols that had been obtained from subjects who had participated in a research project on the major psychoses. The protocol consisted of responses to either a 4- or 10-card Rorschach. The 4-card combinations correlate at acceptably high levels with the full 10 cards for total TDI score as well as for most of the individual categories of thought disorder. We recently compared the amount and kinds of thought disorder identified by the TDI from 4-card combinations of Rorschach cards with that using the full 10 cards. Carpenter et al. (1993) demonstrated that short 4-card forms of the standard 10-card administration of the Rorschach yield excellent-to-good composite indices of total thought disorder, severity levels, and the more frequently occurring qualitative factors (e.g., Irrelevant Intrusions, Idiosyncratic Verbalizations, Confusion, and Combinatory Thinking). The study consisted of making comparisons between assessments of thought disorder using the 10-card and 4-card forms in one sample.

These comparisons were then replicated on a second equivalent sample. The analyses yielded the following results: Intraclass correlations (ICC) for the total TDI score were .81 and .84 for the two samples; ICCs for the three most frequently occurring Severity Level factors ranged from .78 to .90; and ICCs for the previously mentioned qualitative factors ranged from .49 (Combinatory Thinking) to .85 (Irrelevant Intrusions). It therefore appears that for the purposes of scrutinizing interrater reliabilities, the 4- and 10-card protocols are essentially equivalent.

The entire database from which this sample was drawn consists of protocols administered to inpatients who met the criteria of the third edition of the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM-III*; American Psychiatric Association, 1980) for a diagnosis of schizophrenia, schizoaffective disorder, or bipolar disorder, as well as the first-degree biological relatives of these patients. Patient diagnoses were independently assessed by a team of diagnosticians who reviewed all available medical records and the Structured Clinical Interview for *DSM-III* (Spitzer, Williams, & Gibbon, 1988).

Diagnostic reliability, assessed by Cohen's Kappa, was .84 for bipolar disorder and .85 for schizophrenia. The subset of protocols used in this study was obtained from subjects in the following groups: schizophrenia (*n* = 5), schizoaffective disorder (*n* = 4), bipolar disorder (*n* = 8), first-degree relative of schizophrenics (*n* = 2), and first-degree relative of schizoaffective patients (*n* = 1). The 20 protocols used in this reliability study were obtained from unrelated individuals.

Procedure

We randomly selected 20 protocols from the larger database. By this method, both full 10-card as well as 4-card protocols were selected. Thirteen subjects received a 4-card Rorschach; 6 subjects received a 10-card Rorschach, and 1 subject was able to respond to only 5 cards. Those who had received 4 cards were given one of four combinations of a red-and-black card, 2 achromatic cards, and 1 color card (II, IV, VI, IX; III, V, I, VIII; II, VII, IV, X; III, VI, I, IX). Inasmuch as we were interested in the interrater reliability of thought disorder scores generated from the TDI, the variation in the number of cards administered was not considered to be a relevant influence, because all raters scored

the same protocols. Administration of each Rorschach protocol followed the procedures described by Rapaport et al. (1968). All sessions were audiotaped and subsequently transcribed verbatim. The protocols were scored independently for thought disorder according to the TDI manual (Solovay et al., 1986) by four teams of raters, all of whom had been trained in use of the TDI in the McLean Hospital Laboratory of Psychology and had at least 1 year of experience in using the TDI. Three of the four rating teams consisted of 2 to 5 scorers, and one rating team consisted of a single individual with similar training and experience. No attempt was made to achieve equal numbers of scorers on each rating team because these rating teams reflected the naturally occurring combinations of scorers in the collaborating laboratories. The teams scored the protocols at three sites: McLean Hospital (where two teams were based); the Brockton Veterans Administration Hospital; and Hillside Hospital, a division of Long Island Jewish Medical Center. Where two or more raters formed a team, the protocol was read aloud, and when an instance of thought disorder was recognized by any member, a short discussion of its scorability followed and a consensus scoring decision was made.

Statistical Analyses

Intraclass correlation coefficients (ICCs) were calculated to assess the degree of association among the four rating teams for the following variables: (a) Total TDI score, which yields a quantitative index of the amount of thought disorder present; (b) severity levels (.25, .50, .75, 1.0); and (c) scores for the qualitative categories as represented in five post hoc factors (Irrelevant Intrusions, Combinatory Thinking, Fluid Thinking, Confusion, and Idiosyncratic Verbalization) that have been found to discriminate among diagnostic groups (Holzman, Solovay, & Shenton, 1985; Shenton, Solovay, & Holzman, 1987; Solovay, Shenton, & Holzman, 1987).

For these same three variables (total TDI score, severity levels, and post hoc factors), the differences among rating teams were tested by repeated-measures analyses of variance, and contrasts between teams were tested with a Bonferroni correction. Because of the skewed distribution of total TDI scores, log transformations were performed to normalize the distributions. Log-transformed scores were used in the calculation of ICCs. The transformed scores yield more accurate estimates of reliability inasmuch as they reduce the effects of extreme scores. The ANOVAs of severity levels and qualitative factors used the proportion of thought-disordered instances that were assigned to each level or factor. This proportion represents the number of times a rating team assigned scores at a particular level or factor for each subject, divided by the total number of all thought-disordered responses given by that rating team for that subject. These proportions were then averaged for each level and for each factor for all 20 protocols.

Results

Total TDI Scores

As shown in Table 2, the ICC for total TDI scores among the four rating teams is .74. Table 2 also presents the Spearman rank correlations among the six possible pairs of rating teams. These correlations range from .80 to .90, and all are statistically significant (*p* < .01).

Table 3 presents the means, standard deviations, and ranges for the total TDI scores of each rating team. A repeated-measures ANOVA yielded an overall significant effect of rating teams, $F(3,19) = 24.57$, $p < .0001$. Paired contrasts showed significant differences in mean total TDI score for four of the

Table 2
Interrater Reliability for Total TDI Score Among Four Rating Teams for 20 Protocols

Paired rating teams	Spearman rank order correlations ($p < .01$)	Intraclass correlation
1–2	.88	.74
1–3	.90	
1–4	.80	
2–3	.81	
2–4	.88	
3–4	.81	

six possible pairings of rating teams, $p < .05$ with Bonferroni correction. Inspection of Table 3 indicates that some teams detected more thought disorder than did other teams. For example, Team 3 scored the most thought disorder, and Team 4 the least. The rank ordering of the amount of thought disorder in the 20 subjects by each rating team was nevertheless highly consistent, as indicated by the high Spearman coefficients (Table 2).

Severity Levels

Table 4 presents the overall ICCs for the severity levels as well as the Spearman rank order correlations among all combinations of rating teams. Only the data for the .25, .50, and .75 levels are presented because there were too few instances of scores at the 1.0 level for reliability assessments to be made. The ICCs ranged from .72 to .77 and were all highly acceptable. The Spearman coefficients ranged from .86 to .93 for the .25 level (all significant at $p < .01$). For the .50 level, the Spearman coefficients ranged from .50 to .75 (all significant at $p < .02$ or $p < .01$). For the .75 level, the Spearman coefficients ranged from .50 to 1.0 (all significant at $p < .02$ or $p < .01$).

Table 5 presents the means and standard deviations of the proportion of all thought-disordered responses that were assigned at each severity level by the four rating teams. A repeated-measures ANOVA yielded no significant differences among the rating teams in the proportion of scores assigned at the .25 level, $F(3,19) = 1.16$; the .50 level, $F(3,19) = 2.43$; or the .75 level of severity, $F(3,19) = 1.58$. Inspection of Table 5 indicates that each rating team scored a similar proportion of thought-disordered responses at each severity level, that .25-

Table 3
Mean Total Thought Disorder Index (TDI) Scores With Their Standard Deviations and Ranges of 20 Protocols Scored by Four Rating Teams

Rating team	Mean total TDI score	SD	Range
1	35.25	50.20	2.27–177.78
2	22.78	30.06	0.00–113.78
3	37.92	47.29	3.57–180.68
4	18.29	29.15	0.00–107.95

Table 4
Interrater Reliability for Level of Severity of Thought Disorder Among Four Rating Teams for 20 Protocols

Level of severity	Paired rating teams	Spearman rank order correlations	Intraclass correlations
.25	1–2	.91*	.77
	1–3	.93*	
	1–4	.89*	
	2–3	.87*	
	2–4	.86*	
	3–4	.86*	
.50	1–2	.72*	.72
	1–3	.56*	
	1–4	.75*	
	2–3	.50**	
	2–4	.59*	
	3–4	.64*	
.75	1–2	.54*	.77
	1–3	.72*	
	1–4	.54*	
	2–3	.50**	
	2–4	1.00*	
	3–4	.50*	

* $p < .01$. ** $p < .02$.

level responses were the most frequently scored by each team, and that .75-level responses were the least prevalent. Contrasts between teams, tested with Bonferroni correction (requiring $p = .008$), detected no significant differences for the three severity levels among the six possible pairings of rating teams.

Qualitative Factors

Table 6 presents the ICCs and the Spearman rank order correlation coefficients for the three factors on which there were a sufficient number of responses to calculate reliabilities. There were too few instances of scores in the Fluid Thinking and Confusion factors, and therefore reliability was calculated only for the Irrelevant Intrusions, Combinatory Thinking, and

Table 5
Means and Standard Deviations of Proportion Scores for Severity Levels Among Four Rating Teams for 20 Protocols

	.25 level	.50 level	.75 level
Team 1			
M	.88	.09	.03
SD	.19	.13	.07
Team 2			
M	.84	.09	.02
SD	.28	.18	.06
Team 3			
M	.79	.17	.04
SD	.19	.16	.07
Team 4			
M	.87	.06	.02
SD	.26	.11	.06

Table 6

Interrater Reliability for the Post Hoc Qualitative Factors Among Four Rating Teams for 20 Protocols

Post hoc factors	Paired rating teams	Spearman rank order correlations	Intraclass correlations
Irrelevant	1–2	.85*	.86
Intrusions	1–3	.88*	
	1–4	.75*	
	2–3	.79*	
	2–4	.80*	
	3–4	.93*	
Combinatory	1–2	.78*	.76
Thinking	1–3	.78*	
	1–4	.80*	
	2–3	.75*	
	2–4	.85*	
	3–4	.72*	
Idiosyncratic	1–2	.55*	.58
Verbalizations	1–3	.59*	
	1–4	.71*	
	2–3	.52**	
	2–4	.62*	
	3–4	.62*	

* $p < .01$. ** $p < .02$.

Idiosyncratic Verbalizations factors. ICCs for all rating teams averaged are .86 for Irrelevant Intrusions, .76 for Combinatory Thinking, and .58 for Idiosyncratic Verbalizations. Comparisons of paired rating teams yielded highly acceptable ICCs for Irrelevant Intrusions and Combinatory Thinking, and borderline values for Idiosyncratic Verbalizations. For Irrelevant Intrusions, Spearman coefficients ranged from .75 to .93 (all significant at $p < .01$). For Combinatory Thinking, Spearman coefficients ranged from .72 to .85 (all significant at $p < .01$). For Idiosyncratic Verbalizations, Spearman coefficients ranged from .52 to .71 (all significant at $p < .02$ or $p < .01$).

Table 7 presents the means and standard deviations for the proportion of all thought-disordered responses that were as-

signed to each of the five post hoc factors for each of the four rating teams. A repeated-measures ANOVA yielded significant group differences for the factors Combinatory Thinking, $F(3,19) = 4.10$, $p < .05$, and Confusion, $F(3,19) = 2.86$, $p < .05$. No significant differences were found for the Irrelevant Intrusions, Fluid Thinking, or Idiosyncratic Verbalizations factors. Contrasts among all possible combinations of rating teams detected only one statistically significant difference. Rating Teams 2 and 4 differed significantly ($p < .05$ with the Bonferroni correction) in the proportion of thought-disordered responses scored on the Confusion factor, a factor on which a very small proportion of responses was scored by any rating team.

Discussion

A comparison of four rating teams who independently scored thought disorder in 20 Rorschach protocols using the TDI yielded highly acceptable reliability coefficients for total TDI scores, replicating the levels of interrater reliability previously reported for two raters for total TDI (Johnston & Holzman, 1979; Solovay et al., 1986). High reliabilities were also obtained for severity levels and post hoc factors. The rating teams also showed very high agreement in the proportion of thought disorder scored for each factor.

Idiosyncratic Verbalizations, Irrelevant Intrusions, and Combinatory Thinking were the most frequently scored factors by all groups. Fluid Thinking and Confusion, in contrast, occur only rarely. With respect to the severity levels, very acceptable agreement was found for the three levels of severity tested. A review of the proportion of scores at these three severity levels shows that individual scoring groups agree well about the levels of scores.

Although the ICCs for the post hoc factors were well within the acceptable range, interrater reliability between pairs of rating teams, assessed by the Spearman coefficient, is variable. The factor scores are, nevertheless, very comparable across rating teams, indicating that the rating teams detected similar

Table 7

Means and Standard Deviations of Proportion Scores for Post Hoc Factors Among Four Rating Teams for 20 Protocols

	Irrelevant intrusions	Combinatory thinking	Fluid thinking	Confusion	Idiosyncratic verbalizations
Team 1					
M	.11	.12	.00	.06	.36
SD	.15	.19	.00	.11	.34
Team 2					
M	.18	.16	.00	.07	.32
SD	.24	.23	.01	.10	.31
Team 3					
M	.11	.10	.02	.04	.36
SD	.13	.13	.08	.06	.28
Team 4					
M	.11	.21	.01	.01	.41
SD	.16	.31	.02	.03	.37

proportions of thought disturbances reflecting most of the factors.

Absolute amount of thought disorder was found to vary among rating teams. This variation appears to result from threshold differences for tagging thought disorder. Although teams agree well about whether or not a protocol is thought disordered, and about the relative ranking of amount of thought disorder, they may not agree about the exact quantity of thought disorder present.

An examination of the specific protocols that contributed most to variability in total TDI scores among the rating teams showed that they were the most disorganized protocols in this sample of 20 records. All rating teams agreed that they were the most thought disordered, and agreement about their relative ranking was excellent. The discrepancies reflected differences in the number of individual thought disorder scores assigned to these highly disorganized records. Many of these responses appeared to be loose, confused, absurd, and confabulated all at once. To maintain good reliability on these protocols, it is essential to identify and score the dominant process.

More than one category of thought disorder should be scored only when there are distinct and separate disruptions in the organization of thought in the same response. Although there are times, in protocols of very disorganized psychotic patients, when it is difficult to make each scoring decision with certainty, the elevated amount and severity of thought disorder in these protocols is nevertheless quite reliably detected.

In our experience, consensus scoring in groups of well-trained scorers can improve reliability. Periodic review and recalibration are recommended for identifying and correcting sources of scoring errors. Although all of our teams of raters shared the same training, they had worked independently and separately since their initial training and made no attempts to recalibrate their scoring prior to undertaking this study. Periodic recalibration helps to minimize different tolerance levels for deviance in verbalizations, particularly in instances of mild cognitive slippage, and enhances recognition of infrequently occurring kinds of thought disorder. Consistent adherence to the definitions in the scoring manual (Solovay et al., 1986) is essential in this regard.

References

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

Andreasen, N. C. (1979). Thought, language, and communication disorders: I. Clinical assessment, definiton of terms, and evaluation of their reliability. *Archives of General Psychiatry, 36,* 1315–1321.

Andreasen, N. C. (1986). Scale for the Assessment of Thought, Language, and Communication (TLC). *Schizophrenia Bulletin, 12,* 473–482.

Andreasen, N. C., & Powers, P. S. (1974). Overinclusive thinking in mania and schizophrenia. *British Journal of Psychiatry, 125,* 452–456.

Arboleda, C., & Holzman, P. S. (1985). Thought disorder in children at risk for psychosis. *Archives of General Psychiatry, 42,* 1004–1013.

Bleuler, E. (1924). *Textbook of psychiatry* (A. A. Brill, Trans.). New York: MacMillan.

Bleuler, E. (1950). *Dementia praecox or the group of schizophrenias.* New York: International Universities Press. (Original published 1911)

Cameron, N. (1944). Experimental analysis of schizophrenic thinking. In J. S. Kasanin (Ed.), *Language and thought in schizophrenia* (pp. 50–64). New York: Norton.

Carlson, G. A., & Goodwin, F. K. (1973). The stages of mania: A longitudinal analysis of the manic episode. *Archives of General Psychiatry, 28,* 221–228.

Carpenter, J. T., Coleman, M. J., Waternaux, C., Perry, J., Wong, H., O'Brian, C., & Holzman, P. S. (1993). The Thought Disorder Index: Short-form assessments. *Psychological Assessment, 5,* 75–80.

Clayton, P. J., Pitts, F. N., & Winokur, G. (1965). Affective disorder: IV. Mania. *Comprehensive Psychiatry, 6,* 313–322.

Daniels, E. Kestnbaum, Shenton, M. E., Holzman, P. S., Benowitz, L. I., Coleman, M., Levin, S., & Levine, D. (1988). Patterns of thought disorder associated with right cortical damage, schizophrenia, and mania: A comparative study. *The American Journal of Psychiatry, 145,* 944–949.

Edell, W. S. (1987). Role of structure in disordered thinking in borderline and schizophrenic disorders. *Journal of Personality Assessment, 51,* 23–41.

Gershon, S. N., Benson, D. F., & Frazier, S. H. (1974). Diagnosis: Schizophrenia versus posterior aphasia. *American Journal of Psychiatry, 134,* 966–969.

Gold, J. M., & Hurt, S. W. (1990). The effects of haloperidol on thought-disorder and IQ in schizophrenia. *Journal of Personality Assessment, 54,* 390–400.

Goldstein, K. (1944). Methodological approach to the study of schizophrenic thought disorder. In J. S. Kasanin (Ed.), *Language and thought in schizophrenia* (pp. 17–40). New York: Norton.

Haimo, S., & Holzman, P. S. (1979). Thought disorder in schizophrenics and normal controls: Social class and race differences. *Journal of Consulting Clinical Psychology, 47,* 963–967.

Harrow, M., & Quinlan, D. (1977). Is disordered thinking unique to schizophrenia? *Archives of General Psychiatry, 34,* 15–21.

Harrow, M., & Quinlan, D. (1985). *Disordered thinking and schizophrenic psychopathology.* New York: Gardner Press.

Holzman, P. S., Solovay, M. R., & Shenton, M. E. (1985). Thought disorder specificity in functional psychoses. In M. Alpert (Ed.), *Controversies in schizophrenia: Changes and constancies* (pp. 228–245). New York: Guilford Press.

Hurt, S. S., 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: Jossey-Bass.

Kraepelin, E. (1919). *Dementia praecox and paraphrenia* (R. M. Barclay, Trans.). Chicago: Chicago Medical Book. (Original published 1896)

Marengo, J. T., Harrow, M., Lanin-Kettering, I., & Wilson, A. (1986). Evaluating bizarre-idiosyncratic thinking. *Schizophrenia Bulletin, 12,* 497–509.

O'Connell, M., Cooper, S., Perry, J. C., & Hoke, L. (1989). The relationship between thought disorder and psychotic symptoms in borderline personality disorder. *The Journal of Nervous and Mental Disease, 177,* 273–278.

Rapaport, D., Gill, M. M., & Schafer, R. (1968). The Rorschach test. In R. Holt (Ed.), *Diagnostic psychological testing* (pp. 268–463). New York: International Universities Press. (Original published 1945)

Rorschach, H. (1942). *Psychodiagnostics.* New York: Grune & Stratton. (Original published 1921)

Shenton, M. E., Solovay, M. R., & Holzman, P. S. (1987). Comparative studies of thought disorder: II. Schizoaffective disorder. *Archives of General Psychiatry, 44,* 21–30.

Shenton, M. E., Solovay, M. R., Holzman, P. S., Coleman, M., & Gale, H. (1989). Thought disorder in the relatives of psychotic patients. *Archives of General Psychiatry, 46,* 897–901.

Solovay, M. R., Shenton, M. E., Gasperetti, C., Coleman, M., Kestnbaum, E., Carpenter, J. T., & Holzman, P. S. (1986). Scoring manual for the Thought Disorder Index (rev. version). *Schizophrenia Bulletin, 12,* 483–496.

Solovay, M. R., Shenton, M. E., & Holzman, P. S. (1987). Comparative studies of thought disorder: I. Mania and schizophrenia. *Archives of General Psychiatry, 44,* 13–20.

Spitzer, R., Williams, M., & Gibbon, M. (1988). *Structured clinical interview for DSM-III-R: Patient version (SCID-P) and nonpatient (SCID-NP) version.* New York: Biometrics Research Department, New York State Psychiatric Institute.

Spohn, H., Coyne, L., Larson, J., Mittleman, F., Spray, J., & Hayes, K. (1986). Episodic and residual thought pathology in chronic schizophrenics. *Schizophrenia Bulletin, 12,* 394–407.

Watkins, J. G., & Stauffacher, J. C. (1952). An index of pathological thinking in the Rorschach. *Journal of Projective Techniques, 16,* 276–286.

Six Editors Appointed, 1995–2000

The Publications and Communications Board of the American Psychological Association announces the appointment of six new editors for 6-year terms beginning in 1995. As of January 1, 1994, manuscripts should be directed as follows:

- For the *Journal of Abnormal Psychology*, submit manuscripts to Milton E. Strauss, PhD, Department of Psychology, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, Ohio 44106-7123.

- For the *Journal of Applied Psychology*, submit manuscripts to Philip Bobko, PhD, Editor, *Journal of Applied Psychology*, P.O. Box 130, Skillman, New Jersey 08558.

- For the *Journal of Comparative Psychology*, submit manuscripts to Charles T. Snowdon, PhD, Department of Psychology, University of Wisconsin, 1202 West Johnson Street, Madison, Wisconsin 53706-1696.

- For the Attitudes and Social Cognition section of the *Journal of Personality and Social Psychology*, submit manuscripts to Arie W. Kruglanski, PhD, Department of Psychology, University of Maryland, College Park, Maryland 20742.

- For *Professional Psychology: Research and Practice*, submit manuscripts to Patrick H. DeLeon, PhD, JD, Editor, *Professional Psychology: Research and Practice*, APA, Room 3084, 750 First Street, NE, Washington, DC 20002-4242.

- For *Psychological Review*, submit manuscripts to Robert A. Bjork, PhD, *Psychological Review*, Department of Psychology, University of California, Los Angeles, California 90024-1563.

Manuscript submission patterns make the precise date of completion of 1994 volumes uncertain. The current editors, Susan Mineka, PhD; Neal Schmitt, PhD; Gordon G. Gallup, PhD; Abraham Tesser, PhD; Ursula Delworth, PhD; and Walter Kintsch, PhD, respectively, will receive and consider manuscripts until December 31, 1993. Should any 1994 volumes be completed before that date, manuscripts will be redirected to the new editors for consideration in 1995 volumes.