Concurrent Validation of the Affective Scale of the Diagnostic Assessment for the Severely Handicapped (DASH) Scale

Richard R. Ostrom
George Fox University

Follow this and additional works at: http://digitalcommons.georgefox.edu/gscp_fac

Part of the Clinical Psychology Commons

Recommended Citation
http://digitalcommons.georgefox.edu/gscp_fac/59
Concurrent Validation of the Affective Scale of the Diagnostic Assessment for the Severely Handicapped (DASH) Scale

by

Richard M. Ostrom

Presented to the Faculty of George Fox College in partial fulfillment of the requirements for the degree of Doctor of Psychology in Clinical Psychology

Newberg, Oregon

June 15, 1995
ABSTRACT

Sixty-nine severely and profoundly retarded clients in a residential setting, each of whom could ambulate without staff assistance, were administered the Diagnostic Assessment for the Severely Handicapped (DASH) Scale. Three groups of 23 were selected based on previous psychiatric diagnosis and matched for social age. The group with no psychiatric diagnosis served as the nominal control. The second group contained a mix of non-affective psychiatric diagnoses. The third group contained clients with a diagnosis of bipolar disorder. A one-way analysis of variance was conducted. The data demonstrated the DASH effectively discriminated clients with bipolar disorder from those without psychiatric as well as from non-affective psychiatric disorders. The implications of these findings were discussed, and areas for future study were suggested.

Table of Contents

Abstract ....................... ii Table of Contents ................. iii
List of Tables ................. vi
CHAPTER 1: INTRODUCTION ........................ 1  Problem .................

...................... 2  Related literature ................. 3

Content validity ............... 17  Construct validity ..........

. 18  Criterion validity ............... 19  Concurrent validity ......

..... 19  Predictive validity ............... 20  Hypotheses ..............

.................... 21

Summary ....................... 22

CHAPTER 2: METHODS ............... 24  Subjects .................

........... 24

Procedure ....................... 26  CHAPTER 3: RESULTS ........

........... 44  Data summary ............... 44  Statistical methods .

........... 44  Data tables ............... 45  CHAPTER 4:

DISCUSSION ............... 46  Interpretation ............... 46  Implications ...........
List of Tables

Table 1  Types of validity
CHAPTER 1

Introduction

Diagnosis is of great consequence for the prognosis and treatment of psychiatric disorders. Diagnostic consistency was advanced with the development of the American Psychiatric Association's Diagnostic and Statistical Manual 3rd edition - revised (American Psychiatric Association, 1987). Although greater diagnostic consistency has been achieved over the past two to three decades, the process remains largely an art form featuring subjective clinical judgments. When attempting to differentiate between mental retardation and psychiatric disorders (Axis I and II disorders), the process becomes more opaque. In this case, the behavioral manifestations of mental subnormality due to developmental arrest must be differentiated from impairment or disruption in ongoing cognitive and affective functioning, such as the disrupted or regressed intellectual functioning common in psychosis.

Differentiating between mental subnormality and cognitive abnormality is difficult for two reasons. First, many shared characteristics can cloud the key differences between mental subnormality and cognitive abnormality. Second, the lack of a reliable, valid, and standardized method to discriminate objectively between the two conditions (Kazdin, A.E., Matson, J.L., & Senatore, V., 1983).

Increasing attention is being given to the issue of dual diagnosis in the severely/profoundly retarded population. A dual diagnosis occurs when some psychiatric disorder is diagnosed along with the client meeting the criteria for and being diagnosed as mentally retarded.
Recent epidemiological studies (Corbett, 1979; Lewis & MacLean, 1982; Gostason, 1985) have demonstrated that handicapped persons can develop all types of psychiatric disorders. The mental health clinician has to properly assess the presence of developmental disability in conjunction with a psychiatric disorder and vice versa.

Statement of the Problem

Although there is a need for an instrument to assess psychopathology in severely and profoundly mentally retarded persons, relatively little research has been conducted to develop such an instrument (Matson, J.L., Gardner, W.I., Coe, D.A., & Sovner, R., 1991a). The Diagnostic Assessment for the Severely Handicapped (DASH) has recently been developed by Matson at Louisiana State University to address this need. Currently, limited psychometric research has been conducted on the DASH. Matson has conducted construct and content validity studies indicating that the categories in the DASH are distinct factors in the diagnosis of psychiatric disorders in this population. The face validity of the items in the DASH has also been examined (Matson, et al., 1991a).

The present study is a concurrent validity study seeking to demonstrate that the DASH can identify psychiatric disorder in general, and bipolar disorder in specific, in a sample of severely and profoundly mentally retarded clients living in a residential setting. The word nonretarded will be used throughout to denote persons of normal intelligence. This is the standard usage in the literature in this field.

Related Literature
Matson and Frame (1986) and Russell (1988) conducted survey studies which concluded that psychiatric/behavioral symptoms occur frequently within the mentally retarded population. However, as the level of retardation increases, the recognition of particular psychiatric disorders becomes more difficult. Costello (1982) emphasized the difficulty of diagnosing schizophrenia in the severely retarded.

Similarly, Sovner and Hurley (1983) noted that affective disorders (e.g., depression or mania) were very difficult to assess by clinical presentation only. This difficulty in assessment is due largely to the absence or unreliability of verbal input from most of the severely and profoundly retarded clients. The diagnosis of psychiatric disorders is centered around the client's verbal report of thoughts and feelings. Without this input, DSM-III-R and other currently accepted diagnostic approaches are of limited usefulness. ***rewritten at home***

Baumeister and MacLean (1979) suggest that the retarded are more susceptible to psychiatric disorders than the nonretarded population due to their physical vulnerability. As the level of retardation increases, the level of physical vulnerability also increases. The previously cited epidemiological studies show that the lack of coping skills and general intellectual limitations predispose this population to be at risk for psychiatric disorders. The estimates of prevalence of psychiatric disorders in this population range from 7% to 59%, with the higher rates tending to be found in institutions (Costello, 1982). In addition to physical vulnerability and lack of intellectual and coping skills, the severely/profoundly retarded are further restricted in communication and social skills. The
combination of these factors presumably explains the higher prevalence of psychiatric disorders in this group (Helsel and Matson, 1988).

Matson and Sevin (1988) pointed out that the usual assessment techniques (DSM-III-R, clinical interviews, and psychological testing) are not adequate for use with the mentally retarded population. They describe the need for an assessment device for the severely/profoundly retarded client by stating, "...the extreme variability of behaviors in this population suggests the need for precise, detailed descriptions and the strict following of definitional guidelines in diagnosis in order to avoid analogue errors" (page 14).

The crux of the diagnostic/assessment problem is bound up in the very nature of mental retardation; that is, it is difficult to discern what disturbances of behavior are the direct result of subnormal intellectual functioning (Lewis and MacLean, 1982). Corbett has contended that the present diagnostic system does not have the predictive value for mentally retarded individuals that it reportedly has for nonretarded individuals.

Relatively little research has been conducted aimed at developing an instrument specifically for the assessment of psychopathology in the mentally retarded. More generally, very little research has been conducted concerning mental illness in the mentally retarded.

The lack of research in this area was demonstrated by Heaton-Ward (1977). It was noted that from 1962 through 1977, 1300 papers were presented at the International Congress for the Scientific Study of Mental Deficiency (ICSSMD) on mental retardation. Only forty of these presentations addressed
mental illness. Twenty of these forty papers concerned autism, and only two concerned psychosis in the retarded.

Lange (1990) indicates that the situation has not really improved in the past twelve years. He states, "In contrast to the number of studies on lithium for the general population, for the key words lithium and mental retardation, only fifty-one entries were listed during the same time interval. This is compared to 5,416 entries in the literature for the lithium since 1970 (nonretarded population) (page 448)."

Reid (1989) reports that the four major studies that have been conducted in the United Kingdom agree that schizophrenia is a problem of the mildly and moderately retarded people, and cannot be diagnosed on clinical grounds in the more severe and profound ranges of retardation. These were clinical studies and of better design than most other studies in this area. The diagnoses in these studies were made on the basis of a DSM clustering of symptoms approach to diagnosis.

Sovner (1989) outlines several reasons for the development of better psychiatric assessment techniques for use with severely or profoundly retarded persons. First, good psychiatric assessment can make an important difference in the use of psychotropic medication for this population. Currently, psychotropic medications are administered without a reliable data base in many instances. Assessment is a major part of the decision-making process. Second, careful assessment is a prerequisite in the development of an overall treatment plan. Finally, increased diagnostic precision can lead to more precision in the annual plan of care required by law for all mentally retarded persons in the country.
The medical model is based on the principle that a specific condition requires a specific treatment. That is, syndrome X requires drug Y. In the severely/profoundly retarded population, psychiatric disorders have not yet been well described. Therefore, a skilled application of psychotropic medications becomes difficult. Concomitantly, research into drug efficacy is limited. Behavioral treatments and research in the application of behavior modification techniques with the mentally retarded will be better focused with better assessment techniques. Increasing the clinician's ability to see and describe deviations from the expected patterns of behavior would obviously aid in refining treatment. Recognizing behavior clusters helps in creating hypotheses about what is motivating the client's behavior. This is also a more efficient method than looking at a behavior in isolation. Building on this foundation, reinforcers and substitute behaviors can be better selected.

As mentioned previously, the extent to which assessment techniques used with the nonretarded can be applied to the retarded is not known. Evidence exists to support the notion that these techniques are not generalizable to the severely/profoundly retarded population. Even when the assumption is made that the etiology of psychiatric impairment is the same regardless of the level of cognitive/intellectual functioning, it may be that the expression of the impairment takes on a unique form. Because of the limited cognitive ability and non-normal experiences of the severely/profoundly retarded, it can be predicted that psychiatric symptom presentation will be more difficult to differentiate than in the nonretarded population.
The existence of "primitive behaviors" in the severely/profoundly retarded population further complicates the assessment process. These behaviors can be thought of as occurring earlier in life than more complex behaviors. If these primitive behaviors are seen in a nonretarded person, their presence indicates psychiatric impairment. The presence of these behaviors in the severely/profoundly retarded does not necessarily indicate impairment. It depends upon the person's mental age and overall development. Clinicians assessing both nonretarded and retarded children and adolescents encounter similar challenges. This discussion will be confined to adults. Currently, the importance of a co-occurring behavior depends on the significance it has in the severely/profoundly retarded generally.

The DASH

The Diagnostic Assessment for the Severely Handicapped (DASH) Scale is an instrument currently being developed under the direction of Matson at Louisiana State University. The intention is that when it is completed mental health technicians will be able to survey the psychiatric problems of specific people within the severely/profoundly retarded population. Normative data have recently been gathered in various settings (Matson, et al., 1991a). These data will be used to develop a picture of the severely/profoundly retarded person.

The individual DASH questions were selected on the basis of face validity and their presence in diagnostic instruments for other populations. That is, based on clinical experience, other instruments, and the DSM-III-R. Thus, the item content and clusters of the DASH resemble other diagnostic approaches such as DSM-III-R.
The DASH is composed of two parts. The first part is a general information section. Under this heading occurs questions concerning staffing ratio, mental retardation classification, medication, and physical disabilities of the individual. The second part is a behavior rating section. This section contains 86 items, each of which is rated according to three dimensions: frequency, duration (history), and severity. It is administered by a clinician to an informant who knows the client well. The administration takes about 30 minutes.

The items chosen for inclusion in the DASH were selected by its author based on clinical experience, other instruments (used for different populations), and the DSM-III-R. They represent 12 categories of behavior disorder: anxiety, affective disorders, autism, schizophrenia, stereotypies, self-injurious behavior, elimination disorders, eating disorders, sleep disorders, sexual disorders, organic syndromes, and disorders of impulse control.

As mentioned previously, each of the 86 questions in the behavior assessment section of the DASH can be rated on the dimensions of severity, duration, and frequency. Each of these dimensions can receive a 0, 1, or 2 rating. The rating is done according to specific criteria contained in the instructions accompanying the instrument. An individual's score on a psychiatric category is developed by addition of the item scores associated with that category. A higher score indicates a greater probability of that specific psychiatric category. Since these items are ordinal level data, it is not possible to say that a score that is twice another score means that the person possess twice the amount of that psychiatric symptom category.

Validity
Since the present study is an examination of the concurrent validity of one of the categories of the DASH, it is important to examine the subject of validity in general, and concurrent validity in specific. The three main types of validity will be discussed initially, followed by a discussion of concurrent validity.

The Standards for Educational and Psychological Tests and Manuals (1985) specify five areas of technical standards for test construction and evaluation to consider when creating or evaluating a test: 1) validity, 2) reliability and errors of measurement, 3) test development and revision, 4) scaling, norming, score compatibility, and equating; and 5) test publication: technical manuals and user's guide.

According to the Standards, validity is the most important consideration in the test evaluation. Validity refers to the appropriateness, meaningfulness, and usefulness of the inferences made from the instrument. Test validation is a process of accumulating evidence which can support such inferences made from the instrument. Validity refers to the degree to which the evidence collected can support inferences made from scores on the test. That is, the inferences regarding specific uses of the tests are being validated, the test itself is not being validated (Messick, 1975).

The evidence for validity can be gathered in more than one manner. Traditionally, validity is grouped in the following categories: content, construct, and criterion-related. This does not imply that they are completely distinct groups, because it is not possible to make rigorous distinctions among the categories. Table 1 presents a summary of the different types of validity according to various sources.
<table>
<thead>
<tr>
<th>Types of Validity</th>
<th>I. Content</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Face</td>
</tr>
<tr>
<td></td>
<td>Logical</td>
</tr>
<tr>
<td>II. Construct</td>
<td>Developmental Changes</td>
</tr>
<tr>
<td></td>
<td>Correlations with other Tests in the Same Domain</td>
</tr>
<tr>
<td></td>
<td>Factor Analysis</td>
</tr>
<tr>
<td></td>
<td>Internal Consistency</td>
</tr>
<tr>
<td></td>
<td>Convergent</td>
</tr>
<tr>
<td></td>
<td>Discriminant</td>
</tr>
<tr>
<td></td>
<td>Experimental Intervention</td>
</tr>
<tr>
<td></td>
<td>Known-group Differences</td>
</tr>
<tr>
<td>III. Criterion</td>
<td>Concurrent</td>
</tr>
<tr>
<td></td>
<td>Predictive</td>
</tr>
</tbody>
</table>

The Standards state that evidence of validity should be presented along with the rationale used to support the evidence. This process needs to occur for each type of inference that the test is recommended for. Validity relating to the subscales and the procedure, sample composition, and any factors which may influence validity should also be reported.

**Content Validity**

This refers to the systematic examination of test content to determine whether it covers a representative sample of the behavior domain being measured (Kaplin and Saccuzzo, 1982). Content validity is the only type for which evidence is logical rather than statistical. This is the type of validity sought in achievement tests. It is built into the tests during test development through the choice of appropriate items. According to Anastasi (1988), primary reliance on content validity is usually inappropriate for a personality test. Other types of validation are necessary for personality tests to be verified for inferences to be made legitimately.

Personality tests are not based on a specified course of instruction or a known set of prior experiences from which a test author can draw items during test construction. Rather, people are likely to vary in their psychological processes when responding to the test items, thereby measuring different functions. In truth, the content of personality tests reveals more about the hypotheses of the test author in choosing a certain content to measure a certain trait.
There are two types of content validity according to Anastasi: face validity and logical validity. When a reading of the test items leads the reader to conclude that the test is measuring what it is supposed to be measuring, face validity has been achieved. Matson chose items for the DASH that appear to be associated with psychopathology in the severely and profoundly retarded. He relied on his considerable experience in this field, knowledge of other diagnostic tests, and the inclusion of the DSM-III-R criteria (Matson, et al., 1991).

In referring to logical validity, the Standards states that content-related evidence serves as a significant demonstration of validity for a test's use. In addition, a clear definition of the universe represented, its relevance to the proposed use, and the procedures followed in generating test content to represent that universe should be described. When subject matter experts are used to demonstrate logical validity, the qualifications of the experts, along with the procedures used to obtain a consensus on test items, should be reported.

Anastasi asserts that content validity is less applicable to affective trait measurement. This is because it is difficult to define the "universe" of a psychological construct. This is especially difficult when attempting to define psychopathology in the severely/profoundly retarded.

**Construct Validity**

Construct validity is the most general type of validity. Kerlinger (1986) describes construct validity as basically judgmental. It incorporates evidence from studies of the content and criterion related validity of the test (Anastasi, 1988). The construct measured should be embedded in a conceptual framework which specifies the meaning of the construct, distinguishes it from other
constructs, and suggests how measures of the construct should relate to the variables (Standards, 1985).

The construct validity of an instrument refers to how well the instrument represents the construct whose name appears in the title. The concern of construct validity is not to explain a single behavior or item response. Rather, it seeks to account for consistency in item responses which have a small number of determinants, or sometimes one major determinant.

When a test is proposed as a measure of a construct the proposed interpretation of the test score must be explicitly stated. In addition, evidence needs to be presented to support such inferences. Evidence also needs to demonstrate that the test scores are more closely related to variables of theoretical interest than to variables not included in the theoretical framework. This requirement also applies to all subscales.

There are seven specific techniques available for construct validation: developmental changes, correlations with other tests, factor analysis, internal consistency, convergent and discriminant validation, experimental interventions, and known-group differences (Anastasi, 1988; Kerlinger, 1986; Mueller, 1986).

The technique of age differentiation is used to demonstrate developmental changes. Naturally this would not apply to any construct that did not exhibit clear and consistent age changes. Anastasi (1988) notes that the area of personality measurement has limited use for the technique of age differentiation.

Correlations between a new test and earlier well-established tests of a similar nature can be cited as evidence of construct validity. Correlations need to
be moderately high, but not so high that the need for the test is questioned because there is so much similarity (Anastasi, 1988).

**Factor analysis** is a further technique for demonstrating construct validity. Factor analysis is a sophisticated statistical technique used to analyze the interrelationships of data. It displays the common traits that would account for the correlations obtained among measures (tests, test items, or other measures depending on the level of analysis). The process involves reducing the number of variables from, for example, the items of a test down to a small number of factors. After the factors have been identified they can be used to describe the factorial make-up of the test. Once the factorial composition of the test is determined, the items can then be characterized in terms of the major factors determining their scores. Each factor is given a weight, or loading, indicating how highly it correlates with the test.

**Internal consistency** applies to both reliability and validity. It gives evidence of construct validity by showing whether items on a scale have a high level of intercorrelation. High levels of intercorrelation imply that the items are measuring the same underlying variable, that is, that a construct is being measured. Meuller (1986) notes that the demonstrated presence of a construct does not imply it is the correct construct (the one purportedly measured by the instrument). That is, internal consistency cannot be solely relied on in the validation process.

**Convergent validity** is the broad area of how well a test correlates with other variables with which it should theoretically correlate (Anastasi, 1988).
Discriminant analysis is another sophisticated statistical procedure (similar to factor analysis (described above) which is used for predicting group membership from multiple variables. This can be used to find groups of variables or measures that could predict group membership, such as diagnostic category. Thus, it could be a first step in establishing predictive validity (see later in this section).

Another source of data to demonstrate construct validity is experiments on the effect of selected variables on test scores (Anastasi, 1988). This was referred to as a variant of the know group method (Meuller, 1986). This approach entails comparing the mean scores of a group know to be high in the construct with a group known to be low. Naturally, it is assumed that the experimental treatment is effective for this approach to be used.

Finally, construct validity can be demonstrated by an approach called known-group differences. It is assumed when an instrument shows differences between groups that are known to be different that it is demonstrating the construct it purports to.

Criterion Validity

According to Kerlinger (1973), every test instrument requires validation by relating the performance on that instrument to performance on other measures (criteria). Test performance can be evaluated against these other measures. Anastasi (1988) conceives of criterion validity as assessing how effective a test is at predicting a person's behavior in specific situations.

Concurrent Validity If the criterion measure is available at the time of testing, then concurrent validity is being studied. Under appropriate
circumstances, data obtained in a concurrent validation study allows inferences to be made as to the probable present standing of an individual on the criterion. In other words, concurrent validity reflects only the status quo at a certain time. Currently, there are construct and content validity studies available on the DASH, but no concurrent validity studies (Matson, et al., 1991).

Predictive Validity

Predictive validity refers to a test's ability to predict the results of a criterion measure. The better the correlation between the test and a criterion measure, the better the predictive validity of the test. Currently, there are no predictive validity studies available concerning the DASH.

Anastasi (1988) asserts that the test development process should follow these steps: a theoretical description of the test, item development and selection, psychometric investigation, and normative data collection. As mentioned earlier, this process is just beginning with the DASH. Clearly, much investigation of this diagnostic measure will need to occur in the future.

The present study fits into the psychometric investigation step of the above process. Matson et al. (1991a, 1991b) indicate that several factors are needed in the research design for this type of study. These are: 1) developing several comparison groups one of which acts as the control, 2) matching across comparison groups according to social (rather than chronological) age, and ability to ambulate independently, 3) obtaining ratings on the instrument, 4) building in inter-rater and psychiatric diagnosis reliability checks, and 5) statistical analysis.

***

**Three equal-sized groups will be matched for social age (SA).** Additionally, clients will be selected based on whether they are able to ambulate without staff assistance other than by wheelchair. Ratings for each client will
be obtained by interviewing a direct care staff member who has known the client for at least one month.

A one-way analysis of variance will be conducted comparing the mean frequency scores of the three groups on the Affective scale of the DASH. This is referred to as an intact groups comparison.

Hypotheses

It is hypothesized that the DASH will discriminate between clients who have a psychiatric diagnosis of bipolar disorder and those who do not. More specifically, the group previously diagnosed as bipolar will have a higher mean frequency score on the Affective scale of the DASH than either of the two comparison groups.

A second hypothesis is that subjects in the mixed psychiatric disorder group will have higher mean frequency scores on the Affective scale than the non-psychiatric disorder group.

Thus, the first hypothesis is that the DASH can differentially diagnose a particular disorder and the second, more basic hypothesis, is that it can detect psychiatric disorder.

Summary

Diagnosis is of great consequence in the treatment of psychiatric conditions. The usual psychiatric diagnostic procedures, i.e., clinical interview, observation, psychometric assessment, and the application of DSM-III criteria are of doubtful validity with the severely/profoundly retarded population. A further challenge is the difficulty in discriminating between intellectual subnormality and behavioral abnormality.
Relatively little research has been conducted to develop an instrument to specifically assess psychopathology in the severely/profoundly retarded. Matson has developed the Diagnostic Assessment for the Severely Handicapped (DASH) Scale to assess psychopathology in this group. No concurrent validation studies have yet appeared in the literature due to the newness of this instrument.

A concurrent validation study of the Affective scale of the DASH will be conducted. It is hypothesized that the mean frequency score for a specific criterion group previously diagnosed with bipolar disorder (A) will be higher than the mean frequency scores for a mixed psychiatric diagnosis group (B), which in turn will score higher than the non-psychiatric diagnosis group (C). That is, the mean frequency scores for the Affective scale will be ranked: A > B > C.
CHAPTER 2

Methods

The present study examined the concurrent validity of the Affective scale of the Diagnostic Assessment for the Severely Handicapped (DASH) Scale (Matson, et al., 1991a). Severely and profoundly retarded clients in a residential setting were studied.

Subjects

All of the participants in the study were residents of Fairview Training Center in Salem, Oregon. Fairview was founded in 1908 as a hospital for the mentally retarded. At its high point, Fairview was as large as a small town. At that time, there were nearly 3,000 full-time residents, and a similar number of staff. At the present time, Fairview is in the process of significantly reducing the number of clients in keeping with an agreement between the State of Oregon and the United States Department of Justice.

Fairview serves clients who have been found eligible to receive Medicare funding through the Health Care Financing Administration (HCFA). In order to be classified as eligible for Medicare funding, the client must be assessed as being two or more standard deviations below the mean on a standardized intelligence test and on a standardized instrument for assessing adaptive skills, e.g., the Vineland Adaptive Behavior Scale (VABS). In cases where it was not possible to obtain a valid testing on a standardized measure of intelligence, the level of
intellectual functioning is inferred from the results of the standardized test of adaptive skills.

As of March 1, 1995, there are approximately 370 clients residing at Fairview. In the course of the systematic placement of clients into community residences, the higher functioning clients have been placed more easily than those with more difficult care requirements. Because of this process, the majority of the clients now living at Fairview are severely and profoundly retarded.

The terms severely and profoundly retarded, as defined in the DSM-III-R, refer to those people who are roughly four to five standard deviations below the mean on standardized tests of intelligence and adaptive ability. The severely retarded clients are generally regarded as four to five standard deviations below the mean, and the profoundly retarded are considered to be five or more standard deviations below the mean. This roughly corresponds to a full-scale IQ of between 20-39 on the Wechsler Adult Intelligence Scale - Revised (WAIS - R) for the severely retarded, and less than 20 on the WAIS - R for the profoundly retarded. Each category is similar for adaptive skill level, and both intelligence and adaptive skill deficits must be present to make the diagnosis.

Procedure

A list of the clients was obtained from the data base kept by the institution. All of those clients not previously classified in the institutional records as severely or profoundly retarded were excluded from the study. The remaining clients were then classified on the basis of institutional records according to psychiatric diagnosis (or no psychiatric diagnosis) and the ability to ambulate without staff assistance.
It was noted at this point that the most common psychiatric diagnosis given to severely or profoundly retarded clients at Fairview was bipolar disorder. This group provided a natural subgroup of severely or profoundly retarded clients with a current psychiatric diagnosis. An equal number of the remaining psychiatrically diagnosed clients were chosen who could be matched with the bipolar group according to their social age. This group was composed primarily of clients who had been previously diagnosed as schizophrenic. Clients with a variety of personality disorders were also placed in this group.

Finally, the same sized group was selected from the clients listed who did not have a psychiatric diagnosis. These clients were also matched with the bipolar group on the dimension of social age. The eligible clients in this final group were assigned to the group randomly. A random numbers table was used in this process.

Social age was determined by the use of the Vineland Adaptive Behavior Scales (VABS). It is the mean of the age equivalents found in the various skill domains, e.g., communication. VABS scores were obtained from institutional records. Since DASH results have previously been shown to vary with social age, it is believed that social age must be controlled for. Matching is a first control (Matson, et al., 1991a).

Interviews were conducted by psychology undergraduate students on an individual basis. Training of the interviewers included practice administrations of the DASH in the presence of the trainer. The trainer conducted a second administration on every tenth administration to check for inter-rater reliability.
Inter-rater reliability was calculated by setting up a two-by-two contingency table of the possible agreements and disagreements between two observers on the occurrence (+) or nonoccurrence (-) of an event. The percentage of agreement (PO), the percentage nonagreement (PNO), and the overall percentage of agreement (OPA) were then calculated.

How are they calculated?

Materials

The Diagnostic Assessment for the Severely Handicapped (DASH) Scale was developed by Matson (1986) at Louisiana State University for use in assessing symptoms typically reported in research and clinical practice as troublesome for many severely or profoundly mentally retarded persons (Matson, et al., 1991a; Matson et al., 1991b). The items that make up this scale were taken from the American Psychiatric Association's Diagnostic and Statistical Manual, 3rd edition - Revised (American Psychiatric Association, 1987), as well as previously published studies of this population (Aman, et al., 1985; Leuder et al., 1984).

The DASH is a multidimensional instrument that assesses the frequency, severity, and the history of specific symptoms and conditions. Included are anxiety, depression, pervasive developmental disorders, and impulse control problems. Inter-rater reliability when using direct care staff in diagnosing psychiatric conditions was previously studied (Matson, et al., 1991b). It was
shown that direct care staff could provide reliable ratings of the behavior of profoundly and severely retarded clients. Several distinct conditions could be identified including pervasive developmental disorder, impulse-control problems, eating disorders, stereotypies, and tics. However, the DASH subscales were organized along the same lines as the DSM-III-R classifications.

The DASH consists of two sections. The first is a background section composed of questions about the client's medical background, the staffing ratio on the ward, and any physical disabilities. The second part is a behavior rating section. The behavior rating section contains 86 items, each of which is rated on three dimensions: frequency, duration (history), and severity.

The DASH is designed to be administered by interviewing a respondent who has known and directly cared for the client for at least one month. The administrator fills out the DASH based on the responses given by the respondent in the four areas for each scale item: frequency, duration, severity, and whether they know of it occurring within the past year. The responses are structured so that a 0, 1, or 2 score can be entered. This interview format is necessary because nearly all of the clients are nonverbal.

The items on the DASH were chosen based on their face validity from the test author's viewpoint. The test author's clinical experience, knowledge of other diagnostic instruments, and interpretation of the relevant DSM-III criteria were the basis for the decision about the face validity.

Some research has been conducted concerning the content validity of the DASH (Matson et al., 1991a; Matson et al. 1991b). In addition, the 13 subscales have been factor analyzed. Six orthogonal scales representing 39% of the
variance were demonstrated. It was noted that the loadings of the 86 scale items onto these six scales was largely in the vegetative symptoms area. The inter-rater reliability has been demonstrated to be quite high (.90 or higher) indicating that direct care staff are able to provide reliable information concerning the scale items (Matson et al., 1991b).

Procedures

Five junior and senior undergraduate psychology students were trained to properly administer the DASH. The training was conducted by a licensed psychologist who is familiar with the proper administration techniques. Each of these interviewers was randomly assigned to approximately equal sized client groupings. A random number table was used to complete the random assignments.

Three practice administrations were done by each of the interviewers. These were conducted in the presence of the trainer while the trainer filled out the same form. The interrater reliability was calculated, and the interviewers were debriefed.

Every tenth administration was checked for inter-rater reliability. The trainer interviewed the same respondent within 24-48 hours of the original administration. The results obtained by the trainer were compared to those obtained by the other administrators. The inter-rater reliability was computed by a simple ratio of items scored identically compared to the total number of items. All of the protocols were kept in the sample regardless of how well they matched the scoring by the trainer.

Research Design
An intact groups design was implemented. A one-way analysis of variance was conducted comparing the mean frequency scores of the three groups. The nominal independent variable in this design was group membership, since that occurs according to criteria established by the researcher. The groups were: no psychiatric diagnosis, mixed psychiatric diagnosis, and bipolar disorder. The dependent variable was the mean frequency score for each group on the Affective scale of the DASH.

Three groups of clients were matched as closely as possible according to social age. Social age served as a control variable. That is, the amount that social age was allowed to vary was restricted as much as possible given the varities of the sample. This reduced the nonsystematic variance. Nonsystematic variance can mask the effect being studied.

Social age, chronological age, and gender are dimensions that could have been matched across groups. Chronological age was thought to be of little consequence since the age at onset of bipolar disorder is nearly always after the early twenties (Diagnostic and Statistical Manual-Revised, Third Edition, 1987). As mentioned previously, gender has not been shown to be a significant factor (Matson, et al., 1991a). In addition, the groups were not matched for gender because the preponderance of the Fairview clients with bipolar disorder were males (approximately 80%). The sample size available was not large enough to allow matching for gender when the percentage of males in the bipolar disorder group was so high. In addition, the DSM-III-R reports that bipolar disorder occurs about as frequently in males as females.
For the above reasons, it was decided that matching across groups would only be done for social age. Also, social age is the most meaningful of the three dimensions. It must be remembered that these clients are severely and profoundly developmentally delayed. This means that they did not properly develop beyond a very early developmental level. Most of the clients in the study have not developed intellectually beyond the one-year level, although they are chronologically and biologically adults. Clearly, it is more meaningful to compare a client who is at an overall developmental level of 6 months with another client at the 6 month level than to compare two 30-year-olds of vastly different developmental levels.

Each of the 86 items on the DASH is scored for severity, duration (history), and frequency. It was decided that only the frequency score would be used in the analysis. The severity rating refers to such information as the need for medical attention during an episode of the behavior. This was not the aspect of the behavior being focused upon. This study was interested in examining whether or not the symptom was present.

The scores on the Affective scale of the DASH were summed within each group level. The Affective scale is composed of 21 items relating to depression and mania, eg. "is restless or agitated", "wakes up frequently during the night". These items are given a 0, 1, or 2 score by the respondent depending upon the frequency. A score of 2 indicates that the symptom occurred more than ten times in the previous two weeks. The arithmetic mean of the scores was compared using a one-way analysis of variance. A significance level of .05 was used to make decisions about support for the research hypothesis.
CHAPTER 3

Results

Data Summary

The results of the one-way ANOVA conducted on the mean frequency scores for frequency of affective symptomatology for the three treatment groups indicated a statistically significant difference among groups, \( F = ( \text{ }, \text{ } ), p < .05 \). Subsequent \( t \) tests (Tukey-HSD test) indicated that the DASH was sensitive enough to discriminate between the no psychiatric diagnosis group (group one) and the bipolar group (group three), \( F(\text{ } ) = , p < .05 \).

It does not appear sensitive enough to effectively discriminate between the bipolar disordered group and the group contain a variety of psychiatric disorders (group two), \( F = ( \text{ }, \text{ } ), p < . \).

An alternative way of describing the results of the statistical analysis is that the mean frequency score for affective symptomatology for the bipolar disordered group was significantly higher than the mean scores for either of the other groups, \( F = ( \text{ }, \text{ } ), p \text{ .} \). In addition, the mean affective symptomatology frequency scores for the mixed psychiatric disordered group was not significantly higher than the mean score for the no psychiatric diagnosis group, and was not significantly lower than the bipolar disordered group.

A variety of other comparisons were also conducted. The affective scale was divided into the mania and depression scales and one way ANOVAs were
These were conducted to assess for any differences in the sensitivity of the two subscales for discriminating affective symptomatology.

The mania scale was demonstrated to be a better discriminator than the depression scale. The mania scale was sensitive enough to discriminate the control group from the psychiatric and the bipolar groups, $F = (\ ,\ )$, $p < .05$. In addition, the mania scale was nearly able to discriminate between the control group and the mixed psychiatric disordered group, $F = (\ ,\ )$, $p < .0862$.

The mania scale was not able to distinguish between the mixed psychiatric disordered and the bipolar disordered groups, $F = (\ ,\ )$, $p < .3527$.

Comparing the depression scale across groups produced a significant effect, but only between the control group and the bipolar disordered group, $F = (\ ,\ )$, $p < .1209$. In addition, the depression scale was demonstrated to be able to discriminate shown to be sensitive when comparing the control with the mixed psychiatric disordered group indicating that it was not quite able to discriminate these groups, $F = (\ ,\ )$, $p < .1209$.

The depression scale was unable to distinguish between the mixed psychiatric disordered group and the bipolar disordered group, that is groups two and three, $F = (\ ,\ )$, $p < .1209$.

A variety of post hoc comparisons designed to analyze the potential contribution of demographic factors were considered. Three demographic factors were chosen for further analysis. They were: gender, the length of time that the informant had known the participant, and .

Statistical Methods
A one-way analysis of variance (ANOVA) was conducted on the affective symptoms frequency means of three independent treatment groups. The groups had been matched for social age as much as possible to control some of the systematic variances. In addition, all of the subjects were able to ambulate without staff assistance.

Data concerning the duration and severity of the scale items was collected as part of the standard administration procedure for the DASH. There were no a priori reasons to believe that they would contribute to the diagnostic utility of the instrument. Therefore, this data was not analyzed. Severity may represent a "background noise" factor, that is, that there are a variety of symptoms loaded onto a variety of subscales.

There were no a priori reasons to believe that duration data would add to the discriminability of the instrument. In some diagnoses duration may be a consideration. However, the hypotheses in the present study were too global for duration to be of significance.

CHAPTER 4
Discussion

Interpretation

The results of the one-way ANOVA supported only one of the research hypotheses. The Affective scale of the Diagnostic Assessment for the Severely Handicapped (DASH) was demonstrated to be sensitive to differences in affective symptomatology in severely and profoundly mentally retarded clients in a residential setting.

The first research hypothesis was that the Affective scale of the DASH would discriminate between clients who have a psychiatric diagnosis of bipolar disorder and those who do not. A significantly higher mean affective score on the DASH for the bipolar-disordered group than the mean affective scores for either of the other treatment groups was demonstrated. This indicates that the bipolar-disordered group had more affective symptomatology than the other groups.

The second major research hypothesis was that the Affective scale of the DASH would discriminate between a mixed-psychiatric-disordered group and a group without a psychiatric diagnosis of any type. The Affective scale of the DASH failed to yield a mean affective frequency score for the mixed-psychiatric-group which was significantly higher than the mean affective score of the no-psychiatric-diagnosis group. In addition to the main hypotheses, a variety of post hoc comparisons were conducted. The Affective scale was divided into its two components, namely the mania and depression scales. These subscales demonstrated different levels of sensitivity than the Affective scale demonstrated across groups.
The results of the one-way ANOVA conducted on the mean frequency scores for frequency of affective symptomatology for the three treatment groups indicated a statistically significant difference between groups one and three only. Based on this result it does not appear that the DASH is sensitive enough to effectively discriminate between the bipolar disordered group and the group containing a variety of psychiatric disorders (group two).

An alternative way of describing the results of the statistical analysis is that the mean frequency score for affective symptomatology for the bipolar-disordered group was significantly higher than the mean scores for either of the other groups. In addition, the mean affective symptomatology frequency score for the mixed-psychiatric-disordered group was not significantly higher than the mean score for the no-psychiatric-diagnosis group, and was not significantly lower than the bipolar-disordered group.

A variety of other comparisons were also conducted. The affective scale was divided into the subscales of mania and depression, and one-way ANOVAs were run. These were conducted to assess for any differences in the sensitivity of the mania and depression scales.

The mania scale was demonstrated to be a better discriminator than the depression scale. The mania scale was sensitive enough to discriminate the control group from the psychiatric and the bipolar groups. In addition, the mania scale was nearly able to discriminate between the control group and the mixed psychiatric disordered group.

The mania scale was not able to distinguish between the mixed psychiatric disordered and the bipolar disordered groups.
Comparing the depression scale across groups produced a significant effect, but only between the control group and the bipolar disordered group. In addition, the depression scale was shown to be unable to discriminate these groups at a level of great enough statistical significance.

The depression scale was unable to distinguish between the mixed psychiatric diordered group and the bipolar disordered group, i.e. groups two and three.

Therefore, the affective scale generally, and the depression and mania subscales specifically, were not found to be capable of discriminating bipolar disordered participants from those who had a variety of psychiatric disorders.

In addition to examining the potential contribution of these subscales of the Affective scale of the DASH, post hoc comparisons designed to analyze the potential contribution of demographic factors were considered. Three demographic factors were chosen for further analysis. They were gender, the length of time the informant had known the client, and . These three covariates were chosen based on Matson's findings (Matson, et. al., 1991b)

Implications

The results of this concurrent validation study indicate that the DASH has merit in differentiating the presence of psychiatric disorder, specifically bipolar disorder, between groups of severely and profoundly mentally retarded clients in a residential setting. This is a hopeful sign for clinicians working with this population.
There are a number of considerations when discussing the usefulness of the DASH in clinical practice. The DASH is probably most useful when used for intake evaluations where the referral question concerns whether or not psychiatric disturbance is present. Similarly, it would be useful when there has been a substantive clinical change, e.g. the client is no longer receiving psychotropic medication. This is due to the apparent lack of sensitivity to specific diagnostic categories. That is, the DASH appears capable of discriminating the presence of an affective disorder in general, but not it does not appear able to discriminate bipolar disorder in particular.

Due to this lack of specificity, the DASH would not be expected to be particularly useful in ongoing clinical assessments, e.g. annual assessments to decide whether the client remains eligible for continued Medicaid funding.

The generalizability of the results is further limited by the confounding variable of the diagnosis having been made when the client was already receiving psychotropic medication. Most commonly with this population, psychiatric assessments are conducted at least annually as part of the application process for continued Medicaid funding. Most often the client is receiving medication as part of the treatment approach. The confounding effect of medication already being present when the diagnostic assessment is conducted is a two-pronged issue. First, psychiatric symptoms may be suppressed by the presence of the medication, making it more difficult to accurately diagnose. Second, the medication may produce effects that mimic psychiatric symptoms, e.g. increased akathesia appearing as a manic symptom.
Since the DASH is intended to be used as part of the diagnostic process, the reliability of the diagnoses used as the standard for comparison is of significant importance. Although the reliability of the vignette assessments by the two psychiatrists who made the diagnoses used as the standard was high enough it is not clear how valid diagnoses actually are. Also, there are inherent difficulties in diagnosing severely/profoundly retarded people in general. Why? Therefore, there are two possible confounding variables. First, the adequacy of the original diagnoses is difficult to ascertain. Second, even given the best of circumstances, diagnosis is difficult with this population.

In addition, the participant sample used in this study may not be representative of other settings, even residential settings. For example, FTC may have a higher percentage of bipolar disordered clients than others.

**Suggestions for Future Research**

This concurrent validation of the DASH is a beginning to the overall validation process. More research will need to occur prior to the DASH becoming a standard for psychiatric diagnosis of the severely and profoundly mentally retarded in residential settings. Each of the 12 subscales of the DASH will need concurrent and other types of validation. In addition, samples other than the one developed for this study will need to be researched. Does bipolar disorder demonstrate atypically with this population? Is this population more likely to be manic? Does mania present atypically? (probably so) Why does mania appear to be a more sensitive indicator of the presence of bipolar disorder?

It has been previously documented that the severely/profoundly retarded evidence more organicity. Do I need a cite here? Need to read some
Matson/Barret articles. The manic phase of bipolar disorder is more sensitive to organic factors than the depressive phase. Does that imply a greater proportion of mania? Mania is more endogenous than depression. This population is more cognitively limited, may be a factor in being less depressive than manic. Limited social skills development cite the matson article on the need for social skills development in this population, and more difficult to train in this area. This population tends to be either non-verbal, or severely limited in verbal expression. This is why the DASH was designed as an informant instrument. This lack of verbal output implies that memory is less likely to be encoded verbally. It is more likely that memory is encoded imaginally in this population. ***need to cite the 1982 research Bruce mentioned
References


**Need update?***


Appendix A

Approval Page
Approval

Concurrent Validation of the Affective Scale of the Diagnostic Assessment for the Severely Handicapped

by

Richard M. Ostrom

Signatures:

_____________  ____________Committee Chair  Vice President For Academic Affairs
Explanation of Raw Data

Columns 1-3: Identification Number
Column 5: Race
Column 6: Gender
Column 7: Level of Retardation
Column 8: Physical Disabilities
Column 9: Living Situation
Column 10: Informant's Relationship to Client
Column 11: Length of Time Informant Known Client
Column 12: Amount of Contact on Daily Basis
Column 13: Setting Contact Takes Place in
Column 14: Staffing Ratio
Column 15: How Long Informant has Known Client
Columns 27-28:
Columns 30-33:
<table>
<thead>
<tr>
<th>Column 1</th>
<th>Column 2</th>
<th>Column 3</th>
<th>Column 4</th>
<th>Column 5</th>
<th>Column 6</th>
<th>Column 7</th>
<th>Column 8</th>
<th>Column 9</th>
<th>Column 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>001 12163343225 2220 0000 0000 0000 0000 0000 0000 0000 2210 0000</td>
<td>0000 2210 0001 0000 0000 0000 0000 0000 0000 0000 1210 0000 0000 0000 0000</td>
<td>0000 0000 1210 0000 0000 2210 0000 0000 0000 0000 0000 0000 1220 0000</td>
<td>2210 0000 0000 2200 1220 0000 0000 0000 0000 2200 0000 0000 0000 0000</td>
<td>0000 0000 2210 0000 0000 0000 0000 0000 0000 0000 0000 0000 2210 0000 0000</td>
<td>0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000</td>
<td>0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000</td>
<td>0000 0000 1210 0000 0000 0000 0000 0000 0000 0000 0000 0000 1200 0000 0000</td>
<td>002 12163332235 1220 0000 0001 1210 1210 0000 2210 0000 1200 0000</td>
<td>0000 2210 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000</td>
</tr>
</tbody>
</table>
030 11163342235 2222 2222 0000 2212 2212 0000 2212 0000 2222 0000
0000 2222 0000 0000 0000 1211 0000 0001 0000 2222 2222 1101 2202 0000
0000 0000 2222 0000 1201 0000 2222 2202 2222 0000 0000 2222 2212 0000
2202 0000 0000 0000 0000 0000 0000 0000 0000 2222 0000 0000 2222 0000
0000 0000 0000 0000 0000 1001 1211 0000 0000 0000 0000 0000 0000 0000
0000 0000 0000 0000 0000 0000 0000 0000 0000 2222 0000 0000 2222 0000 0001
0000 1201 0000 0000 0000 2222 2222
031 31133353235 1110 0001 0002 0000 2210 0000 0000 0001 0000 0000
0000 1210 0000 1210 0000 0000 0000 0000 0000 0000 1220 2200 0000 0000
0000 0000 0001 0000 0000 0000 0001 2200 0000 0001 0000 0000 0000 2210
0000 0000 0000 0000 0000 0000 0000 0000 0000 1210 2200 0000 0000 0000
0000 0000 0000 0000 0001 0000 0000 0001 0000 1010 0002 0000 0000
0000 0000 0000 1210 0000 0000 0000 0001 0000 0000 0000 0000 1210
0000 0000 0000 0000 1210 0000 0001
032 12163342234 1201 0000 0000 0000 0000 0000 0000 0000 0000 1211 0000 0000
0000 0000 0000 2202 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000
0000 0000 1001 0000 0000 0000 0000 0000 0000 0000 0000 0000 2212
0000 0000 0000 0000 0000 0000 0000 0000 0000 2212 0000 0000 0000
0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000
0000 0000 0000 0000 0000 0000 0000 0000 0000 1211
0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000
VITA

Richard M. Ostrom  Age:  43
8817 SE Morrison  Marital Status: Married
Portland, OR 97216  Health: Excellent
(503) 257-8345

EDUCATION

Doctor of Psychology:  (Anticipated) (1995); George Fox College, Newberg, Oregon.

Master of Arts in Clinical Psychology:  (1980-1983); Western Baptist Seminary, Portland, Oregon.

Bachelor of Arts in Psychology:  (1969-1973); University of Minnesota, Minneapolis, Minnesota.

Graduate Studies in Clinical Psychology:  (1973-1974); Mankato State University, Mankato, Minnesota.  No degreee awarded.


INTERNSHIP

1991-93:  Counseling Center of Vancouver, Vancouver, WA
  Supervisor:  Dennis Olson, Ph.D.

PRACTICUM EXPERIENCE IN PSYCHOLOGY
1982-84: Experiences included didactic sessions, group supervision, and individual supervision. Psychotherapy with geriatric clients, adults, adolescents, and children was supervised by various psychologists. Experience was gained in psychotherapy, psychological testing, and report writing.

**DISSERTATION**

"Concurrent Validation of the Affective Scale of the Diagnostic Assessment for the Severely Handicapped (DASH) Scale".

**PUBLICATIONS**


**REFERENCES**

Available upon request.
WORK EXPERIENCE

Clinical Psychologist I:  (1985-present)  I have been employed at Fairview Training Center providing psychological services to developmentally disabled adults.  I have served as the chairman of the Psychotropic Review Committee, and as a member of the Behavior Modification Review Committee.

Phone Duty Counselor:  (1983-84)  Clackamas County Mental Health Center employed me as the first line of contact for potential clients.  I was responsible for diagnostic triage using DSM-III criteria for admission into the system.  I also conducted counseling sessions with walk-in clients.

Data Tables

Table: Affective X No-psychiatric X Bipolar-disordered

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>2</td>
<td>.4084</td>
<td>4.6084</td>
<td>.0134</td>
</tr>
<tr>
<td>Within Groups</td>
<td>66</td>
<td>2.9242</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>68</td>
<td>3.3325</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Tukey - HSD Procedure

Mean       Group 1 2 3

.3168       1
.4141       2
.5052       3  *

* Denotes pairs of groups significantly different at the .050 level
<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main Effects</td>
<td>.408</td>
<td>2</td>
<td>4.608</td>
<td>.013</td>
</tr>
<tr>
<td>GROUP</td>
<td>.408</td>
<td>2</td>
<td>4.608</td>
<td>.013</td>
</tr>
<tr>
<td>Explained</td>
<td>.408</td>
<td>2</td>
<td>4.608</td>
<td>.013</td>
</tr>
<tr>
<td>Residual</td>
<td>2.924</td>
<td>66</td>
<td>.044</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>3.333</td>
<td>68</td>
<td>.049</td>
<td></td>
</tr>
</tbody>
</table>

Correlations: MANIA  DEPRESS  AFFECTIVE

D2   .1374  .2145  .1990  Group 2 only
D12  .2797  .2777  .3293
D7   -.2041 -.2640 -.2678

Correlations: MANIA  DEPRESS  AFFECTIVE

D2   -.1400 .0351  .0063  Group 3 only
D12  -.0844 .0716  .0681
D7   -.1003 .2931  .2577

D2   -.0357 .0961  .0516  All groups
D12  .0105 .1714  .1431
D7   -.0752 .0544  .0355

Table  Analysis of Variance  Affective X Psychiatric - disordered X Bipolar - disordered

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>.0954</td>
<td>1</td>
<td>2.0034</td>
<td>.1640</td>
</tr>
</tbody>
</table>
Table Analysis of Variance  Affective X No - Psychiatric X Bipolar - disordered X Medication

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Covariates</td>
<td>30.017</td>
<td>1</td>
<td>1.557</td>
<td>.217</td>
</tr>
<tr>
<td>D12</td>
<td>30.017</td>
<td>1</td>
<td>1.557</td>
<td>.217</td>
</tr>
<tr>
<td>Main Effects</td>
<td>182.502</td>
<td>2</td>
<td>4.718</td>
<td>.012</td>
</tr>
<tr>
<td>GROUP</td>
<td>182.502</td>
<td>2</td>
<td>4.718</td>
<td>.012</td>
</tr>
<tr>
<td>Explained</td>
<td>212.610</td>
<td>3</td>
<td>3.665</td>
<td></td>
</tr>
<tr>
<td>Residual</td>
<td>1257.043</td>
<td>65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1469.652</td>
<td>68</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table Analysis of Variance  Mania X No - Psychiatric X Bipolar - disordered X Medication

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Within Groups</td>
<td>2.096</td>
<td>44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>2.191</td>
<td>45</td>
<td></td>
<td>.049</td>
</tr>
</tbody>
</table>

Main Effects    .095  1  2.003  .164
GROUP           .095  1  2.003  .164
Explained       .095  1  2.003  .164
Residual        2.096  44  .048
Total           2.191  45  .049
<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Covariates</td>
<td>23.816</td>
<td>1</td>
<td>2.196</td>
<td>.143</td>
</tr>
<tr>
<td>D12</td>
<td>23.816</td>
<td>1</td>
<td>2.196</td>
<td>.143</td>
</tr>
<tr>
<td>Main Effects</td>
<td>82.332</td>
<td>2</td>
<td>3.796</td>
<td>.028</td>
</tr>
<tr>
<td>GROUP</td>
<td>82.332</td>
<td>2</td>
<td>3.796</td>
<td>.028</td>
</tr>
<tr>
<td>Explained</td>
<td>106.148</td>
<td>3</td>
<td>3.263</td>
<td>.027</td>
</tr>
<tr>
<td>Residual</td>
<td>704.925</td>
<td>65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>811.072</td>
<td>68</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The following are one-way ANOVAs containing data on frequency for the previous two weeks as well as any occurrence during the entire year.

Table  Analysis of Variance  Mania X No - Psychiatric X Bipolar - Disordered

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>87.352</td>
<td>2</td>
<td>6.251</td>
<td>.0036</td>
</tr>
<tr>
<td>Within Groups</td>
<td>384.300</td>
<td>55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>471.6552</td>
<td>57</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Tukey - HSD Procedure

<table>
<thead>
<tr>
<th>Mean</th>
<th>Group</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.7500</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.8500</td>
<td>2</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.6667</td>
<td>3</td>
<td>*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Denotes pairs of groups significantly different at the .050 level
Table  Analysis of Variance  Affective X Psychiatric - Disordered x Bipolar - Disordered

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>182.0690</td>
<td>2</td>
<td>3.3694</td>
<td>.0417</td>
</tr>
<tr>
<td>Within Groups</td>
<td>1486.0000</td>
<td>55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1668.0690</td>
<td>57</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Tukey - HSD Procedure

Mean Group 1 2 3

6.2000 1
9.4000 2
10.3333 3 *

* Denotes pairs of groups significantly different at the .0505 level
## Table: Analysis of Variance

### Depression X All Groups

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Groups</td>
<td>106.6957</td>
<td>2</td>
<td>3.5992</td>
<td>.0328</td>
</tr>
<tr>
<td>Within Groups</td>
<td>978.2609</td>
<td>66</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1084.9565</td>
<td>68</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Tukey - HSD Procedure

<table>
<thead>
<tr>
<th>Mean</th>
<th>Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.7826</td>
<td>1</td>
</tr>
<tr>
<td>6.7391</td>
<td>2</td>
</tr>
<tr>
<td>7.7826</td>
<td>3</td>
</tr>
</tbody>
</table>

* Denotes pairs of groups significantly different at the .050
<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>44.1000</td>
<td>1</td>
<td>5.9786</td>
<td>.0192</td>
</tr>
<tr>
<td>Within Groups</td>
<td>280.3000</td>
<td>38</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>324.4000</td>
<td>39</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table Analysis of Variance  Affective X No Psychiatric X Psychiatric - Disordered

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>102.4000</td>
<td>1</td>
<td>3.1030</td>
<td>.0862</td>
</tr>
<tr>
<td>Within Groups</td>
<td>1254.0000</td>
<td>38</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1356.4000</td>
<td>39</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table Analysis of Variance  
Depression X No Psychiatric X Psychiatric - Disordered

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>44.0217</td>
<td>1</td>
<td>2.5014</td>
<td>.1209</td>
</tr>
<tr>
<td>Within Groups</td>
<td>774.3478</td>
<td>44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>818.3696</td>
<td>45</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table Analysis of Variance  
Mania X Psychiatric - Disordered X Bipolar - Disordered

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>6.3184</td>
<td>1</td>
<td>.8866</td>
<td>.3527</td>
</tr>
<tr>
<td>Within Groups</td>
<td>256.5500</td>
<td>36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>262.8484</td>
<td>37</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Table: Analysis of Variance - Affective X Psychiatric - Disordered X Bipolar - Disordered

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>8.2526</td>
<td>1</td>
<td>.2559</td>
<td>.6160</td>
</tr>
<tr>
<td>Within Groups</td>
<td>1160.8000</td>
<td>36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1169.0526</td>
<td>37</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table: Analysis of Variance - Depression X Psychiatric - Disordered X Bipolar - Disordered

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>12.5217</td>
<td>1</td>
<td>.6989</td>
<td>.4077</td>
</tr>
<tr>
<td>Within Groups</td>
<td>788.3478</td>
<td>44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>800.8696</td>
<td>45</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>