

# Indices of Psychopathology in the Rorschachs of Boys With Severe Gender Identity Disorder: A Comparison With Normal Control Subjects

Steven Tuber  
*City University of New York*

Susan Coates  
*Roosevelt Hospital*

Recent clinical and empirical studies of boys with Gender Identity Disorder (GID) of Childhood suggest that severe behavioral disturbance exists beyond their feminine preoccupations. This study examines the thought organization and object representational paradigms of a sample of GID boys ( $n = 26$ ) in a blind comparison with a normal subgroup ( $n = 18$ ). Rorschach test protocols were used to compare the groups along these two more internal measures of personality organization. As hypothesized, the GIC group was found to have more pathological scores than the normal group on the Mutuality of Autonomy (MOA) Scale and a thought disorder hierarchy. The findings suggest that GID boys do not suffer solely from gender symptomatology but are disturbed in other aspects of their ego functioning as well.

Research on extreme boyhood femininity has demonstrated that this syndrome does not emerge *de novo* as an isolated disorder as has previously been thought, but is part of a pervasive psychiatric disorder (Bates, Bentler, & Thompson, 1973, 1979; Bates, Skilbeck, Smith, & Bentler, 1974; Bradley, Doering, Zucker, Finegan, & Gonda, 1980).

Stoller's (1975) original clinical investigation of these boys concluded that extreme boyhood femininity was a dysfunction specific only to gender identity and not part of a more global or pervasive disturbance. Stoller stated "These mothers do not cripple the development of ego functions in general or even body

ego, except in regard to femaleness. . . . None of these boys has shown the slightest evidence of psychosis or precursors of psychosis" (p. 54). This finding has not been corroborated by subsequent systematic research in this area. All investigators of this problem who have systematically evaluated the psychiatric status of these boys have found them as a group to be as disturbed behaviorally as boys referred for psychiatric help (Bates et al., 1973, 1974; Bradley et al., 1980). This applies to the group of boys who have the extreme version of the syndrome, and not boys who may display either isolated feminine behavior or those who display feminine behavior during a time-limited period, for example, in response to a life stress.

The *Diagnostic and Statistical Manual of Mental Disorders* (3rd ed., *DSM-III*; American Psychiatric Association, 1980) diagnostic criteria for gender identity disorder of childhood for boys is: A plus either B or C.

- A. A strongly and persistently stated desire to be a girl or insistence that he is a girl.
- B. A persistent repudiation of male anatomic structures as manifested by the assertion that:
  1. He will grow up to be a woman (not merely in role);
  2. That his penis or testes are disgusting or will disappear;
  3. That it would be better not to have a penis.
- C. Preoccupation with female stereotypical activities as manifested by a preference for either cross-dressing or simulating female attire, or by a compelling desire to participate in the games and practices of girls.

The onset of the disturbance must occur before puberty. These criteria have remained unchanged in the revised *DSM-III-R* (American Psychiatric Association, 1987).

Most researchers agree that these boys' gender identification includes elements of both masculinity and femininity. Although they present clinically with hyperfeminine stereotypical interests and behaviors, two important differences between gender-confused boys and normal girls are evident: (a) the boys' feminine play and interests are rigid caricatures of stereotypical female mannerisms; and (b) despite the intensity of these mannerisms, they are aware that they are biologically male. Whereas earlier research with these children has focused largely on descriptive/demographic aspects of the disorder, our study is part of a more recent effort to delineate both the behavioral and emotional correlates of this extreme effeminacy (Coates & Person, 1985).

Systematic research on the extremely feminine boy has produced a consistent picture. In their sample of 25 GID boys, Coates and Person (1985) found that 84% were behaviorally disturbed children in need of mental health services. In addition, 60% of these children had a *DSM-III* diagnosis of separation anxiety disorder. Many others had symptoms of separation anxiety but did not reach

the criteria for a *DSM-III* diagnosis. Other *DSM-III* disorders that emerged in this group were dysthymic disorder and overanxious disorder. Pervasive behavioral disturbances have also been identified in similar samples of others (Bates et al., 1973, 1974, 1979; Bradley et al., 1980).

In an attempt to amplify the findings of pervasive behavioral difficulties in these children, Tuber and Coates (1985) studied the internalized object relations of these boys. They reported a preliminary study in which they compared the object relations as manifested in the Rorschachs of GID boys to the nonclinical sample reported by Ames, Metraux, Rodell, and Walker (1974). Their major results can be summarized as follows:

1. Feminine boys had a preponderance of quasi-human percepts in comparison to fully human percepts.
2. When percepts had individuals in interaction, a preponderance of these interactions were malevolent.
3. Boundary disturbances, as indicated by disturbances in thought organization, occurred more frequently in feminine boys.
4. Last, overt gender confusion was seen in single percepts of humans. This occurred in three different ways: some single percepts combined female and male elements into a single response. Second, in some single percepts the gender of the person changed from one to another. Third, some boys had great difficulty making up their mind as to whether a single percept is either exclusively female or male.

Tuber and Coates (1985) concluded that "The prevalence of boundary disturbances and lapses of thought organization, combined with the predominance of malevolent object representations, indicate the presence of severe psychopathology" (p. 261). Tuber and Coates's work was an attempt to extend the findings of Coates and Person (1985) on behavioral problems of GID boys into the internal representational world of the child. Tuber and Coates's (1985) study was limited by its lack of access to the actual normative Rorschach protocols of Ames et al. (1974), and their cross-group comparisons were, therefore, based only on the summary measures reported by Ames et al. Differences in socioeconomic status or IQ between groups were also not discerned. It seemed most appropriate to view their study as a pilot investigation exploring potential intergroup distinctions. Our study attempts to place their Rorschach findings on firmer empirical grounds. To this end, we chose to compare the thought organizations and inner representational worlds of a larger group of boys meeting the *DSM-III* criteria for gender identity disorder with a more carefully selected group of normal subjects.

### HYPOTHESES

We hypothesized that the GID boys would have significantly more examples of disordered thinking on their Rorschach protocols than would a normal sample.

Similarly, we also hypothesized that the GID boys would reveal their psychopathology through impairments in the quality of their object representations. We specifically hypothesized that they would show a greater number of malevolent, engulfing object representational themes than a normal sample. We also hypothesized that their overall level of object representations would be more primitive than the normal sample.

## METHOD

### Subjects

Twenty-six subjects who met the *DSM-III* criteria for gender identity disorder of childhood comprise the gender-confused population of our study. The children had all been referred to an outpatient psychiatric clinic for the first time when their diagnostic evaluations were conducted. Assessment of these children's verbal and play interviews, along with an interview with their parents provided the data from which the *DSM-III* diagnosis was made. These interviews and diagnoses were conducted prior to the conception of this study. As in our prior research, the subjects were all boys ranging in age from 5 to 12. Their average age was 8.3 years, with a mean Wechsler Full Scale IQ of 102. As assessed by Hollingshead's (1975) "Four Factor Index of Social Class," this sample covered the range from lower to upper economic class. Ethnic patterns also varied and were in keeping with the hospital catchment area.

The normal control sample consisted of 18 male subjects who were derived from local public, parochial, and private schools and were paid for participation in the study. Subjects were included only if they met the following criteria: (a) they did not have a history of chronic medical difficulties; (b) they had not been previously referred to a psychiatric unit; (c) they did not have a history of major academic difficulties requiring social class placement; and (d) they did not have a history of suspension for antisocial acts. This sample was also recruited and interviewed (both parents and child) prior to the conception of this study. These criteria were also met by the GID sample, although only their diagnosis and not these four criteria were the basis for their inclusion in the study.

The two groups were compared on both their mean Wechsler Full Scale IQ scores ( $M$  gender = 102,  $SD$  = 7.41;  $M$  controls = 106.2,  $SD$  = 8.49) and chronological age ( $M$  gender = 8.3,  $SD$  = 3.22;  $M$  controls = 8.4,  $SD$  = 3.05). No significant differences between the two groups were found on these two measures.

Comparing the two groups on chronological age was particularly important given Ames et al.'s (1974) findings that "odd combination" responses (i.e., thought disorder responses) do occur in normal populations with children aged 4 to 5½ years, although they "virtually disappear" by age 6. A post hoc review of our two samples revealed 4 subjects in each group between the ages of 5 and 6,

suggesting that age range differences between the groups would not affect possible group differences on the thought disorder measure. All 44 subjects were derived from the same hospital catchment area and similarly reflected the ethnic and socioeconomic patterns of that catchment area.

### Procedures for Scoring Rorschach Protocols

The GID sample was given an entire psychological test battery as part of their psychological intake assessment at the Roosevelt Hospital. The entire psychological test battery included Wechsler Intelligence Scale for Children-Revised (WISC-R) or Wechsler Preschool and Primary Scale of Intelligence (WPPSI), Rorschach, Thematic Apperception Test, Sex-typed Animal Preference Test (Coates & Zucker, 1987), and the Draw a Person Test. Their verbatim Rorschach protocols, administered by clinical psychology interns according to the standard clinical method as described by Rapaport, Gill, and Schafer (1945), comprised the data through which an assessment of their thought organization and object representations would be derived. Because of our interest in fantasied representations of human and animal figures, Rorschach inquiry regarding these responses were modified. Those responses potentially scoreable for human, animal, or inanimate interaction were further delineated by inquiry that in a specific manner addressed the quality of the interaction. If a subject spontaneously described a movement response, but a Mutuality of Autonomy (MOA) score was still unclear, the examiner repeated the response in the inquiry but added the phrase "as if. . ." in an attempt to elicit further information in a reasonably nondirective manner (see Tuber, 1988, for further discussion of this method). The normal control sample was also administered the same battery of psychological tests and their Rorschachs were administered in an identical fashion to those of the clinical group. None of the 44 subjects had been previously tested. It is important to note that the administration of all 44 protocols was conducted prior to the conception of our study so that the administrators of the test batteries, including the Rorschach, were blind to both the purpose of the study and the group identity of the subjects.

The 44 Rorschach protocols were then disguised as to individual identity and diagnostic group and were blindly scored by two PhD-level psychologists. Following an assessment of the interrater reliability of these two raters, disputed scores were reconciled by discussion between the raters and by Tuber (the first author of this study). The raters were aware of the general area of interest of the study, but did not know that two different groups of children were being assessed. The raters of the protocols did not know the administrators of the protocols. In addition, although Tuber was obviously aware of the nature of the study, he was blind to the group identity of the subjects when he reconciled the disagreements between the raters.

Due to the ordinal nature of the Rorschach scales utilized in the study, the

two groups were compared with one another by means of the Mann-Whitney U Test. In addition, using the Shapiro-Wilkes Test for normality, the distribution of both thought disorder and MOA Scale scores were found to not be normally distributed,  $p < .05$ , (on both distributions) thus mandating the use of nonparametric tests.

### The Rorschach Scales

The delineation of psychopathological thinking, in which separate percepts become fused or overelaborated so that intense affects blur the distinction between fantasy and reality, has long been a focus of Rorschach investigators (B. Klopfer, Ainsworth, Holt, & W. Klopfer, 1954; Meyer & Caruth, 1965; Quirk, Quarrington, Neiger, & Slemon, 1962; Rapaport et al., 1945; Watkins & Stauffacher, 1952). These studies have been successful in differentiating varying levels of pathology in clinical populations. Blatt and Berman (1984) and Blatt and Ritzler (1974) elaborated on the original work of Rapaport et al. (1945) and described a variety of thinking impairments which they depicted as impairments in maintaining distinct boundaries between ideas, images, and affects. Blatt and Ritzler (1974), Blatt and Berman (1984), Goldfried, Stricker, and Weiner (1971), and Harder and Ritzler (1979), among others, showed that these examples of pathological thinking can successfully differentiate among subtly different diagnostic groups. Blatt's hierarchy of pathological thinking uses Rapaport et al.'s (1945) original description of three discrete types of thought disturbance: contamination, confabulation, and fabulized combination responses, as its starting point. The fabulized combination responses are then divided into two subgroups: a fabulized combination *regular* response occurs when the spatial contiguity of two or more percepts is taken as evidence of a natural relationship between them, although each distinct image retains its separateness. An example would be "A rabbit being supported by two worms" to Card X. A fabulized combination *serious* response also uses spatial contiguity to depict an unnatural relationship, but here the disturbance occurs within a single unit so that the integrity of both percepts is violated. An example would be "a penguin with a man's legs" to the upper red area of Card II. Contamination responses involve the apparently ego-syntonic fusion or merger of two independent percepts into a single, bizarre, and impossible unit, such as the example "a bloody island" to the middle detail of Card III.

Confabulation responses are instances where percepts become imbued with overelaborated, often bizarre, personalized meanings to the subject. An example would be "two masked men, they're messengers of death for me" to Card I. Blatt and Ritzler (1974) provided evidence suggesting that contamination, fabulized combination serious, confabulation, and fabulized combination regular responses are associated with more serious to milder forms of psychopathology, respectively. Contamination and fabulized combination serious responses are

considered to be failures in the boundary between self and other. Confabulation responses are deemed failures in the boundary between reality and fantasy, whereas fabulized combination regular responses are viewed as connoting a more general laxness of boundaries. Each of these four types of thinking disturbance is further refined by the inclusion in the hierarchy of "tendency" responses in which some type of "disclaimer" is described by the patient after the response is given, implying some maintenance of a critical distance between the subject and the inappropriateness of his response. For the purpose of this study, the subjects were compared on the total number of these pathological thinking responses described by Blatt and Ritzler (1974).

Subjects' object representation scores were assessed via the MOA Scale developed by Urist (1977) and Urist and Shill (1982). This 7-point scale places Rorschach object representation responses along a continuum ranging from mutual, reciprocal, empathic relatedness, to themes of malevolent, overwhelming engulfment, and destruction.

A scale score of 1, the highest level of object relatedness, reflects interacting figures in which reciprocity and mutual acknowledgement are the key elements of the response. An example would be "two people dancing, sticking their tongues out at each other" to Card II. A scale score of 2 depicts interaction of a parallel sort, no reciprocity is described but no violation of the integrity of either "participant" exists. An example would be "two ladies cooking something" to Card III. Scale scores of 3 and 4 both reveal an emerging loss of autonomy in the interaction between the figures. Both scores imply a need for another figure to permit a sense of structural cohesion to exist. A score of 3 is used when figures are described as leaning on each other and/or must be externally supported. A score of 4 is employed where one figure is only the reflection or shadow of another. An example of a 3 response would be "two dead trees leaning on each other" to Card IV. An example of a score of 4 would be "a double image of a woman" to Card III. Points 5, 6, and 7 reflect the increasing malevolence of one figure toward another, so that the autonomy of one or more figures is intentionally violated. A score of 5 is reserved for responses in which fighting is clearly one-sided and hurtful or where one figure is malevolently controlling or influencing another. Examples would be "two cats fighting, the red is their blood" to Card II or "puppets, like someone has them on strings" to Card III. A scale score of 6 extends the score 5's theme of malevolence to instances of active attacking, destroying, and dominating by one figure over another. An example would be "a leech sucking up the blood of that man" to the lower detail of Card X. Last, a score of 7 extends the malevolence of point 6 to the extent that figures are engulfed by overpowering forces, not mere figures. This "larger than life" destructiveness is exemplified by the response "something being consumed by fire, can't even see what it is, just a raging fire" to the top of Card IX.

A score is given to each Rorschach response that explicitly or implicitly describes an exchange or interaction between two or more animate and/or inanimate percepts. Its inclusion of animate and inanimate as well as human

movement responses has made it of particular usefulness in the assessment of children's object representations (Coates & Tuber, 1988; Ryan, Avery, & Grolnick, 1985; Tuber, 1981, 1983, 1988; Tuber & Coates, 1985). Subjects were compared on their mean object representation scores as derived from the MOA Scale. Inasmuch as previous research with a GID sample revealed a particular emphasis on malevolent and/or toxic representations (Tuber, 1985), the three groups were also compared on their total number of malevolent MOA scores (those scores rated as a 5, 6, or 7 on the MOA Scale).

Thus the 44 subjects were compared on two Rorschach indices of psychopathology: a thought disorder index and an object representation index. Two object representation scores, a mean overall score, and the total number of malevolent scores were used to compare the groups to one another, whereas the total number and different types of thought disorder responses were used to compare the degree of thinking disturbance in each of the three groups.

## RESULTS

### Reliability

We assessed interrater reliability by comparing the ratings of the two PhD-level psychologists on both the thought disorder hierarchy and the MOA Scale. Percentage of agreement levels were obtained for the Blatt and Ritzler (1974) scale (84%) and for the Urist scale (79% "exact hits" and 88% agreement within 1 scale point). These results suggest a solid degree of interrater reliability on these two measures.

We first compared the groups on their total number of Rorschach responses. No significant differences were found on their total number of responses ( $M$  gender = 18.12,  $SD$  = 6.78;  $M$  controls = 16.55,  $SD$  = 4.66,  $t$  = .84). These means were both comparable to those reported by Ames et al. (1974), whose subjects in this age range had an average of 16 responses per protocol. Following a method of controlling for Rorschach productivity devised by Kalter and Marsden (1970), the number of thought disorder and object relation responses were separately subtracted from the total number of responses and then compared. None of the comparisons between the two groups reached significance.

These findings suggest that Rorschach productivity did not significantly influence the patterning of either the thought disorder or object representation scores. No correction for Rorschach productivity was deemed necessary.

### Thought Disorder

Table 1 describes the distribution of thought disorder response across groups. The GID group had a median value of 5.5 thought disorder responses. The distribution of thought disorder scores were then compared between groups



TABLE 1  
Distribution of Thought Disorder Responses

<i>Number of Thought Disorder Responses</i>	<i>GID<sup>a</sup></i>	<i>Controls<sup>b</sup></i>
0	1	4
1	3	2
2	1	4
3	2	1
4	2	1
5	4	3
6	2	1
7	1	0
8	4	0
9	1	2
10	3	0
19	1	0
20	1	0

<sup>a</sup>*n* = 26. <sup>b</sup>*n* = 18.

using the Mann-Whitney Test. Because the total distribution of scores was not normally distributed, the Kruskal-Wallis Test (chi-square approximation) was employed. As hypothesized, the GID group was significantly more likely to have a larger number of disruptions in thinking as measured by their larger number of thought-disordered responses  $\chi^2(1, N = 44) = 5.65, p < .01$ . To rule out the possible influence of age on these results, the two groups were subdivided into a younger (ages 5 to 7, with 12 GID boys and 11 controls) and an older (ages 8 to 12, with 14 GID boys and 7 controls) age category. Within each age category, the GID boys had between 2 and 2½ times the median number of thought disorder responses as the control subjects, suggesting that the age of the subjects did not influence the group differences between the subjects. When the children were compared across the varying types of thought disordered responses, the GID boys had a greater median number of contamination, fabulized combination serious and regular responses, and confabulation responses. Only the confabulation ( $Z = 2.27, p = .01$ ) and fabulized combination regular ( $Z = 2.32, p = .01$ ) response category differences were significant.

### Mutuality of Autonomy

Table 2 compares the two groups on the total number of malevolent object representation scores, as measured by the MOA Scale. The GID group had, per subject, over 1½ times as many depictions of malevolent interactions per Rorschach protocol as the normal group. When the distribution of malevolent MOA scores were ranked across groups, the GID group, as hypothesized, had significantly more disrupted object relations scores (using the Kruskal-Wallis Test,  $\chi^2(1, N = 44) = 2.75, p < .05$ ). The two groups of children were again

TABLE 2  
Distribution of Malevolent MOA Scale Scores<sup>a</sup>

MOA Scale Score	GID <sup>b</sup>	Controls <sup>c</sup>
5	48	21
6	39	17
7	6	0

<sup>a</sup>Malevolent MOA Scale scores refer to scores of 5, 6, or 7 on the Mutuality of Autonomy Scale (Urist, 1977). <sup>b</sup>*n* = 26. <sup>c</sup>*n* = 18.

compared across the younger and older age categories. Within both age categories, the GID boys had between 1½ and 2 times the number of malevolent interactions per Rorschach protocol as the normal subjects. The age of the children did not appear to influence the results across groups on the MOA variable.

Table 3 compares the two groups on their overall distribution of MOA Scale scores. The GID group had a median score of 5 which is in the malevolent range, whereas the control group had a median MOA of 3. When the distributions of mean scores were combined and ranked across the two groups, the GID groups' ranked scores were, as hypothesized, more disturbed than the normal controls, but this intergroup difference did not reach the preestablished  $p < .05$  level (using the Kruskal-Wallis Test,  $\chi^2(1, N = 44) = 2.45, p < .06$ ). The two groups did not differ significantly on their total number of MOA Scale scores (using Kruskal-Wallis, chi-square approximation,  $\chi^2(1, N = 44) = .42$ ).

## DISCUSSION

The findings of this study suggest that GID boys, in relation to their behavioral symptomatology, manifest psychopathology on another more internalized dimension as well. Rorschach assessments of the quality of thinking disturbance and the range and modal manner of depicting object representations suggest that two central aspects of personality organization are abnormal in the GID

TABLE 3  
Distribution of MOA Scales Scores

MOA Scale Points	GID	Controls
1	3	6
2	43	33
3	13	10
4	4	5
5	48	21
6	39	17
7	6	0

boys. These two dimensions involve the organization of thought and the quality of object relations.

As hypothesized, the gender-confused group has a significantly greater number of both thought disorder responses and of malevolent object representation responses than did the normal subjects. They had greater difficulty maintaining intact boundaries between disparate ideas, images, and affects. Their percepts were more frequently confused, merged, and overelaborated. These results replicate the findings of our previous pilot study. The current results are now on firmer ground because a control group was used in this study and the Rorschachs were rated blindly.

The distribution of malevolent scores across both groups suggests important differences in the way these children may experience interpersonal interaction. The gender-confused boys were significantly more likely to depict interactions characterized by control, dominance, or violence. When interactive responses were given on the Rorschach, the psychological and/or physical existence of one or more of the figures represented was threatened.

These results are compatible with previous investigations that strongly support the notion that the behavior of GID boys, beyond their stereotypic feminine interests and mannerisms, is disturbed as well. These studies report that GID boys suffer significantly disrupted peer relationships, display a greater degree of behavioral disturbances than normal subjects, and evidence both depressive and separation anxious symptomatology.

The etiology of severe boyhood femininity is presently not determined. Of the theorists who have speculated about etiological influences, the most influential has probably been Stoller (1975). As we described earlier, he hypothesized that the psychopathology in GID boys is specific only to gender identity and not part of a more pervasive ego dysfunction. He also hypothesized five major sources of influence for the etiology of extreme boyhood femininity:

1. Maternal depression.
2. Sexual identity problems in the mother.
3. Paternal distance or absence.
4. Striking beauty in the boy.
5. Blissful symbiosis between the mother and son.

The results of this investigation are incompatible with Stoller's two major hypotheses. We find no evidence for Stoller's blissful symbiosis hypothesis. We suggest that if a period of blissful symbiosis did serve as the principle organizer of the gender-confused boys' experience, then we would expect plentiful examples of benign reciprocity and a dearth of malevolent percepts in their Rorschach representations of animate and inanimate interaction.

The Rorschach data are also compatible with our inability to validate the blissful symbiosis hypotheses in clinical interviews (Coates, 1985; Coates &

Person, 1985; Coates & Zucker, 1987). The pervasive disturbances of cognition and object relations suggest that GID is part of a more global syndrome and not, as Stoller proposed as his other major hypothesis, an isolated symptom.

### ACKNOWLEDGMENTS

This research was supported in part by National Institute of Health Biomedical Research Support Grant 1-S07-RR05840 to Dr. Susan Coates.

Special thanks are due to Dr. John Fogelman, Director of Child and Adolescent Psychiatry at Roosevelt Hospital, for his generous support of this project. We particularly thank Drs. Robert Sherman, Rodger Goddard, Elise Lewin, Ellen Taylor, and Mark Glassman for their important contributions to this research effort. We also appreciate the assistance of several research staff members, including Penny Donnenfeld and Drs. Nancy Schultz and Toby Miroff.

### REFERENCES

- American Psychiatric Association. (1980). *Diagnostic and Statistical Manual of Mental Disorders* (3rd ed.). Washington, DC: Author.
- American Psychiatric Association. (1987). *Diagnostic and Statistical Manual of Mental Disorders* (3rd ed., rev.). Washington, DC: Author.
- Ames, L. B., Metraux, R., Rodell, J. L., & Walker, R. N. (1974). *Child Rorschach responses*. New York: Brunner/Mazel.
- Bates, J. E., Bentler, P. M., & Thompson, S. K. (1973). Measurement of deviant gender development in boys. *Child Development*, 44, 591-598.
- Bates, J. E., Bentler, P. M., & Thompson, S. K. (1979). Gender-deviant boys compared with normal and clinical control boys. *Journal of Abnormal Child Psychology*, 7, 243-259.
- Bates, J. E., Skilbeck, W., Smith, K. V. R., & Bentler, P. M. (1974). Gender role abnormalities in boys: An analysis of clinical ratings. *Journal of Abnormal Child Psychology*, 2, 1-16.
- Blatt, S. J., & Berman, W. M. (1984). A methodology for the use of the Rorschach in clinical research. *Journal of Personality Assessment*, 48, 226-239.
- Blatt, S. J., & Ritzler, B. (1974). Thought disorder and boundary disturbance in psychosis. *Journal of Consulting and Clinical Psychology*, 42(3), 370-381.
- Bradley, S. J., Doering, R. W., Zucker, K. J., Finegan, J. K., & Gonda, G. M. (1980). Assessment of the gender disturbed child: A comparison to sibling and psychiatric controls. In J. Samson (Ed.), *Childhood and sexuality* (pp. 554-568). Montreal: Editions Études Vivantes.
- Coates, S. (1985). Extreme boyhood femininity: Overview and new research findings. In Z. DeFries, R. Friedman, & R. Corn (Eds.), *Sexuality: New perspectives* (pp. 101-124). Westport, CT: Greenwood Press.
- Coates, S., & Person, E. (1985). Extreme boyhood femininity: Isolated finding or pervasive disorder? *Journal of the American Academy of Child Psychiatry*, 24(6), 702-709.
- Coates, S., & Tuber, S. (1988). Object relations psychopathology in gender confused boys. In H. Lerner & P. Lerner (Eds.), *Primitive mental states on the Rorschach* (pp. 647-654). New York: International Universities Press.
- Coates, S., & Zucker, K. (1987). Gender identity disorders. In C. Kestenbaum & D. Williams (Eds.), *The clinical assessment of children and adolescents* (pp. 893-914). New York: New York University

- Press.
- Goldfried, M., Stricker, G., & Weiner, I. (1971). *Rorschach handbook of clinical and research applications*. New York: Prentice-Hall.
- Harder, D., & Ritzler, B. (1979). A comparison of Rorschach development level and form-level systems as indicators of psychosis. *Journal of Personality Assessment*, 43, 347-354.
- Hollingshead, A. B. (1975). *Four factor index of social status*. Unpublished manuscript, Yale University, Department of Sociology, New Haven, CT.
- Kalter, N., & Marsden, G. (1970). Response productivity in Rorschach research. *Journal of Personality Assessment*, 34, 10-15.
- Klopfer, B., Ainsworth, M., Holt, R., & Klopfer, W. (1954). *Developments in the Rorschach technique* (Vol. 1). New York: Harcourt.
- Meyer, M., & Caruth, E. (1965). Rorschach indicators of ego processes. *Journal of Personality Assessment*, 29, 200-218.
- Quirk, D., Quarrington, M., Neiger, S., & Slemon, A. (1962). The performance of acute psychotic patients on the index of pathological thinking and on selected signs of idiosyncrasy on the Rorschach. *Journal of Personality Assessment*, 26, 431-441.
- Rapaport, D., Gill, M. M., & Schafer, R. (1945). *Diagnostic psychological testing*. Chicago: Yearbook Publishers.
- Ryan, R., Avery, R., & Grolnick, W. (1985). A Rorschach assessment of children's mutuality of autonomy. *Journal of Personality Assessment*, 49, 6-12.
- Stoller, R. J. (1975). *Sex and gender* (Vol. 2). *The transsexual experiment*. New York: Aronson.
- Tuber, S. (1981). *Children's Rorschachs as predictors of their later adjustment*. Unpublished doctoral dissertation, University of Michigan, Ann Arbor.
- Tuber, S. (1983). Children's Rorschach scores as predictors of later adjustment. *Journal of Consulting and Clinical Psychology*, 51, 379-385.
- Tuber, S. (1985, April). *An empirical study of the object representations of extremely feminine boys*. Paper presented at the meeting of the American Psychology Association, Division of Psychoanalysis, New York.
- Tuber, S. (1988). An extension of the mutuality of autonomy scale in the assessment of children's Rorschachs. In H. Lerner & P. Lerner (Eds.), *Primitive mental states on the Rorschach* (pp. 655-664). New York: International Universities Press.
- Tuber, S., & Coates, S. (1985). Interpersonal phenomena in the Rorschachs of extremely feminine boys. *Psychoanalytic Psychology*, 2, 251-265.
- Urist, J. (1977). The Rorschach test and the assessment of object relations. *Journal of Personality Assessment*, 41, 3-9.
- Urist, J., & Shill, M. (1982). Validity of the Rorschach mutuality of autonomy scale: a replication using excerpted responses. *Journal of Personality Assessment*, 46, 450-454.
- Watkins, J., & Stauffacher, J. (1952). An index of pathological thinking on the Rorschach. *Journal of Personality Assessment*, 16, 276-286.

Steven Tuber  
 Department of Psychology  
 City University of New York  
 New York, NY 10031

Received July 23, 1987  
 Revised December 14, 1987