The Psychological Assessment of Impulsivity: A Review

Peter Oas

Journal of Psychoeducational Assessment 1985; 3; 141
DOI: 10.1177/073428298500300205

Journal of Psychoeducational Assessment
1985, *3,* 141-156

THE PSYCHOLOGICAL ASSESSMENT OF IMPULSIVITY: A REVIEW

Peter Oas
California School of Professional Psychology, San Diego

A review of the clinical assessment of impulsivity with common psychological tests is presented. Although the term *impulsivity* is well known in the field of psychology and psychiatry, there are often misconceptions of the definition of the term and related psychometric aspects on both the empirical and clinical levels. Proposed is a definition that is in agreement with the different perspectives of the concept such that impulsivity is viewed as disordered behavior occurring with little or no premeditation or psychological capacity for delay. The assessment of impulsivity with the Bender Gestalt and Draw-A-Person tests, Wechsler scales, Rorschach, Self-Report, and Behavior Rating Scales cited in *Psychological Abstracts* and *Dissertation Abstracts* through 1984 is reviewed. It is concluded that the validity of these tests for assessing impulsivity has yet to be shown due to methodological and conceptual problems.

In contrast to the popularity of the *concept* of impulsivity as central to the conceptualization of the dynamics of various disorders, the clinical *assessment* of impulsivity has been slowly developing since the turn of the century. In defining the term *impulsivity,* there has been great disparity on all conceptual levels including etiologic, dynamic, and descriptive, leading to a great disparity in methods of psychologically assessing "impulsive behavior." Typically, impulsivity has been assessed by observing the presence of numerous "signs of impulsivity" emanating from a number of psychological tests in a standard psychological test battery. Evidence for impulsivity comes from signs on tests such as the Rorschach, WAIS-R, and Bender Gestalt and Draw-A-Person tests.

As indicated by Shapiro (1977), the Rorschach, which was first described in 1921, could be scored by five separate scoring systems developed between 1936 and 1945, all of which contained at least one score to measure impulsivity or impulse control. Typically, impulsivity was measured by the subject's reaction times to the blots or the prevalence of color responses. The fast speed of the subject's response was assumed to indicate a short-circuiting of analytic or reflective thought processes (Rapaport, Gill, & Schafer, 1968).

The Porteus Maze Test was introduced as a test of "feeblemindedness" in 1913 and eventually expanded by the author in 1942 to measure impulsivity in delinquents (Porteus, 1965). The scoring system of the test was based on the subject's

The author wishes to acknowledge the help of Alan S. Kaufman of the California School of Professional Psychology, San Diego, for his reviews of the manscript.
Address reprint requests to Dr. Peter T. Oas, 5208 Hahn Avenue, Fairborn, OH 45324.

0734-2829/85/030141$03.00/0

lack of attention to detail, carelessness, and lack of planning, depicted by errors in solving mazes (Porteus, 1965). Porteus's focus on lack of planning as a cognitive or affective deficit paralleled the shift in psychoanalytic theory from seeing impulsivity as a heightened drive state to explaining impulsivity as a "neurotic style" characterized by ego deficits such as poor judgment and inability to plan (Shapiro, 1965). The bulk of research on impulsivity then shifted from intrapsychic deficit models, or lack of "ego control," to emphasis on observable behaviors that defined impulsivity.

The most recent trend in the literature pertaining to the construct of impulsivity has emphasized cognitive differences of impulsive and nonimpulsive children (Kagan, Rosman, Day, Albert, & Phillips, 1964). In 1964, Jerome Kagan coined the expression "conceptual tempo" to refer to the cognitive dimension of reflection-impulsivity. Impulsivity, or cognitive tempo, was operationally defined as a subject's performance on the Matching Familiar Figures Test (MFFT). The MFFT is a visual, 12-item match-to-sample task that requires the subject to choose from an array of eight closely similar variants the one picture that is identical to a standard. Impulsive subjects are those who respond so quickly that they make errors in identification of the correctly matching picture. Latency to first response and total errors are recorded for each of the 12 items on the test.

The clinical validity of the MFFT has been demonstrated for children up to 12 years of age (Arizmendi, Paulsen, & Domino, 1981; Becker, Bender, & Morrison, 1978; Bentler & McClain, 1976; Egeland & Weinberg, 1976; Epstein, Cullivan, & Lloyd, 1977; Messer, 1976). Its usefulness with older subjects is currently being examined, as is a modified version of the MFFT, the MFF20 (Messer & Brodzinski, 1981; Cairns & Cammock, 1978).

A recent review of the literature on the conceptualization of "impulsivity" has provided a definition of the term that is in agreement with all theoretical viewpoints (Oas, 1984b). Impulsivity is defined as "behavior that is socially inappropriate or maladaptive and is quickly emitted without forethought." This definition takes into account that impulsive behavior is cognitively mediated and that it is maladaptive or inappropriate to the situation, as opposed to behavior that is more positively described as intuitive, instinctive, or spontaneous. The definition rightly makes no pretense as to etiologic factors because impulsive behavior can be induced both psychologically and physiologically (Monroe, 1970). However, it is one's definition of impulsivity that is particularly crucial to the psychological assessment of impulsivity and the validation of tests for assessing this behavior.

RATINGS OF OBSERVED IMPULSIVE BEHAVIOR

In studies that use some type of behavior rating scale as a concurrent measure, subjects are typically rated as impulsive on the MFFT, in conjunction with rating scales designed to assess hyperactivity or "acting out," not impulsivity per se. Typically used rating scales are those such as Conners's Teacher Rating Scale (Conners, 1969) or the Behavior Problem Checklist (Quay & Peterson, 1967). As a result, the construct validity of the MFFT has been firmly established for groups of children when the criteria are rating scales of *hyperactivity*. However, the construct

validity of the MFFT with adolescents and adults is not established, and validity has not been established with rating scales specific to impulsivity in clinical and many educational settings.

Recently, researchers have begun to question the use of the diagnosis of attention deficit disorder (ADD) in the third edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM III; APA, 1980) as a criterion measure of impulsivity in conjunction with the more traditional rating scales of hyperactivity (Lahey, Green, & Forehand, 1980; Levy & Hobbes, 1981; Ownby, 1983). There have been two studies that have shown that raters could accurately distinguish between impulsive behavior as defined by the ADD diagnosis as a criterion measure of impulsivity. Lahey et al. (1980) studied teachers' ratings and direct observational measures of 109 third-grade children. Factors related to ADD criteria and hyperactivity scores did not correlate, and the authors suggested that ADD criteria and ratings of hyperactivity be considered separately. King and Young (1982) studied the diagnosis of attention deficit disorder *with* hyperactivity (ADDH) as distinct from the diagnosis of attention deficit disorder *without* hyperactivity (ADD). With scores on the Conners rating scale as the criteria, subjects rated as ADD were not rated as hyperactive on the Conners scale, and subjects rated ADDH *were* rated as being hyperactive based on Conners's criteria for hyperactivity.

DSM III makes a clinical distinction among hyperactivity, impulsivity, and attention deficits. It seems reasonable that an accepted clinical standard of impulsivity, such as the diagnostic criterion of attention deficit disorder *without* hyperactivity, could be compared to the widely accepted test of impulsivity, the MFFT. In a doctoral dissertation, Ozawa (1980) developed a rating scale based on the attention deficit disorder diagnosis; ratings on this scale for 120 elementary school children correlated .21 with the MFFT ($p < .05$).Although this study was important in its attempt to validate the MFFT with clinical ratings of behavior, the correlation was low despite its statistical significance. This low relationship may have resulted because of the normality of the sample; that is, too few of the children may have been truly impulsive. Higher relationships between the MFFT and the DSM III criterion with hospitalized adolescents have been found in studies by Oas (1983, 1984a, 1984c). In these studies, samples of 20, 30, and 214 "normal" and hospitalized adolescents were tested on the MFFT and were rated on a Behavior Rating Scale (BRS) developed to include the DSM III diagnosis of attention deficit disorder *without* hyperactivity. Correlations between the BRS scores and MFFT scores ranged from .62 to .92 ($p < .01$), suggesting that both the MFFT and BRS could be viable criterion measures of impulsivity in populations of hospitalized adolescents and nonhospitalized adolescents.

In other studies recognizing the importance of obtaining observed ratings of impulsive behavior, teachers' ratings are typically recorded, but only a few behavior rating scales have been developed. One such scale is the Self-Control Rating Scale (Kendall & Wilcox, 1979), a 33-item scale designed to measure impulsive behaviors emitted by elementary school children in classroom-type settings. This scale has been shown to have sound psychometric properties and to be moderately correlated with MFFT scores, although its validity has yet to be confirmed. Another scale designed for use with children, called the Kansas Reflection-

Impulsivity Scale (Wright, 1973), also has not been extensively studied, although studies by McClanahan (1979) and Mistry (1975) have reported moderate validity and reliability coefficients of .40 to .52.

In developing behavior rating scales designed to assess impulsiveness, researchers have not been careful to isolate and use factors that differentiate impulsivity from other constructs and syndromes such as hyperactivity, delinquency, acting out, and aggression. Further construction of "impulsivity rating scales" necessitates the inclusion of items that are both valid and able to discriminate among these similar constructs and syndromes.

SELF-REPORT MEASURES

Self-report questionnaires are a commonly used technique for measuring impulsivity among adolescents and adults (Zuckerman, 1983). Most of these scales have been developed as part of more global measures of personality traits or dimensions, but a few have been designed specifically to measure impulsivity.

One such scale is the Impulsiveness Scale of the Eysenck Personality Inventory (Eysenck & Eysenck, 1977). This scale and modifications of it have been reported to have moderate to high reliability coefficients (Eysenck, 1981; Eysenck & Eysenck, 1977; Saklofske & Eysenck, 1983), although in none of these studies have there been reported validity coefficients with other criterion measures of impulsivity. In most studies listing correlations between impulsiveness as assessed by Eysenck's test and observed impulsive behavior, correlations are modest in some (Cairns & Harbison, 1975) and insignificant in most (Bentler & McClain, 1976; Edman, Schalling, & Levander, 1983; Glow, Lange, Glow, & Barnett, 1983; Gudjonsson, 1980). Not only is the validity of this scale suspect but it suffers from the problems inherent in the validity of the self-report of behavior itself. These same problems are apparent in research reports of other self-report measures of impulsivity (Hirschfield, 1965; Kipnis, 1971; Murray, 1938; Schalling, 1975).

Other self-report measures of impulsivity have been researched more extensively. The Barratt Impulsiveness Scale (BIS) was developed in 1965 (Barratt, 1965). In reviewing this instrument it is apparent that the validity of this scale as a measure of impulsiveness is still suspect. In Barratt's own words (Barratt & Patton, 1983):

> From our research interrelating the BIS with other measures of impulsivity, two general conclusions are obvious: (a) most questionnaire measures of impulsivity are significantly intercorrelated, usually at an average to high level, and (b) the questionnaire measures usually have low order and often insignificant correlations with the nonquestionnaire measures of impulsivity. (p. 86)

Although they believe this indicates that all techniques that purport to measure impulsivity are not measuring the same variable or construct, in my opinion it is more a difficulty with their instruments than a problem with trying to identify a proclivity toward impulsiveness.

Where the clinical importance of identifying impulsiveness is paramount, the value of an instrument lies in its ability to identify not only biological or cognitive aspects of impulsivity but observed impulsiveness as well. It is for this reason that although Zuckerman's Sensation Seeking Scale (Zuckerman, Kolin, Price, &

Zoob, 1964) has been designated an important measure of impulsiveness, it too must be shown to correlate significantly with other nonquestionnaire indices of impulsivity if it is to be ultimately useful.

IMPULSIVITY AND COMMON PSYCHOLOGICAL TESTS

A major gap in the research literature pertaining to the assessment of impulsivity now exists. This gap has to do not only with a lack of unity in developing rating scales and self-report measures of impulsivity but also with the *validity* of inferences of impulsivity made with the most common psychological tests. Measurement with clinical instruments such as the Rorschach, Wechsler scales, Porteus mazes, and especially graphomotor instruments such as the Bender Gestalt test or the Draw-A-Person test has infrequently been studied empirically and rarely validated. In a doctoral dissertation, Redmountain (1976) attempted to assess the validity of these instruments in measuring impulsivity based on the sign approach. He studied 522 children and adolescents and found that "with respect to the signs of impulsivity, no significant differences were found between subjects rated as impulsive and not impulsive. . . . In other words signs of impulsivity neither reflected actual behavior, nor were they consistently used by psychologists to identify or predict such behavior." Although there were methodological flaws in his study, he pointed out, significantly, that there is a paucity of research to validate the use of signs of impulsivity on these tests to infer impulsivity and that for the most part clinicians and researchers have been content with their own theoretical presuppositions.

A few theoretical reviews of the concept of impulsivity exist (Monroe, 1970; Oas, 1984b), but with the exception of Redmountain's study, no one has attempted to examine the available data to determine the validity of such commonly used tests for developing inferences of impulsivity as the Wechsler scales, Rorschach, DAP, or BG. The following is an attempt to provide a comprehensive examination of the research results of these tests since their development, with particular attention paid to methodological issues and the lack of appropriate criterion measures.

IMPULSIVITY AND THE DRAW-A-PERSON AND BENDER GESTALT TESTS

Drawing tests such as the Draw-A-Person (DAP; Goodenough, 1926; Harris, 1963) and the Bender Gestalt Test (BG; Bender, 1938) were cited by Lubin, Wallis, and Pain (1971) as two of the five most frequently used assessment devices by psychologists. Wade, Baker, Morton, & Baker (1978) list both tests in the top seven of clinical tests used most frequently by psychologists. These tests have been used frequently to study aspects of visual-motor integration on one extreme and to make assessments of personality dynamics or traits on the other.

A survey cited in Redmountain (1976) of 78 child guidance and community mental health centers revealed that the tests most often used by psychologists and other health professionals to diagnose impulsivity were the Wechsler Intelligence Scale for Children (WISC) and the Bender Gestalt and Draw-A-Person tests. The Bender Gestalt and Draw-A-Person tests have also been used frequently to assess

impulsivity by such noted researchers as Bender (1938), Hammer (1965), Koppitz (1968), and Machover (1949). Many of their conclusions regarding the measurement of impulsivity on these tests have been based on the prevalence of various signs of impulsivity such as collision of figures or rounded angles (BG), lack of symmetry or line discontinuity (DAP), and poor quality in general. Some have even concluded that because of the lack of validity found in DAP and BG research, the use of these two instruments in any way is a violation of APA ethics (Martin, 1983). Yet none of these researchers have published acceptable empirical studies to assess the validity of their claims.

One problem is that research with the DAP and BG is methodologically poor. With respect to the DAP, Swenson (1968) and Roback (1968) cited numerous methodological flaws in research studies with this instrument. They stated that poor research methodology is the contributing factor related to almost 30 years of inconclusive evidence for the validity of inferences made using the DAP in clinical research. Their main criticism is that researchers have inappropriately used the sign approach to test hypotheses by ignoring a fundamental approach to inference making with projective psychological test data. Inferences (conclusions) are made not by comparing the *relationship* between signs on these tests but by measuring the frequencies of individual signs (Falk, 1981; Roback, 1968). Other criticisms of DAP research methodology include the use of broad diagnostic labels and heterogeneous groups of subjects (instead of specific behaviors and homogeneous groups of subjects), the low reliability of specific signs, lack of cross-validation procedures, and lack of controls for artistic quality, organicity, IQ, and age differences (Falk, 1981; Hammer, 1965; Roback, 1968; Sonnenblick, 1980; Swenson, 1968). Many of these same criticisms have been applied to BG research as well (Blackman & Goldstein, 1982; Buckley, 1978).

Research studies of these tests used to discriminate between impulsive and nonimpulsive subjects have the same methodological problems. In psychological journals, only six studies could be found that attempted to look at the relationship between accepted criterion measures of impulsivity (such as the MFFT) and BG or DAP test data. Studies looking at the relationship between the BG and impulsivity (Blaha, Fawaz, & Wallbrown, 1979; Brannigan, Barone, & Margolis, 1978; Wallbrown, Wirth, & Engin, 1975) and the relationship between the DAP and impulsivity (Blaha et al., 1979; Brannigan et al., 1979; Schecter, 1981) have yielded inconclusive evidence, suggesting that there may not be any relationship between criterion measures of impulsive behavior and drawings. Furthermore, these studies have all used children as subjects, and most have methodological weaknesses such as small sample size and lack of controls for age differences, IQ differences, and level of pathology.

A summary of studies on the DAP is presented in Table 1. These studies were evaluated along several important methodological lines trying to show the vast discrepancy in procedural differences. In general, the results of studies attempting to determine if signs of impulsivity on the DAP test are valid have yielded insignificant results. The results, whether achieved by correlational or group comparison methods, must be considered lightly due to the unsystematized use of criterion measures, the varied ages of subjects, the lack of controls for IQ or perceptual-motor ability, and the correlational nature of the statistical comparisons.

Table 1
Methodological Differences in Studies of the Validity of the Draw-A-Person Test

Study	Subject Comparison Groups	Impulsivity Criterion	Dependent Variables	Uncontrolled Variables	Results
Redmountain (1976)	Correlations among scores of 522 clinic children aged 6–18 years	Checklist scores of "impulsive acts"	Various DAP variables	All[a]	Not significant
Blaha, Fawaz, & Wallbrown (1979)	Correlations among scores of 74 first-grade children	MFFT	Completion time and Harris (1963) DAP score	All but IQ	Not significant
Branningan, Margolis, & Moran (1979)	28 impulsive vs. 28 reflective kindergarten children	MFFT	Koppitz (1968) DAP emotional indicators"	All	Not significant
Schecter (1981)	Correlations among scores of 51 behavior-disordered male adolescents	MFFT, freedom distraction index (Cohen, 1957), behavior observation	Handler (1965) DAP "anxiety indicators"	All	Not significant
Oas (1984b)	73 impulsive vs. 69 nonimpulsive hospitalized and nonhospitalized adolescents	MFFT, behavior ratings	DAP Impulsivity score	—	Significant at $p < .001$

[a]IQ, age, drawing ability, visual motor ability, and gender.

The evidence presented in the foregoing description of empirical studies is also contradictory to results of studies that have examined MFFT scores and psychomotor performance on other measures such as the Draw-A-Line (DAL; Harrison & Nadelman, 1972) or the Porteus Mazes Test (PMT; Homatidis & Konstantareas, 1981; Paulsen & Johnson, 1980). In these studies, MFFT performance correlated significantly with drawing a straight line or completing mazes. This empirical evidence of drawing performance and impulsivity is also in direct contradiction to Kagan's (1965) hypothesis that conceptual tempo is measured only in situations with "high response uncertainty." It seems reasonable to infer that increased response uncertainty is involved in the completion of the DAP test more so than in the DAL or PMT and that DAP performance would be more apt to correlate with MFFT scores if Kagan's hypothesis were true, since drawing any person involves more reflectiveness than simply drawing a line on a sheet of paper. Blaha et al. (1979) also hypothesized that impulsivity is measured in situations with high response uncertainty and concluded that this relationship was true based on the fact that Bender Gestalt (BG) test scores correlated with MFFT performance but DAP scores did not. Certainly one could argue that there is more "response uncertainty" involved in drawing a human figure than in reproducing geometric designs.

There has been considerably more research looking at the validity of BG test scores as indicants of impulsivity. Yet the validity of signs of impulsivity on the BG is still undetermined primarily because of methodological and theoretical flaws

Table 2
Methodological Differences in Studies of the Validity of the Bender Gestalt Test

Study	Subject Comparison Groups	Impulsivity Criterion	Dependent Variables	Uncontrolled Variables	Results
Donnelly & Murphy (1974)	37 manic-depressive vs. 30 major depressed adults	DSM II diagnosis	BG "sequencing" and "placement" variables	All[a]	Significant at $p < .01$
Wallbrown & Wallbrown (1975)	Correlations among scores of 76 first-grade children	MFFT	BG completion time and Koppitz (1975) error scores	Drawing ability, visual motor ability, gender	Not significant
Wallbrown, Wirth, & Engin (1975)	Correlations among scores of 48 first-grade children	MFFT	BG completion time and Koppitz (1975) error scores	Drawing ability, visual motor ability, gender	Not significant
Brannigan, Barone, & Margolis (1978)	22 impulsive vs. 18 reflective 10-year-old schoolchildren	MFFT	Hutt (1969) BG variables	All but age	Significant at $p < .01$
Blaha, Fawaz, & Wallbrown (1979)	Correlations among scores of 74 first-grade black children	MFFT	BG completion time and Koppitz (1975) error scores	Drawing ability, visual motor ability, gender	Not significant
Trahan & Stricklin (1979)	49 impulsive vs. 44 nonimpulsive children aged 8–15 years	Behavior rating scale scores	Koppitz (1975) BG "emotional indicators"	IQ, drawing ability, visual motor ability	Not significant
Oas (1984b)	73 impulsive vs. 69 nonimpulsive hospitalized and nonhospitalized adolescents	MFFT, behavior ratings	BG Impulsivity score	—	Significant at $p < .001$

[a]IQ, age, drawing ability, visual motor ability, and gender.

similar to those inherent in DAP research, which undermine the validity of the procedures in the studies that have been done. These studies are presented in Table 2.

In addition to the methodological problems of research on impulsivity and drawing performance, no theoretical rationale is demonstrated in the choice of so-called signs of impulsivity on these tests. It is quite possible that there is nothing inherent in drawing per se to which impulsive and reflective individuals respond differentially. It may be that an impulsive individual would do poorly on any task and that something other than impulsivity—such as anxiety, intelligence, or motivation—is the operative factor.

A recent, well-controlled study on the DAP and BG has demonstrated that it may just be that impulsivity is an important factor in drawing performance (Oas, 1984c). In this study, 100 psychiatrically hospitalized and 114 nonhospitalized adolescents were designated as impulsive or nonimpulsive (reflective) based on two measures of impulsivity and one measure of impulsivity/delinquency: Kagan's Matching Familiar Figures test scores, ratings of observed impulsive behavior, and checklist scores of reported impulsive/delinquent acts. Correlations among these

three measures for both hospitalized and nonhospitalized groups ranged from .479 to .893. The impulsive and nonimpulsive subject groups were compared to determine the validity of so-called signs of impulsivity on the Draw-A-Person and Bender Gestalt tests. Signs of impulsivity and signs of nonimpulsivity on these two tests discriminated between impulsive and nonimpulsive adolescents with surprising accuracy. The results of this study also show that in research on the DAP and BG, impulsive drawing performance is related to a general inability to delay responses and that reflectiveness and cognitive mediation are important factors.

IMPULSIVITY AND THE RORSCHACH TEST

Many writers contend that the major criticisms of the Rorschach test as an instrument for predicting overt behavior center around problems of validation of test factors. Schafer (1949) clarified this argument by pointing out that interpretations rather than scores are the proper units for research on projective tests and stated that the actual need is for validations of these interpretations.

The definition of impulsivity based on Rorschach performance comes from the writings of Rapaport, Gill, and Schafer (1968) on the test. Their definition of impulsivity concurs with earlier psychoanalytic theory, which contends that ego function is a modulation of control and delay over internal drives or pressures via fantasy or cognition. The rationale for the predominance of color-dominated versus form-dominated responses on the Rorschach as being indicative of impulsivity comes from thinking that the perception of color is much easier and more readily available to the unspeculative mind than reflections on variation in possible form differences. The prevalence of color-dominated responses as indicative of impulsivity has received the most empirical support over the years and seems to have a moderate amount of validity, but only when clusters of scores, rather than any one particular score, are interpretated (Gardner, 1951; Holtzman, 1950; Mandeville, 1954; Ostrov, Offer, Marohn, & Rosenwein, 1972).

Studies of the relationship between observed ratings of impulsive behavior and the predominance of color responses have generally yielded moderately significant relationships (Gardner, 1951; Holtzman, 1950; Mandeville, 1954; Verrill, 1958). A group of studies by Singer and various colleagues (Singer, Meltzoff, & Goldman, 1952; Singer & Spohn, 1954; Singer & Herman, 1954; Singer, Wilensky, & McCraven, 1956) have revealed that subjects who produce more M responses tend to be rated as more inhibited and overcontrolled, yet there is no difference in motor inhibition or planning between those rated as inhibited and those rated as impulsive. The only study to find results discrepant to these was by Hardison (1964), who found no significant differences in color responses among inhibited and noninhibited army trainees.

Most studies on the Rorschach have used adults as subjects, unlike studies of other psychological tests, and there appears to be much room for validation of previous work on the Rorschach with children and adolescents. A recent study by Ikegami (1979) reported that increased numbers of W and color responses of 151 kindergartners were correlated significantly with MFFT scores. A small body of research on incarcerated adolescent delinquents has revealed moderate evidence

for a relationship between scores on such measures as the MFFT, behavior ratings, and self-report measures of impulsivity, with Rorschach performance (Ostrov et al., 1972), although the review of this literature by Ostrov et al. pointed out that results are mixed and depend on the samples and variables chosen. As with the research on the DAP and BG, methodological problems are apparent in this body of research as well. Criterion measures of impulsivity were suspect, as were a limited use of control groups or controls for age, gender, IQ, and so on.

IMPULSIVITY AND THE WECHSLER SCALES

Messer (1976) provided a review of the relationship between WISC and WISC-R performance and scores on the MFFT. In a summary of his findings he reported:

> When the content of an I.Q. test is primarily non-verbal and the format requires decision about alternatives (multiple choice), correlations of MFFT response time to I.Q. are higher than when the test calls for verbal responses, especially for items with minimal response uncertainty . . . (p. 1036)
>
> . . . The correlation of MFFT response time with I.Q. is small (.16); the correlation of MFFT errors with I.Q. is moderate (– .31). At the same time, there is considerable variability in the size of those relationships, stemming in part from the subject's age and nature of intelligence test employed. In children 5 years old and younger and on intelligence tests with format and content similar to the MFFT, the relation of response time to I.Q. is higher. In children over 5 and on verbally oriented I.Q. tests, the relation is lower. (p. 1047)

Since Messer's review, others have found impulsives to score lower than reflectives on measures of WISC or WISC-R IQ. Brannigan and Ash (1977) studied 41 reflective and 32 impulsive children aged 8 to 11. The mean Full Scale IQ scores were 88.6 for impulsives and 99.9 for reflectives. This difference was significant, as were the differences between both groups on the Information, Comprehension, Digit Span, and Picture Completion subtests, with impulsives scoring significantly ($p < .05$) lower.

Mollick and Messer (1978) studied 53 school girls with a mean age of 10 years. They found that MFFT errors correlated – .36 with WISC Full Scale IQ scores and – .35 with Performance IQ scores.

Jackson, Farley, Zimet, and Gottman (1979) studied 75 male and 26 female schoolchildren both with a mean age of 11.2. They found that "low impulsive" children consistently achieved higher WISC-R Verbal IQ, Performance IQ, and Full Scale IQ scores ($F = 4.45$, $df = 1/9$, $p < .03$; $F = 5.69$, $df = 1/9$, $p < .01$; $F = 6.78$, $df = 1/9$, $p < .01$) than "high impulsive" children.

Brannigan, Ash, and Margolis (1980) studied 36 reflective and 22 impulsive 8- to 11-year-old schoolchildren and found significant differences on WISC-R Attention-Concentration subtests ($t = 2.29$, $df = 56$, $p < .05$) and Visual Organization subtests ($t = 2.70$, $df = 56$, $p < .01$). Interestingly, they also found that impulsive and reflective children did not differ significantly on Full Scale, Performance Scale, or Verbal Scale IQ scores. This is the only study that failed to find significant differences on WISC-R IQ scores between impulsive and reflective subjects.

Finally, Finch, Spirito, and Brophy (1982) studied children aged 8 to 15 years. They also found that impulsives and reflectives differed significantly on WISC-R

Full Scale, Performance Scale, and Verbal Scale IQ scores ($F = 7.11$, $df = 41$, $p < .01$; $F = 7.70$, $df = 41$, $p < .01$; $F = 4.34$, $df = 41$, $p < .05$).

To date, no one has systematically studied performance on the WAIS-R, nor have there been studies of other age groups from adolescence and older.

A major problem with research on the Wechsler scales is that no one has thoroughly focused on a theoretical rationale for why impulsive individuals do better on performance items than on verbal items. How does impulsivity differentially relate to performance versus verbal abilities and why? Alan Kaufman (1984) has suggested that differences in performance by impulsive children may be due more to poor verbal skills related to poor academic achievement rather than better performance capabilities. In fact, one would expect the gap between Verbal and Performance IQ to *increase* with age as performance skills are acquired. Yet a review of research on DAP and BG performance points out performance *deficits* of impulsive children to those of reflective children. These issues have not been addressed primarily because of the haphazard rationale in choosing criteria for impulsiveness and making inferences of impulsiveness based on WISC-R Verbal and Performance scale differences. There is a need to determine specifically which factors of impulsivity are operative and the actual definition of impulsivity in these studies.

DISCUSSION

In reviewing the vast number of studies of impulsivity on psychological tests, one may correctly conclude that although this body of research is large, the evidence for what constitutes impulsive behavior on common psychological tests has yet to be determined.

Even though poor methodology is readily apparent upon closer examination of individual studies, the focus in applied clinical arenas tends toward a naive assumption of the acceptance of the validity of "signs of impulsivity" on the tests described in this review. Additionally, resources most readily available to the clinician, such as Ogden's (1981) and Gilbert's (1980) "cookbooks" for interpreting psychological test data may elicit a false sense of clinical acumen in the clinician by reporting "nonempirical" data.

In reviewing research on instruments commonly used by clinicians to assess impulsivity, it appears that what is being studied are a tendency for individuals to show a certain hesitancy to reflect and a lack of persistence on tasks. This behavior is also associated with poorer performance than in more reflective individuals and is consistent with studies examining the various effects impulsivity has on school performance (Messer, 1976; Thompson, Teare, & Elliot, 1983). It does seem that what is being assessed in most studies of the DAP, BG, Wechsler scales, and Rorschach is a general aspect of impulsivity that has to do with a consistent incapacity for delay. Further research efforts need to adhere to more strict definitions so that it is the multidimensional nature of impulsivity, and not the multidimensional nature of its definition, that is studied.

CONCLUSION

This review was an effort at trying to clear up the confusion over the concept of impulsivity by attempting to discuss the concept from empirical perspectives perti-

nent to the field of psychoeducational assessment. It is not meant to be completely comprehensive, as current research on the treatment of impulsivity and on etiologic factors is only briefly mentioned. It is my hypothesis that if definitional problems are clarified and distinctions from other terms are made, both the empirical and clinical aspects of impulsivity can be more fruitfully examined further and "impulsive behavior disorders" more effectively treated. It is my hope that future research will take into account the current status of the validity of common psychological instruments and that researchers and clinicians will seek and strive for a better methodology and conceptualization of this concept.

References

American Psychiatric Association. (1980). *Diagnostic and Statistical Manual of Mental Disorders* (3rd. ed.). Washington, DC.

Arizmendi, T., Paulsen, K., & Domino, G. (1981). The Matching Familiar Figures Test: A primary, secondary, and tertiary evaluation. *Journal of Clinical Psychology, 37,* 812-818.

Barratt, E. (1965). Factor analysis of some psychometric measures of impulsiveness and anxiety. *Psychological Reports, 16,* 544-547.

Barratt, E., & Patton, J. (1983). Impulsivity: Cognitive, behavioral, and psychophysiological correlates. In M. Zuckerman (Ed.), *Biological bases of sensation seeking, impulsivity, and anxiety.* Hillsdale, NJ: Erlbaum.

Becker, L., Bender, N., & Morrison, M. (1978). Measuring impulsivity-reflectivity: A review. *Journal of Learning Disabilities, 11,* 24-30.

Bender, L. (1938). *A visual motor gestalt test and its clinical use* (Monograph No. 3). New York: American Orthopsychiatric Association.

Bentler, P., & McClain, J. (1976). A multitrait-multimethod analysis of reflection-impulsivity. *Child Development, 47,* 218-226.

Blackman, S., & Goldstein, M. (1982). Cognitive styles and learning disabilities. *Journal of Learning Disabilities, 15,* 106-115.

Blaha, J., Fawaz, N., & Wallbrown, F. (1979). Information processing components of Koppitz errors on the Bender Visual-Motor Gestalt Test. *Journal of Clinical Psychology, 35,* 784-790.

Brannigan, G., & Ash, T. (1977). Cognitive tempo and WISC-R performance. *Journal of Clinical Psychology, 33,* 212.

Brannigan, G., Ash, T., & Margolis, H. (1980). Impulsivity-reflectivity and children's intellectual performance. *Journal of Personality Assessment, 44,* 41-43.

Brannigan, G., Barone, R., & Margolis, H. (1978). Bender Gestalt signs as indicants of conceptual impulsivity. *Journal of Personality Assessment, 42,* 233-236.

Brannigan, G., Margolis, H., & Moran, P. (1979). Cognitive tempo and children's human figure drawings. *Perceptual and Motor Skills, 49,* 414.

Buckley, P. (1978). The Bender Gestalt Test: A review of reported research with school-age subjects, 1966-1977. *Psychology in the Schools, 15,* 327-336.

Cairns, E., & Cammock, T. (1978). Development of a more reliable version of the Matching Familiar Figures Test. *Developmental Psychology, 5,* 555-560.

Cairns, E., & Harbison, J. (1975). Impulsivity: Self-report and performance measure. *British Journal of Educational Psychology, 45,* 327-329.

Connors, C. (1969). A teacher rating scale for use in drug studies with children. *American Journal of Psychiatry, 126,* 884.

Donnelly, E., & Murphy, D. (1974). Primary affective disorder: Bender Gestalt sequence of placement as an indicator of impulse control. *Perceptual and Motor Skills, 38,* 1079-1082.

Edman, G., Schalling, D., & Levander, S. (1983). Impulsivity and speed and error in a reaction-time task. *Acta Psychologia, 53,* 1-8.

Egeland, B., & Weinberg, R. (1976). The Matching Familiar Figures Test: A look at its psychometric credibility. *Child Development, 47,* 483-491.

Epstein, M., Cullivan, D., & Lloyd, J. (1977). Reliability of the Matching Familiar Figures Test scores of learning-disabled children. *Perceptual and Motor Skills, 45,* 56-58.

Eysenck, S. (1981). Impulsiveness and antisocial behavior in children. *Current Psychological Research, 1,* 31-37.

Eysenck, S., & Eysenck, H. (1977). The place of impulsiveness in a dimensional system of personality description. *British Journal of Social and Clinical Psychology, 16,* 57-58.

Falk, J. (1981). Understanding children's art: An analysis of the literature. *Journal of Personality Assessment, 45,* 465-472.

Finch, A., Spirito, A., & Brophy, C. (1982). Reflection-impulsivity and WISC-R performance in behavior problem children. *The Journal of Psychology, 111,* 217-221.

Gardner, R. (1951). Impulsivity as indicated by Rorschach test factors. *Journal of Consulting Psychology, 15,* 464-468.

Gilbert, J. (1980). *Interpreting psychological test data* (Vol. 2). New York: Van Nostrand Reinhold.

Glow, R., Lange, R., Glow, R., & Barnett, J. (1983). Cognitive and self-reported impulsiveness: Comparison of Kagan's MFFT and Eysenck's EPQ impulsive measures. *Personality and Individual Differences, 4,* 179-187.

Goodenough, F. (1926). *Measurement of intelligence by drawings.* New York: World Book.

Gudjonsson, G. (1980). The relationship between the EPI extraversion score and impulsiveness on a perceptual motor task. *Personality and Individual Differences, 1,* 177-180.

Hammer, E. (1965). Acting out and its prediction by projective drawing assessment. In L. Abt & S. Weissmann (Eds.), *Acting out.* Orlando: Grune & Stratton.

Handler, L. (1965). Figure drawing anxiety indices: A review of the literature. *Journal of Projective Techniques, 29,* 305-313.

Hardison, J. (1964). *An investigation of the relationships between several inkblot variables and impulse control.* Unpublished doctoral dissertation, University of Kentucky.

Harris, D. (1963). *Children's drawings as measures of intellectual maturity: A revision and extension of the Goodenough Draw-A-Man Test.* Orlando: Harcourt Brace Jovanovich.

Harrison, A., & Nadelman, L. (1972). Conceptual tempo and inhibition of movement in black preschool children. *Child Development, 43,* 657-668.

Hirschfield, P. (1965). Response set in impulsive children. *Journal of Genetic Psychology, 107,* 117-126.

Holtzman, W. (1950). Validation studies of the Rorschach test: Impulsiveness in the normal superior adult. *Journal of Clinical Psychology, 6,* 348-351.

Homatidis, S., & Konstantareas, M. (1981). Assessment of hyperactivity isolating measures of high discriminant ability. *Journal of Consulting and Clinical Psychology, 49,* 533-541.

Hutt, M. (1969). *The Hutt adaptation of the Bender Gestalt Test* (2nd ed). Orlando: Grune & Stratton.

Ikegami, T. (1979). Cognitive style in children: The relation between reflection-impulsivity and Rorschach scores. *Psychologia, 22,* 207-221.

Jackson, A., Farley, G., Zimet, S., & Gottman, J. (1979). Optimizing the WISC-R test performance of low- and high-impulsive emotionally disturbed children. *Journal of Learning Disabilities, 12,* 56-59.

Kagan, J. (1965). Impulsive and reflective children: Significance of conceptual tempo. In J. O. Krumboltz (Ed.), *Learning and the educational process.* Chicago: Rand McNally.

Kagan, J., Rosman, B., Day, D., Albert, J., & Phillips, W. (1964). Information processing in the child: Significance of

analytic and reflective attitudes. *Psychological Monographs, 78*(1, Whole No. 578).

Kaufman, A. (1984, May 21). Personal communication.

Kendall, P., & Wilcox, L. (1979). Self-control in children: Development of a rating scale. *Journal of Consulting and Clinical Psychology, 47,* 1020-1029.

King, C., & Young, R. (1982). Attentional deficits with and without hyperactivity: Teacher and peer perceptions. *Journal of Abnormal Child Psychology, 10,* 483-496.

Kipnis D. (1971). *Character structure and impulsiveness.* New York: Academic Press.

Koppitz, E. (1968). *The psychological evaluation of children's human figure drawings.* Orlando: Grune & Stratton.

Koppitz, E. (1975). *The Bender Gestalt Test for children.* Orlando: Grune & Stratton.

Lahey, B., Green K., & Forehand, R. (1980). On the independence of ratings of hyperactivity, conduct problems, and attention deficits in children. *Journal of Consulting and Clinical Psychology, 48,* 566-574

Levy, F., & Hobbes, M. (1981). The diagnosis of attention deficit disorder (hyperkinesis) in children. *Journal of the American Academy of Child Psychiatry, 20,* 376-384.

Lubin, B., Wallis, R., & Paine, C. (1971). Patterns of psychological test usage in the United States, 1935-1969. *Professional Psychology, 2,* 70-74.

Machover, K. (1949). *Personality projection in the drawings of the human figure.* Springfield, IL: Thomas.

Mandeville, P. (1954). *A study of the relationship between responsiveness to color on the Rorschach examination and impulsive behavior.* Unpublished doctoral dissertation, University of Nebraska.

Martin, R. (1983, October). The ethical issues in the use and interpretation of the Draw-A-Person test and other similar projective procedures. *Division of School Psychology Newsletter,* American Psychological Association.

McClanahan, R. (1979). *The Kansas Reflection-Impulsivity Scale for preschoolers: A construct validity study.* Unpublished doctoral dissertation, University of Kansas.

Messer, S. (1976). Reflection-impulsivity: A review. *Psychological Bulletin, 83,* 1036-1052.

Messer, S., & Brodzinsky, D. (1981). Three-year stability of reflection-impulsivity in young adolescents. *Developmental Psychology, 17,* 848-850.

Mistry, V. (1975). *Dimensions of reflectivity and impulsivity as correlated to cognitive, psychomotor, and affective domains of preschool children.* Unpublished doctoral dissertation, University of Southern Illinois.

Mollick, L., & Messer, S. (1978). The relation of reflection-impulsivity to intelligence tests. *Journal of Genetic Psychology, 132,* 157-158.

Monroe, R. (1970). *Episodic behavioral disorders: A psychodynamic and neurophysiological analysis.* Cambridge, MA: Harvard University Press.

Murray, J. (1938). *Explorations in personality.* New York: Oxford University Press.

Oas, P. (1983). Impulsive behavior and assessment of impulsivity with hospitalized adolescents. *Psychological Reports, 53,* 764-766.

Oas, P. (1984a). An investigation of the clinical validity of Kagan's Matching Familiar Figures Test with hospitalized adolescents. *Educational and Psychological Research, 4,* 177-183.

Oas, P. (1984b). *The psychological conceptualization of impulsivity.* Manuscript submitted for publication.

Oas, P. (1984c). Validity of the Draw-A-Person and Bender Gestalt tests as measures of impulsivity with adolescents. *Journal of Consulting and Clinical Psychology, 52,* 1011-1019.

Ogden, D. (1981). *Psychodiagnostic and personality assessment: A handbook.* Los Angeles: Western Psychological Services.

Ostrov, E., Offer, D., Marohn, R., & Rosenwein, T. (1972). The impulsivity index: Its application to juvenile delinquency. *Journal of Youth and Adolescence, 1,* 179-196.

Ownby, R. (1983, August). *Development of a scale to assess attention deficit disorder in children.* Paper presented at the annual convention of the American Psychological Association, Anaheim, CA.

Ozawa, J. (1980). *Attentional deficit disorder (DSM III): The diagnosis of distractability and impulsivity in learning-disabled and normal children.* Unpublished doctoral dissertation, University of Southern California.

Paulsen, K., & Johnson, M. (1980). Impulsivity: A multidimensional concept with developmental aspects. *Journal of Abnormal Child Psychology, 8,* 269-277.

Porteus, S. (1965). *The Maze Test and clinical psychology.* Palo Alto, CA: Pacific Books.

Quay, H., & Peterson, D. (1967). *Behavior problem checklist.* Urbana, IL: Children's Research Center, University of Illinois.

Rapaport, D., Gill, M., & Schafer, R. (1968). *Diagnostic psychological testing.* New York: International Universities Press.

Redmountain, A. (1976). *Validity of the sign approach in the psychological assessment of impulsivity among children and adolescents.* Unpublished doctoral dissertation, George Washington University.

Roback, H. (1968). Human figure drawings: Their utility in the clinical psychologist's armamentarium for personality assessment. *Psychological Bulletin, 70,* 1-19.

Saklofske, D., & Eysenck, S. (1983). Impulsiveness and venturesomeness in Canadian children. *Psychological Reports, 52,* 147-152.

Schafer, R. (1949). Psychological tests in clinical research. *Journal of Consulting Psychology, 13,* 328-334.

Schalling, D. (1975). *Psychopathic behavior: Personality and neuropsychology.* Paper presented at the conference on psychopathic behavior, Les Arcs, France.

Schecter, J. (1981). *Impulsivity: Its assessment by psychometric, clinical psychomotor, and situational measures in a behavior disorder population.* Unpublished doctoral dissertation, Loyola University of Chicago.

Shapiro, D. (1965). *Neurotic styles.* New York: Basic Books.

Shapiro, L. (1977). *The validation of the construct of cognitive impulsivity in normal and delinquent adolescent boys.* Unpublished doctoral dissertation, University of Colorado at Boulder.

Singer, S., & Herman, J. (1954). Motor and fantasy correlates of Rorschach human movement responses. *Journal of Consulting Psychology, 18,* 325-331.

Singer, J., Meltzoff, J., & Goldman, G. (1952). Rorschach movement responses following motor inhibition and hyperactivity. *Journal of Consulting Psychology, 16,* 359-364.

Singer, J., & Spohn, H. (1954). Some behavioral correlates of Rorschach's experience-type. *Journal of Consulting Psychology, 18,* 1-9.

Singer, J., Wilensky, H., & McGraven, V. (1956). Delaying capacity, fantasy, and planning ability: A factorial study of some basic ego functions. *Journal of Consulting Psychology, 20,* 375-383.

Sonnenblick, M. (1980). *A normative approach to human figure drawing in adolescence.* Unpublished doctoral dissertation, St. John's University.

Swenson, C. (1968). Empirical evaluations of human figure drawings, 1957-1966. *Psychological Bulletin, 70,* 20-44.

Thompson, R., Teare, J., & Elliot, S. (1983). Impulsivity: From theoretical constructs to applied interventions. *Journal of Special Education, 17,* 157-169.

Trahan, D., & Stricklin, A. (1979). Bender Gestalt emotional indicators and acting-out behavior in young children. *Journal of Personality Assessment, 43,* 365-375.

Verrill, B. (1958). *An investigation of the concept of impulsivity.* Unpublished doctoral dissertation, University of Houston.

Wade, T., Baker, T., Morton, T., & Baker, L. (1978). The status of psychological testing in clinical psychology: Relationships between test use and professional activities and orientations. *Journal of Personality Assessment, 42,* 3-10.

Wallbrown, J., & Wallbrown, F. (1975). Further evidence concerning the validity

of Kagan's comments on the clinical interpretation of the BenderGestalt. *Perceptual and Motor Skills, 41,* 51-54.

Wallbrown, F., Wirth, E., & Engin, A. (1975). Some evidence relevant to Kagan's comments on the clinical interpretation of the Bender Gestalt. *Journal of Clinical Psychology, 31,* 345-347.

Wright, J. (1973). *The KRISP: A technical report.* Unpublished manuscript, University of Kansas.

Zuckerman, M. (Ed.). (1983). *Biological bases of sensation seeking, impulsivity, and anxiety.* Hillsdale, NJ: Erlbaum.

Zuckerman, M., Kolin, E., Price, L., & Zoob, I. (1964). Development of a sensation-seeking scale. *Journal of Consulting Psychology, 28,* 477-482.