The Use of the Rorschach Comprehensive System as an Assessment of Depression in Adolescents

Robert G. Van Noord
THE FLORIDA STATE UNIVERSITY

COLLEGE OF EDUCATION

THE USE OF THE RORSCHACH COMPREHENSIVE SYSTEM

AS AN ASSESSMENT OF DEPRESSION

IN ADOLESCENTS

By

ROBERT G. VAN NOORD

A Doctoral Dissertation submitted to the Department of Educational Psychology and Learning Systems in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Degree Awarded:
Spring Semester, 2006

Copyright ©2006
Robert G. Van Noord
All Rights Reserved
The members of the Committee approve the dissertation of Robert G. Van Noord
defended on December 12, 2005.

______________________________
Frances Prevatt
Professor Directing Dissertation

______________________________
Thomas Joiner, Jr.
Outside Committee Member

______________________________
Gary Peterson
Committee Member

______________________________
F. Donald Kelly
Committee Member

Approved:

______________________________
Frances Prevatt, Chair, Department of Educational Psychology and Learning Systems

The Office of Graduate Studies has verified and approved the above named committee
members.
I dedicate this dissertation to my wife, Cheryl. Your love, patience, understanding, encouragement, endless support, knowledge and insight made it possible to complete this project. Words cannot sufficiently describe the feelings, respect, and admiration I have for you. I remain grateful to you and for you, and love you with all that I am.

I also dedicate this dissertation to my parents, Robert and Sally Van Noord, who have ceaselessly believed in me, encouraged me, and supported me throughout this process and throughout my life. Dad, you are my role model, mentor, and support. Mom, you are my comfort, my guide, and my confidante. You are both inspirations to me; your insight, understanding, and emotional and conventional wisdom paved the way for my completion of this project. Among so many other things, thank you for teaching me through your actions and words to believe in, trust in, and challenge myself. I love you both.
ACKNOWLEDGEMENTS

I gratefully acknowledge the guidance, assistance, and generosity of Dr. Larry Kubiak, without which this study would not have been possible. I also express my gratitude and respect to Dr. Frances Prevatt for her corroboration, direction, support, and encouragement throughout the course of my graduate education. Further, my heartfelt appreciation is due to Dan Ward and Stephen Hill for their with technical, procedural, and editing assistance, as well as their ample moral support, advice, and friendship.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of Tables</td>
<td>viii</td>
</tr>
<tr>
<td>Abstract</td>
<td>ix</td>
</tr>
<tr>
<td><strong>1. INTRODUCTION</strong></td>
<td>1</td>
</tr>
<tr>
<td>Social Significance of the Study</td>
<td>3</td>
</tr>
<tr>
<td>Statement of the Problem</td>
<td>5</td>
</tr>
<tr>
<td>Purpose of the Study</td>
<td>7</td>
</tr>
<tr>
<td>Research Questions</td>
<td>7</td>
</tr>
<tr>
<td>Assumptions</td>
<td>9</td>
</tr>
<tr>
<td>Definition of Terms</td>
<td>10</td>
</tr>
<tr>
<td><strong>2. REVIEW OF THE LITERATURE</strong></td>
<td>12</td>
</tr>
<tr>
<td>Introduction</td>
<td>12</td>
</tr>
<tr>
<td>Historical perspectives on childhood and adolescent depression</td>
<td>12</td>
</tr>
<tr>
<td>Criteria for assessing depressive disorders</td>
<td>15</td>
</tr>
<tr>
<td>Current Epidemiology of Depression</td>
<td>17</td>
</tr>
<tr>
<td>Current Problems in the Conceptualization of Depression</td>
<td>18</td>
</tr>
<tr>
<td>Categorical vs. Dimensional Methodology</td>
<td>19</td>
</tr>
<tr>
<td>The Problem of Comorbidity</td>
<td>20</td>
</tr>
<tr>
<td>Developmental Considerations</td>
<td>22</td>
</tr>
<tr>
<td>Problems in the Assessment of Depression</td>
<td>24</td>
</tr>
<tr>
<td>Internalizing vs. Externalizing disorders</td>
<td>24</td>
</tr>
<tr>
<td>Categorical vs. Dimensional Methodologies</td>
<td>26</td>
</tr>
<tr>
<td>Integration of Multiple Assessment Modalities and Respondents</td>
<td>27</td>
</tr>
<tr>
<td>Cognitive Theoretical Perspectives on Child and Adolescent Depression</td>
<td>30</td>
</tr>
<tr>
<td>Beck’s Theory of Depression</td>
<td>31</td>
</tr>
<tr>
<td>The Learned Helplessness Model</td>
<td>33</td>
</tr>
<tr>
<td>The Tripartite Model of Depression</td>
<td>34</td>
</tr>
<tr>
<td>Cognitive Information Processing and Resource Allocation Models</td>
<td>35</td>
</tr>
<tr>
<td>Commonly Used Self-Report-Format Instruments</td>
<td>37</td>
</tr>
<tr>
<td>The Children’s Depression Inventory</td>
<td>38</td>
</tr>
<tr>
<td>The Reynolds Child Depression Scale (RCDS)</td>
<td>39</td>
</tr>
<tr>
<td>The Reynolds Adolescent Depression Scale (RADS)</td>
<td>40</td>
</tr>
<tr>
<td>Alternative Assessment: The Rorschach Comprehensive System (CS)</td>
<td>42</td>
</tr>
<tr>
<td>Historical and Present Use of the Rorschach</td>
<td>43</td>
</tr>
<tr>
<td>Overview of The Exner Comprehensive System (CS)</td>
<td>44</td>
</tr>
<tr>
<td>Administration of the CS</td>
<td>44</td>
</tr>
<tr>
<td>Scoring the CS</td>
<td>45</td>
</tr>
<tr>
<td>Interpreting the CS: the structural summary</td>
<td>46</td>
</tr>
<tr>
<td>Theoretical Basis and Applications</td>
<td>46</td>
</tr>
<tr>
<td>Perception and Cognition</td>
<td>47</td>
</tr>
<tr>
<td>The Limit of Constructivism</td>
<td>48</td>
</tr>
<tr>
<td>The Depression Index (DEPI)</td>
<td>48</td>
</tr>
</tbody>
</table>
### 3. METHODOLOGY

**Statement of the Hypotheses**

**Population**

**Sample**

**Research Design**

**Variables**

- DEPI components
- CDI components
- MMPI-A scale 2 (Depression scale)
- Reynolds Adolescent Depression Scale (RADS)

**Instrumentation**

- Rorschach Inkblot Test
- MMPI-A
- RADS

**Procedures**

**Data Analysis**

- Analyses set 1
- Analyses set 2
- Analyses set 3
- Analyses set 4

### 4. RESULTS

**Introduction**

**Sample demographic Data**

**Results for Analyses set 1**

- Hypotheses 1 through 3

**Results for Analyses set 2**

- Hypothesis 4
  - Pearson product moment results
Hypothesis 5  
Results of analyses of variance  
Results of Data Analyses Set 3  
Hypothesis 6  
Diagnostic efficiency results  
DEPI  
CDI  
Combined DEPI and CDI  
Results of Data Analyses Set 4  
Preliminary Analysis  
Results of Hierarchical Logistic Regression Analysis 1  
Results of Hierarchical Logistic Regression Analysis 2  
Hypothesis 7  

5. DISCUSSION  
Summary of Results  
Interpretation of Results  
Differences in Scores by Diagnosis  
Clinical vs. Non-Clinical  
Clinical Depressed vs. Clinical Non-Depressed  
Overall Interpretation  
Relationship Between DEPI and CDI  
Main Effect of Age On DEPI and CDI Scores  
Diagnostic Efficiency of the DEPI and CDI  
DEPI  
CDI  
Combined DEPI and CDI  
Comparison with other research  
Incremental Validity of the DEPI and CDI  
Clinical Implications  
Recommendations for Future Research  
Limitations of the Study  
Conclusions  

APPENDIX  
REFERENCES  

BIOGRAPHICAL SKETCH
LIST OF TABLES

1. Demographic Characteristics of Clinical Sample 90
2. Total Sample By Diagnosis and Age 91
3. Non-Depressed Psychiatric Sample by Age and Diagnosis 91
5. T-Scores, Means, and Standard Deviations For CDI Pair-Wise Comparisons 95
6. Pearson Correlation Coefficients, Means, & Standard Deviations for DEPI and CDI Total Score Comparisons 96
7. Descriptive Data for ANOVA Results, Combined Clinical Groups 97
8. Diagnostic Efficiency Statistics for Different DEPI Total Score Cutoffs 101
9. Diagnostic Efficiency Statistics for Different CDI Total Score Cutoffs 101
10. Diagnostic Efficiency Statistics for Different DEPI and CDI Total Score Combined Cutoffs 102
11. Pearson and Point Biserial Correlation Coefficients, Means, & Standard Deviations Among MMPI-A, RADS, DEPI, CDI, and Diagnosis 106
12. Logistic Hierarchical Regression Analysis 1, with Conditional Variable Entry of Variables in the Following Order: RADS, MMPI-A, DEPI, CDI 109
13. Logistic Hierarchical Regression Analysis 2, with Forced (i.e., Non-Conditional) Variable Entry in Order of Appearance 110
ABSTRACT

The current study was conducted to evaluate the overall effectiveness of the Rorschach Comprehensive System’s (CS) (Exner, 1993) Depression Index (DEPI) and Coping Deficit Index (CDI) in assessing depression among a clinical sample of adolescents. DEPI and CDI archival data collected from roughly 340 inpatient adolescents were compared via independent, pair-wise t-tests with data from Exner’s (1993) non-clinical, CS normative sample of roughly 450 adolescents to investigate degree of co-relationship for both DEPI and CDI by diagnostic group (i.e., clinical depressed, clinical non-depressed, and non-clinical). Pearson coefficients assessed the degree of co-relationship between DEPI and CDI total scores, and ANOVA analyses tested the overall effect of age. Diagnostic efficiency coefficients were calculated for clinical DEPI and CDI scores, and variable combinations based on Exner’s (1993) cutoff points. Also, the incremental validity of the DEPI and CDI towards predicting depression was evaluated using hierarchical logistic regression, relative to the contributions of the Minnesota Multiphasic Personality Inventory-Adolescent (MMPI-A) (Butcher, Williams, Graham, Archer, Tellegen, Ben-Porath, and Kaemmer, 1992) and the Reynolds Adolescent Depression Scale (RADS) (Reynolds, 1987). For the most part, t-test results revealed no significant differences in DEPI total scores between clinical and non-clinical adolescents, or depressed-only and non-depressed clinical adolescents. Both depressed and non-depressed clinical groups demonstrated significantly higher CDI scores than the non-clinical adolescent group. When DEPI and CDI scores were compared, significant correlations were observed for the depressed-only group and combined depressed and non-depressed clinical adolescents, but not for non-depressed clinical adolescents. No significant main effect of age was identified on either DEPI or CDI total scores. Overall, the DEPI yielded diagnostic efficiency considerably inferior to that anticipated by Ganellen’s (1996) roughly comparable study involving adults. Overall, the CDI did not contribute meaningfully towards enhancing the diagnostic efficacy of the DEPI, as Exner (1993) has predicted. Findings provided partial corroboration of Archer and Krishnamurthy’s (1997) study in indicating that neither the DEPI nor CDI total scores appeared to produce incremental increases towards
improvement in diagnostic classification. Limitations of the current study, implications for the assessment of depression for adolescent populations, and recommendations for future research are discussed.
CHAPTER 1

INTRODUCTION

Within the past 20 years, researchers and clinicians have come to appreciate the prevalence and role of depressive disorders among children and adolescents, rather than among exclusively adult populations (Parry-Jones, 1995). Further, the unique developmental variability of depressive symptoms exhibited by children and adolescents has gained greater attention (Cicchetti & Toth, 1998; Hammen & Rudolph, 1996; Weiss, Weisz, Politano, Carey, Nelson, & Finch, 1991; Worchel, Hughes, Hall, Stanton, Stanton, & Little, 1990). The current Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) (American Psychiatric Association, 1994) includes some age-based allowances for depressed children and adolescents, and a multitude of different assessment instruments have been developed to better assess depressive symptomatology; however, numerous problems in the assessment of depressive disorders persist.

The practice of assessing depression continues to be hindered by the inability to accurately classify depression as a discrete categorical disorder (Poznanski & Mokros, 1994; Parry-Jones, 1995), as well as by problems inherent in dimensional assessment methodology (Angold & Costello, 1995; Cytryn & McKnew, 1974; Hammon & Rudolph, 1994). These problems have been compounded by factors such as lack of empirical research (Angold & Costello, 1995; Weiner, 1989), depression’s high rate of comorbidity with other disorders (Hammen & Compas, 1994; Hammen & Rudolph, 1996), the difficulty of differentiating internalizing and externalizing symptoms of depression (Compas, Ay, and Grant, 1993), and discrepancies in child/adolescent vs. adult manifestation of depressive symptoms (Cicchetti & Toth, 1998).

Perhaps it is not surprising, given these confounds, that attempts to integrate disparate dimensional assessments of depression often result in lack of convergence (Kazdin, 1987; Reynolds and Johnston, 1994; Worchel, Little, & Alcala, 1990); however, psychometric shortcomings of dimensional assessments, themselves, have also been implicated. For one, it has been posited that different instruments measure different constructs, as opposed to the same construct at differing levels of severity and specificity.
(Hammen and Rudolph, 1996). Scores from instruments completed by different raters (i.e., teachers, parents, peers, and self) also tend to demonstrate limited relationship (Asarnow & Carlson; Kazdin, 1987), which may be attributed to rater psychopathology, as well as measurement error (Cole, Hoffman, Tram, & Maxwell, 2000; Crowley and Worchel, 1993; Hammon & Rudolph, 1996; Joiner, Catanzaro, & Laurent, 1996; Worchel et al., 1990).

Overall, no solitary dimensional instrument, administered independent of categorical criteria, has been found capable of validly indicating the presence of depression in children and adolescents (Kazdin, 1987; Reynolds and Johnston, 1994; Worchel et al., 1990); however, a number of popular self-report instruments, such as the Children’s Depression Inventory (Kovacs, 1992) and the Reynolds Child Depression Scale (Reynolds, 1989a) have been found to demonstrate acceptable psychometric stability and validity. Because a higher percentage of individuals receive diagnoses of depression as a result of being administered a full psychometric battery of multiple instruments (i.e., assessment of broad psychopathology based on multiple domains and informants) this type of assessment has become common practice.

Alternative instruments, including the Rorschach Inkblot Test, have also been developed in attempt to develop more effective and conclusive means of assessing the presence of depression among children and adolescents (Cole, Martin, Peeke, Seroczynski, & Hoffman, 1998; Garber & Kaminski, 2000; Joiner & Lonigan, 2000; Stark, Schmidt, & Joiner, Jr., 1996). Although the Rorschach has been administered and interpreted in a variety of controversial ways, numerous clinicians currently use the Exner Comprehensive System (CS) (Exner, 1993) to assess broad personality characteristics, as well as specific psychopathology (Blatt, 1990; Exner, 1988; Weiner, 1997). Scored according to the CS, the Rorschach is a popular instrument that represents a potentially valuable, and oft-neglected complement to criterion-based assessment.

Despite being one of the most commonly used instruments in clinical facilities and graduate academic training settings (McCann, 1998; Weiner, 1997), the Rorschach has been widely criticized due to its purported lack of theory base (Weiner, 1989; 1997) and controversy regarding its psychometric stability (Garb, Wood, Nezworski, & Grove, 2001; Wood, Lilienfeld, Garb, & Nezworski, 2000; Wood, Nezworski, & Stejskal, 1997).
Ongoing controversy continues to surround the Rorschach CS (Archer & Meyer, 2001; Viglione and Hilsenroth, 2001). Its use has been supported by judicial precedents through numerous court cases (McCann, 1998; Meloy, Hansen, & Weiner, 1997; Weiner, Exner, & Sciara, 1996); however, it remains uncertain whether the CS represents a valid instrument with which to assess depression, particularly among populations of children and adolescents. For one, there is a lack of published normative data for clinical samples of depressed children, as well as limited normative data pertaining to adolescents administered the CS. Moreover, the Rorschach’s predictive, convergent, incremental, and overall psychometric validity remain largely unexplored with regard to assessing depression in relation to these populations. Pertaining to subjects exhibiting categorically viable symptoms of depression, how well does the Rorschach differentiate children and adolescents from their non-depressed counterparts? How accurate and consistent is the Rorschach in identifying depressed non-adults across childhood and adolescence? Does the Rorschach contribute meaningful incremental data towards formulating valid depressive diagnoses? These questions warrant further study, and will serve as the basis of this study.

Social Significance of the Study

Contrary to past theoretical assumptions, current epidemiological research indicates that childhood and adolescent depression represents a significant concern to society. Throughout the United States, approximately two to six percent of children, and approximately ten to twenty percent of adolescents have been estimated to exhibit sufficient depressive symptoms to constitute a categorically validated depressive disorder (Poznanski & Mokros, 1994; Reynolds and Johnston, 1994; Snyder, 1999; Steele, Armistead & Forehand, 2000). Further, among depressed individuals referred for inpatient treatment, estimates as high as 59 percent have been cited pertaining to children, and as high as 27 percent for adolescents (Poznanski & Mokros, 1994). Silver (as cited in Reynolds & Johnston, 1994) has found that nearly one out of every six juveniles admitted to psychiatric hospitals in the United States receive a diagnoses of a depressive disorder.
Many detrimental aspects of child and adolescent depression comprise an overt, integral part of DSM-IV criteria (APA, 1994) for depressive disorders. While depression’s direst outcome remains heightened probability of suicide, non-lethal symptoms of depression significantly also detract from the health, well-being, and functioning of children and adolescents. Irrespective of the age at which depression occurs, its repercussions are evident; however, contrasted with adults, depressive disorders during childhood and adolescence have the potential to impede intellectual, affective, and social functioning, and hence retard and/or impede “normal” development within these realms (Cicchetti, Rogosch, & Toth, 1994; Cicchetti & Toth, 1998).

Similarly, depression occurring during youth has the potential to impede academic progress, impeding and/or limiting future educational and vocational progress across essential social, emotional, cognitive, educational, and developmental areas of life. Further, early treatment intervention during childhood and adolescence can forestall impairment in recurrence and/or exacerbation of depressive disorders across the lifespan (APA, 1994; Kovacs, 1996; Masi, Favilla, Mucci, Poli, & Romano, 2001).

From a sociological perspective, non-adult depression can be conceptualized as a detriment to national industrial productivity. Given the prevalence of child and adolescent depressive disorders, the financial costs to society of providing treatment interventions (i.e., psychological, pharmacological, and possibly medical as a result of self-harm) are undoubtedly high. Finally, the emotional and fiscal ramifications of depressive disorders for the parents and families of children and adolescents cannot be understated, and further detract from national productivity (i.e., increased family and marital stress, resulting absences from work, financial costs, etc.).

Thus, the professional problem addressing the social problem of non-adult depression consists of the need to accurately and consistently identify depressive disorders as they occur among children and adolescents. Among psychological professionals, this problem is commonly addressed via the use of dimensional assessment instruments. These instruments must be capable of consistently and accurately signaling the presence of depressive symptoms that are sufficiently persistent and severe to warrant categorical diagnoses of depression. Valid and reliable assessment is necessary to ensure predictable response to treatment (Trad, 1987), including the application of specific...
forms of treatment that effectively differentiate and target depressive symptoms (Goodman, Schwab-Stone, Lahey, Shaffer, & Jensen, 2000).

**Statement of the Problem**

The Rorschach Inkblot Test, administered and scored according to the guidelines of the Exner Comprehensive System (CS; Exner, 1993), has been found to be one of the most prominent psychometric instruments utilized in empirical research studies, graduate training programs, and clinical settings (McCann, 1998; Weiner, 1997). The CS is administered most often to evaluate broad personality characteristics (Blatt, 1990; Exner, 1988; Weiner, 1997); however, within clinical settings the CS Depression Index (DEPI) and Coping Deficit Index (CDI) commonly are used as criterion-based measures to assess the presence and severity of depression.

Virtually no empirical research has evaluated the psychometric validity of the CS depression indices using populations of children and adolescents (Exner, 1993; 1997; Exner & Weiner, 1995; Weiner, 2001); however, considerable evidence indicates that the CS is routinely administered to these populations in a wide range of clinical settings (Exner, 1993; Exner & Weiner, 1995; Finch, Imm, & Belter, 1990; L. Kubiak, personal communication, June 2000). This is problematic, as even the most recent versions of the DEPI and CDI were standardized using a non-clinical (i.e., non-depressed) sample of children and adolescents (Exner, 1993; Exner & Weiner, 1995). Moreover, among this non-clinical standardization sample, the specified number of child and adolescent participants within each age group varied considerably, and was unacceptably small for several of the represented ages.

For both children and adolescents, research is needed to compare Exner’s (1993) descriptive data with clinically depressed Rorschach responses (i.e., the DEPI and CDI) using an acceptably large sample and categorically validated diagnoses. To date, the only Rorschach study providing published descriptive data for depressed non-adult subjects involved 100 depressed adolescents diagnosed using DSM-III (American Psychiatric Association, 1980) criteria. This study was conducted at the time of the most recent revision of the DEPI and CDI (Exner & Weiner, 1995). Research is also needed to address the significant problem of depression’s high comorbidity rate by comparing
DEPI and CDI scores provided by depressed non-adults, as compared with non-adults diagnosed with other disorders commonly found to be comorbid with depression.

Reconciling the results of dimensional assessment instruments with categorical diagnoses of depression (i.e., diagnoses based on the DSM-IV; APA, 1994) is also problematic, as these distinct methodologies often yield discrepant results (Hammon & Rudolph, 1994; Kazdin, 1987; Sherak, Speier & Cantwell, 1994; Weiner, 1989; Worchel, Little, & Alcala, 1990). In other words, dimensional instruments designed to measure depression often exhibit poor predictive and convergent validity. The limited number of studies that have examined the predictive validity of the DEPI yielded similarly disappointing results (Meyer & Archer, 2001; Wood, Lilienfeld, Garb, and Nezworski, 2000; Wood, Nezworski, and Stejskal; 1997), and almost all of these studies failed to include non-adult subjects.

Further, research is needed that explicitly addresses the role of the CDI in Rorschach CS studies, as only Exner’s (1993) CDI validation study has examined the predictive and/or incremental validity of this index (Exner & Weiner, 1995). This research found that the CDI increased the predictive validity of the CS in 79% of clinically depressed subjects exhibiting false negative results based only the DEPI; however, this study was limited to adults diagnosed with depressive disorders according to DSM-III (APA, 1980) criteria. Research is needed to validate these promising results for clinically-referred populations of depressed children and adolescents.

Correspondingly, categorical classification of depression encompasses a broad range of symptom typologies, not all of which must be observed to administer a diagnosis of depression (Rudolph, 1994; Sherak et al., 1994; Weiner, 1989). For example, DSM-IV (APA, 1994) criteria for Major Depressive Disorder necessitates the presence of five or more of the following types of symptoms: depressed mood, feelings of worthlessness or excessive or inappropriate guilt, diminished ability to think or concentrate/indecisiveness, recurrent thoughts of death, and physiological symptoms such as significant weight loss, insomnia or hypersomnia, psychomotor agitation or retardation, and fatigue or loss of energy (p. 327).

Children and adolescents display some, but not all of the symptoms encompassed by categorical depressive taxonomy to such varying degrees of severity that no solitary
dimensional instrument is felt to be capable of consistently assessing their breadth and depth with sufficient accuracy (Kazdin, 1987; Worchel et al., 1990). As such, respective dimensional instruments are commonly restricted to measuring specific types of depressive symptoms to the exclusion of others, warranting an integrated approach that involves multiple dimensional assessment instruments. Research is warranted to elucidate the role of the Rorschach CS, as a complement to other traditional assessment instruments, in providing incremental data with which to facilitate increasingly valid diagnoses of depression among children and adolescents.

In summary, this study will evaluate the DEPI and CDI to satisfy the following needs: (a) the need to evaluate the strengths, weaknesses, diagnostic efficiency, and predictive validity of these indices with regard to the assessment of adolescent depression; (b) the need to clarify the broad influence of development (i.e., as represented by chronological age) on CS predictive validity; and (c) the need to investigate the influence of comorbid non-depressive disorders on CS depression-related index scores. Moreover, the study will evaluate the incremental validity of the DEPI and the CDI, both independently and co-jointly, in conjunction with the Reynolds Adolescent Depression Scale (RADS) and the MMPI-A depression scale (scale 2).

**Purpose of the Study**

The purpose of this study is to contribute to the understanding of the relationship between Rorschach DEPI and CDI total scores provided by clinical and non-clinical populations of adolescents. It is also conducted to contribute to the understanding of the relationship between chronological age, and non-adult DEPI and CDI responses and total scores. Moreover, this research will contribute to the understanding of the relationship between DEPI and CDI responses and measured depression by evaluating the diagnostic efficiency and incremental validity of both scales.

**Research Questions**

Given the aforementioned statement of the problem, the following research questions have been formulated:
1. Pertaining to total scores of adolescents between the ages of 13-16 on the DEPI and CDI, what is the relationship between non-depressed “normal” controls, depressed psychiatric inpatients, and non-depressed psychiatric inpatients?
   a. In terms of adolescents’ DEPI total scores, what is the relationship between non-depressed “normal” controls, depressed psychiatric inpatients, and non-depressed psychiatric inpatients within each age group (i.e., 13-16 years old)?
   b. In terms of adolescents’ CDI total scores, what is the relationship between non-depressed “normal” controls, depressed psychiatric inpatients, and non-depressed psychiatric inpatients within each age group (i.e., 13-16 years old)?

2. What is the relationship between DEPI and CDI total scores?
   a. What is the relationship between DEPI and CDI total scores for depressed psychiatric adolescents?
   b. What is the relationship between DEPI and CDI total scores for non-depressed psychiatric inpatients?

3. For psychiatric adolescents between the ages of 13-17, what is the relationship between DEPI and CDI total scores and chronological age?
   a. For depressed psychiatric inpatients, what is the relationship between DEPI total scores and age?
   b. For depressed psychiatric inpatients, what is the relationship between CDI total scores and age?
   c. For non-depressed psychiatric inpatients, what is the relationship between DEPI total scores and age?
   d. For non-depressed psychiatric inpatients, what is the relationship between CDI total scores and age?

4. For depressed and non-depressed psychiatric adolescents between the ages of 13-17, what is the relationship between DEPI and CDI total scores and categorical (i.e., DSM-IV-TR) diagnostic criteria?
   a. What is the diagnostic efficiency of the DEPI?
b. What is the diagnostic efficiency of the CDI?
c. What is the diagnostic efficiency of combined DEPI and CDI?

5. For depressed psychiatric inpatient adolescents, what is the relationship between DEPI Total Scores, CDI Total Scores, MMPI-A Scale 2 T-Scores, RADS scores, and categorical (i.e., DSM-IV-TR) diagnostic criteria?
   a. What is the incremental validity of the MMPI-A Scale 2 with respect to DSM-IV diagnoses of depression?
   b. What is the incremental validity of the RADS with respect to DSM-IV diagnoses of depression?
   c. What is the incremental validity of the DEPI, in conjunction with the MMPI-A Scale 2 and the RADS, with respect to DSM-IV diagnoses of depression?
   d. What is the incremental validity of the DEPI and CDI, in conjunction with the MMPI-A Scale 2, with respect to DSM-IV diagnoses of depression?

Assumptions

This study will be conducted under the following broad assumptions:

1. Symptoms of depression manifest uniquely among children and adolescents, as opposed to adults.
2. For individuals of all ages, depression is defined as a discrete disorder on the basis of categorical taxonomy (i.e., DSM-IV).
3. Dimensional and categorical assessment methodologies lack acceptable convergent validity, and are inherently limited to assessing some (i.e., one or more), but not all typologies of depressive symptoms.
4. Only a combination of dimensional and categorical assessment methodologies is capable of validly assessing depression as a discrete disorder; conversely, no unitary dimensional instrument is capable of validly assessing depression as a discrete disorder.
5. Despite measurement error, the Rorschach Comprehensive System (Exner, 1993) and the MMPI-A (Butcher, Williams, Graham, Archer, Tellegen, Ben-Porath, &

**Definitions of Terms**

In order to ensure a common understanding among readers, the definitions below are presented to clarify the terminology used in the study:

**Percept**: the aspects, including characteristics, of Rorschach inkblots perceived and verbalized by subjects in response to viewing the stimulus.

**Categorical Assessment Methodology**: type of assessment designed to verify the presence of a discrete psychological disorder, based on specified criteria (i.e., the DSM-IV) (Kazdin, 1987; Reynolds, 1994).

**Dimensional Assessment Methodology**: type of assessment designed to determine the extent or severity of observed depressive symptoms according to a pre-specified continuum (i.e., self-report instruments, the Rorschach CS) (Kazdin, 1987; Reynolds, 1994).

**Nomothetic interpretation**: a method of test interpretation that empirically evaluates individual performance on a psychometric instrument relative to the performance of normative populations of individuals (i.e., standardized tests) (Weiner, 2000).

**Idiographic interpretation**: a method of test interpretation that non-empirically evaluates individual performance on a psychometric instrument independent of the performance of normative populations of individuals (i.e., projective tests) (Weiner, 2000).

**Sensitivity**: the actual rate at which an assessment tool identifies subjects diagnosed with a particular categorical (i.e., DSM-IV-based) disorder of interest, not taking into account the actual prevalence of the disorder among the population of interest (Ganellen, 1996).

**Specificity**: the actual negative rate at which an assessment tool identifies subjects not diagnosed with a particular categorical (i.e., DSM-IV-based) disorder of interest, not taking into account the actual prevalence of the non-target disorder among the population of interest (Ganellen, 1996).

**False Positive Rate**: the proportion of test subjects who do not meet categorical criteria
for membership within a particular diagnostic group, but who rather have been
incorrectly classified within that group by a particular assessment tool (Streiner,
2003).

**False Negative Rate**: the proportion of test subjects who meet categorical criteria for
membership within a particular diagnostic group, but who have been incorrectly
excluded from membership within that group by an assessment tool (Streiner,
2003).

**Positive Predictive Power**: the ability of an assessment tool to identify subjects
diagnosed with a particular categorical (i.e., DSM-IV-based) disorder of interest,
taking into account the prevalence rate of the target disorder in the overall sample
of interest (Ganellen, 1996).

**Negative Predictive Power**: the ability of an assessment tool to exclude test
subjects not diagnosed with a particular categorical (i.e., DSM-IV-based) disorder
of interest, taking into account the prevalence rate of the target disorder in the
overall sample of interest (Ganellen, 1996).

**Overall Correct Classification Rate**: the total rate at which an assessment tool
is able to correctly identify test subjects as both members and non-members of a
particular diagnostic group (as previously determined by categorical/DSM-IV
criteria) (Streiner, 2003).

**Kappa**: the statistically represented level of agreement between a dimensional assessment
tool and diagnostic (i.e., DSM-IV) criteria beyond that accounted for by chance
alone (Streiner, 2003).
CHAPTER 2

REVIEW OF THE LITERATURE

Introduction

The process, methodology, and theoretical assumptions on which assessment of depression in children and adolescents has been based have undergone dramatic changes within the past twenty years. Rather than advancing and continuing the historical body of knowledge in a linear manner, these recent changes are found to break with historical perspectives in a manner described by Kuhn (1996) as a shift in paradigm. Conceptualized as a change in theoretical knowledge, this paradigm shift has led to changes in the way science has come to understand childhood/adolescent depressive illness. As a result, increasing inclusion of children and adolescents in empirical research has occurred. For one, while retaining various historical theoretical perspectives, research has served to alter current conceptualization of childhood and adolescent depression by challenging formerly-held theoretical beliefs. From a historical perspective, clinicians have often neglected to consider children and adolescents as being potentially depressed individuals. While this exclusion is somewhat surprising, given past documentation of numerous case studies suggestive of the presence and legitimacy of depressive symptoms (Parry-Jones, 1995), Kuhn (1996) notes that this finding is instead typical of the functioning of “normal science.” In other words, the concept of depression in children and adolescents remained incompatible with the largely rigid structure imposed by the diverse theoretical paradigms throughout the past.

Historical Perspectives On Childhood and Adolescent Depression

Publications dating prior to the nineteenth- and twentieth-centuries have acknowledged the presence of depressive symptoms, if not discrete disorders, in children and adolescents. In comparison with modern standards of established empirical methodology, pre-nineteenth century conceptualization of depression in non-adult populations was significantly lacking in empirical rigor. For one, attempts to diagnose depression, regardless of the age of the afflicted, lacked clinical consensus regarding
acceptable diagnostic criteria (Parry-Jones, 1995; Poznanski & Mokros, 1994; Trad, 1987). Such diagnoses also relied exclusively on inductive reasoning, based on clinical observation of a particular patient. Historical literature suggests that general consensus regarding the etiology of depression in children largely implicated parenting practices, such as over-indulgence and excessively stern discipline, and hypothesized genetic transmission of parental dispositions (Burton, as cited in Parry-Jones, 1995).

Additionally, in 1782, Arnold (as cited in Parry-Jones, 1995) wrote that childhood depression stemmed to a large degree from environmental, stress-inducing factors such as parental separation, which he hypothesized resulted in “nostalgic insanity” (p. 6).

Pre-nineteenth century literature also reveals that practitioners struggled to discern pathological, as opposed to ‘normal,’ parameters for defining depression in children and adolescents (Parry-Jones, 1995). As a result, many clinical observations of depressive symptoms were dismissed as typical, albeit unfavorable, elements of maturation. By modern standards, this problem is felt to have resulted from a lack of empirically-based agreement concerning depression’s nosology, including symptom prevalence and epidemiology; however, as will be addressed, similar timeless challenges to accurate assessment of childhood and adolescent depression continue to persist, despite the availability of modern scientific advancements.

Theoretical models improved throughout the nineteenth century, incorporating rough distinctions between adult and childhood course of depression/melancholia. Among other theorists living during that time period, Crichton-Browne (as cited in Parry-Jones, 1995) overtly acknowledged the plausibility of childhood and adolescent depression based on the severity, type, and developmental onset of depressive symptoms. Further, he hypothesized a myriad of potential etiological factors such as physical illness, inherent predisposition, moral disturbances, and separation from parental figures (Parry-Jones, 1995). Interestingly, it is noted that attempts were undertaken as early as 1879 by Maudsley (as cited in Parry-Jones, 1995) to developmentally categorize depressive symptoms by age of onset. Additionally, throughout the latter half of the nineteenth century the causal significance of puberty began to receive increasing attention as a result of its hypothesized physiological contribution towards affective symptoms.
At the turn of the twentieth century, however, the concept of childhood and adolescent depression received little to no support from the prominent clinical community. The emergence of Psychoanalysis as the dominant theoretical modality precluded acceptance of childhood depression as a valid pathological condition. Psychoanalytic theory contends that depression is the result of aggression turned inward against the self; hence, as Rochlin notes (as cited in Hammen & Rudolph, 1996), the presence of depression in children was shunned on the basis that children lack a sufficiently developed superego. Thus, no attempts were undertaken to assess or treat this disorder, as its presence in young children was felt to be impossible. Depression in older adolescents was deemed equivalent to that of adults without heed to age-appropriate, developmental factors. Similar to the informal criteria of the past, psychoanalytic clinicians on the whole deemed that depressive symptoms causing sufficient impairment to the individual’s functioning, albeit for unspecified time or severity, constituted acceptable criteria for diagnosis and treatment of depression (Arlow, 1989).

While psychoanalytic conceptualization of depressive disorders persisted, revisions in clinicians’ theoretical view of childhood depression began to occur. In the 1960’s, differences in symptom expression at different stages of an individual’s development became a focus of increased attention. A then-prominent position, referred to by Trad (1987) as the “Depressive Equivalents” theory, postulated that children overtly express depression in a manner highly atypical than that of adults. While this theory did not exclude “traditional” depressive symptoms more commonly observed in adulthood (e.g., flat or dysphoric affect, anhedonia, etc.), children’s manifestations of depression were thought to consist primarily of externalizing types of behavior such as aggression and hyperactivity, as well as physical symptoms such as enuresis, somatic problems, and sleeping and eating disturbances. For children, these types of symptoms were essentially felt to represent “masked depressions,” or “depressive equivalents” of symptoms traditionally observed in adult depression (Hammen & Compas, 1994; Hammen & Rudolph, 1994; Nurcombe, 1994; Trad, 1987, p. 28).

The depressive equivalents theory held considerable value by delineating the differences between children’s vs. adults’ overt expression of depressive symptoms;
however, a primary criticism of this theory attacked its lack of empirically-based
nosological criteria for discriminating depression from other childhood disorders (Trad,
1987). In particular, this theory held obvious potential for misdiagnosis (i.e., false
positives), resulting in significant questions regarding its validity (Cytryn & McKnew, as
cited in Trad, 1987). Increasingly, contemporary theory began to differentiate itself from
the Depressive Equivalents theory by conceptualizing childhood and adolescent
depression as an internalizing, rather than externalizing disorder (Cytryn & McKnew,
cited in Trad, 1987). Correspondingly, much contemporary research continues to target
internalizing symptoms, rather than the common symptoms depression was previously
thought to share with exclusively externalizing types of disorders (Angold & Costello,
1995; Cicchetti & Toth, 1998; Reynolds, 1990b). Prominent internalizing symptoms for
children and adolescents consist of intense subjective discomfort, lack of feelings of
personal self-worth, and feelings of distress (Reynolds, 1990b; Reynolds & Johnston,
1994).

The need to consider internalizing symptoms of depression has become
undeniable due to these symptoms’ detrimental influence on the subjective self-
perception of the child or adolescent, particularly due to their potentially disastrous
behavioral consequences. Thus, the onus of obtaining an assessment of depression that
accurately accounts for negative self-perception is crucial. It must be noted, however,
that classification of childhood and adolescent depression as a primarily internalizing
disorder does not presuppose the non-existence of externalizing behavior. Current
conceptualization of depression acknowledges that the concept of “masked depression”
represented an erroneous theoretical conclusion due to its unparalleled high comorbidity
with other childhood disorders, particularly externalizing (Hammen & Compas, 1994;
symptoms are found to be strong predictors of externalizing behavior, and note that the
two commonly co-occur.

Criteria for Assessing Depressive Disorders

In the 1980’s, research efforts attempted to establish greater empirical specificity
regarding the stable characteristics and symptoms associated with depressive disorders.
Addressing this problem first in adults, research practitioners began to delineate more concrete diagnostic criteria (Feighner, Robins, Guze, Woodruff, Winokur, & Munoz, cited in Trad, 1987); however, a subsequent attempt was soon made to modify these adult criteria for child populations (Weinberg, as cited in Trad, 1987). Both criteria were found to be highly ineffective, however, as they were based upon commonly agreed-upon observations of symptoms characteristic of clinical depression, rather than empirical research studies (Angold & Costello, 1995). Using the child depression criteria, for example, clinicians were able to identify previously diagnosed cases of depression in only slightly over half of a sample of referred children (Trad, 1987). Both sets of depressive criteria combined (i.e., child and adult) were adapted and became universal for both children and adults within the Diagnostic and Statistical Manual of Mental Disorders, Third Edition (DSM-III) (American Psychiatric Association, 1980).

The revised DSM-III criteria were considerably more effective and reliable than those previously utilized (Trad, 1987). Reynolds and Johnston (1994) note that, although not unilaterally adopted by all researchers and clinicians, DSM criteria have come to comprise the most commonly accepted clinical system for classifying depressive disorders in North America. It is noted, however, that lack of a coherent set of diagnostic criteria for depression was extremely problematic for some time due to findings of low concordance among diagnoses using DSM-based vs. non-DSM-based criteria (Angold & Costello, 1995). Poznanski and Mokros (1994) note that the most recent revision of the DSM, the DSM-IV (American Psychiatric Association, 1994), “has become the standard for clinical practice” (p. 21). The DSM-IV, remaining consistent with its DSM-III and III-R predecessors, divides depressive disorders into two distinct diagnoses: (a) Major Depressive Disorder and (b) Dysthymic Disorder. These two classifications are distinguished on the basis of acuteness and chronicity of symptoms, although both require the presence of “clinically significant distress or impairment in social, occupational, or other important areas of functioning” (APA, 1994). Additionally, the DSM-IV includes a third diagnostic category, Depressive Disorder Not Otherwise Specified, which encompasses depressive features unable to fulfill criteria specified by the other two diagnoses.
Current Epidemiology of Depression

The presence of depressive disorders among children and adolescents, based on DSM-IV criteria, has been amply demonstrated. Current research estimates cite that as many as 5% of children and between 10% and 20% of adolescents have experienced symptoms of sufficient severity to substantiate a diagnosis of a depressive disorder (Reynolds, as cited in Reynolds and Johnston, 1994; Shatté, Reivich, Gillham, & Seligman, found in Snyder, 1999). Some estimates relating to children’s epidemiology are more conservative, ranging from 1.8% to 6% among samples of school-age children (Steele, Armistead & Forehand, 2000). Poznanski and Mokros (1994) cite several research findings supporting the prevalence of depression among both children and adolescents, respectively, and within both clinical and non-clinical populations.

Pertaining to non-clinical populations of preschool-aged children, prevalence rates of unspecified depression range from 0.4% to 4%, depending on the type of assessment methodology utilized (Kashani & Carlson; Kashani, Holcomb, & Orvaschel; Kashani and Ray; Kashani, Ray & Carlson, cited in Poznanski & Mokros, 1994). Poznanski & Mokros concluded that the highest prevalence of depression among non-clinical children resulted from utilization of multiple assessment tools, including DSM-IV criteria. These ratings increased slightly among elementary school-aged children, roughly approximating 6% of subjects sampled (Lefkowitz & Tesiny; Strauss, Forehand, Frame, & Smith, cited in Poznanski & Mokros, 1994). Prevalence of depression in non-clinical adolescents ranged from 3.3% to 6.8% using combined parent interview formats and DSM-III criteria (Deykin, Levy & Wells; Kashani et al., cited in Poznanski & Mokros, 1994).

Estimates among in-patient populations were found to be considerably higher. Silver (as cited in Reynolds & Johnston, 1994) notes that nearly 1 in 6 juveniles admitted to psychiatric hospitals in the United States received intake diagnoses of a depressive disorder. Across a number of research studies conducted with prepubertal clinical populations, prevalence rates ranged widely from 13% to 59%, based on DSM-IV criteria (Alessi & Magen; Carlson & Cantwell; Kashani, Cantwell, Shekim, & Reid; Ney, Colbert, Newman & Young; Petti, cited in Poznanski & Mokros, 1994), while additional research estimates involving older adolescents of unspecified ages ranged from 18.7% to
27% (Robbins, Alessi, Cook, Pooznanski, & Yanchysyn; Strober, Green, & Carlson, cited in Poznanski & Mokros, 1994). Although a wide variety of epidemiological surveys indicate that adult depression is twice as prevalent in females as it is in males (Harvard Medical School, 2002; Maier, Gansicke, Gater, Rezaki, Tiemens & Urzua, 1999), this finding is only partially true for younger individuals. Overall, before puberty (i.e., below 12 years of age) depressive disorders appear to be more common among boys than girls (Rutter, 1986); however, prevalence rates appear to increase for females after the onset of puberty (i.e., above 12 years of age) (Angold & Costello, 1995).

Overall, the wide range of epidemiological estimates proffered by different studies within recent years is highly problematic, and clearly demonstrates a pervasive sense of ambiguity surrounding the construct of depression. Considerable improvements in consistent taxonomy have been made, culminating in the DSM-IV (APA, 1993); however, numerous problems continue to limit a satisfactory definition of depression and, hence, the ability to assess it. As evidenced by the aforementioned discussion of the history of depression, the necessity of establishing reliable diagnostic criteria for this disorder, from which to base sound assessment methodology, remains a crucial goal of future research towards understanding child and adolescent depression. As will be discussed, shortcomings in both conceptualization and assessment are found to contribute to the problematically wide range of estimated epidemiology of these depressed individuals.

Current Problems in the Conceptualization of Depression

Despite advances in nosological specificity, numerous problems in current conceptualization of depressive disorders persist. Trad (1987) defines a disorder, or syndrome, as “a cluster of symptoms causing varying degrees of incapacity, having a commonly occurring clinical picture, a natural history, a number of bio-behavioral correlates, and a predictable response to treatment” (p. 26). Clinical acceptance of this definition, as Trad notes, is important not only to accurately classify a depressive problem and its severity, but also to ensure the efficacy of empirically-based treatment applications. Assessment instruments must be capable of identifying depressive disorders by their symptoms, which must be stable over time, and characteristic of the
construct in question. In other words, assessment instruments must demonstrate adequate construct validity, which requires some consensus regarding acceptable diagnostic criteria.

**Categorical vs. Dimensional Methodology**

Two types of methodologies are commonly employed to assess depression: (a) categorical, and (b) dimensional (Sherak, Speier & Cantwell, 1994). Categorical classification systems, which are formed on the basis of observed symptoms, (e.g., the DSM-IV), are ideally founded on accepted, empirically-based theories with which to predict and explain the etiology, course, and manifestation of consistent patterns of symptoms. Due to their consistency among different individuals over time, these symptoms are categorically associated with specific disorders. Although necessary for diagnostic purposes, categorical systems lack the reliability of dimensional classification systems, which include specific assessment measures. For example, as mentioned, lack of empirical support for current DSM-IV (APA, 1994) diagnostic criteria remains the most prominent impediment to utilizing categorical criteria for assessment purposes (Weiner, 1989).

Dimensional measures, in contrast, are designed to verify the presence and severity of categorically-based diagnoses. This is usually accomplished by quantitatively comparing subjects’ responses to empirically-validated, normative sample-based “cutoff points.” Unfortunately, the parameters used by dimensional measures to confirm the presence of depression as a discrete disorder also often suffer from lacking empirical validation (Angold & Costello, 1995; Cytryn & McKnew, 1974; Hammon & Rudolph, 1994; Sherak et al., 1994; Weiner, 1989). This issue resides at the core of reliability, and continues to require considerable improvement. In other words, dimensional measures lack unanimous categorical agreement of which symptoms must be present, and to what degree, to verify the presence of a depressive disorder. In general, consensus regarding acceptable diagnostic criteria for depression has been linked with controversy regarding selection of assessment measures (Sherak, et al., 1994), which remains a key topic addressed by this manuscript; however, numerous other issues also serve to confound this problem.
The Problem of Comorbidity

Estimates of the prevalence of comorbid externalizing and internalizing disorders with children and adolescents range somewhat, considering the myriad of research studies conducted within recent years. Current longitudinal studies estimate the prevalence of a single comorbid disorder to occur in between 40-70% of depressed children and adolescents, with 20-50% of the population exhibiting two or more comorbid diagnoses (Harrington, Rutter, & Fombonne; Kovacs, as cited in Cicchetti & Toth, 1998; also, see Hammen & Rudolph, 1994 for a summary of specific studies). Externalizing disorders most commonly found to be comorbid among non-clinical samples of subjects include conduct disorder (17-79%), oppositional defiant disorder (0-50%), attention-deficit hyperactivity disorder (0-57%), and substance abuse disorders (23-25%) (Fleming & Offord, 1990; Hammen & Rudolph, 1994). Higher estimates of externalizing comorbid disorders have been found to accompany clinical, as opposed to non-clinical subjects (Fleming & Offord, 1990). A wide array of internalizing anxiety disorders, including separation anxiety, generalized anxiety disorder, phobias, and obsessive-compulsive disorder are also found to commonly accompany depression in children and adolescents (Cicchetti & Toth, 1998; Hammon & Rudolph, 1996). A review of recent studies revealed rates of anxiety disorders, comorbid with unspecified diagnoses of depression, ranging between 30% to 75% among their evaluated populations of children and adolescents (see Hammen & Rudolph, 1994 for a detailed review of these studies).

Depression’s lack of diagnostic autonomy in children and adolescents represents a particularly daunting challenge to its assessment, and stems largely from its high comorbidity with other disorders (Cicchetti & Toth, 1998; Compas, Ay & Grant, 1993; Fleming & Offord, 1990; Hammen & Compas, 1994; Hammen & Rudolph, 1996). Empirical validation of depression as a disorder requires symptom consistency. Thus, symptoms that are commonly comorbid with depressive disorders must either be conceptually distinguished from depression or, conversely, accounted for as age-related developmental manifestations. Of note, this assertion does not necessarily refute the current presence of “shared” symptoms among DSM-IV (APA, 1994) disorders;
however, comorbidity does represent a conceptual challenge to the current state of DSM diagnostic taxonomy.

Theoreticians have addressed the issue of depression’s comorbidity in three ways. For one, given such a high prevalence of comorbid disorders, Angold and Costello (1995) question the existence of a “pure” depression entirely. Perhaps, given the many causal factors implicated in the development of depressive disorders, it is possible that more than one discrete disorder typically emerges, rather than simply depression in isolation. Secondly, and conversely, Hammen and Rudolph (1996) hypothesize that the high consistency among comorbid findings may more accurately reflect limitations in current DSM-IV definitions of childhood and adolescent depression. In other words, rather than conceptually treating externalizing and anxious symptoms as comorbid factors, they may instead represent typical manifestations of depression, particularly in children and adolescents (Patterson, Greising, Hyland & Burger, 1997; Renouf & Harter, 1990).

Hammen and Rudolph’s (1996) position is bolstered by recent findings of high reciprocal comorbidity among Major Depressive Disorder and Dysthymic Disorder, raising issues regarding the diagnostic clarity, autonomy, and utility of these classifications (Goodman, Schwab-Stone, Lahey, Shaffer & Jensen, 2000). Additionally, Clark, Steer, and Beck (1994) provide some support for this position by postulating a theoretical construct present in both depression and anxiety. These researchers have empirically demonstrated the presence of this underlying construct, which they refer to as negative affectivity (NA) (Clark & Watson, cited in Clark, Steer, & Beck, 1994), among a sample of depressed individuals. Angold and Costello (1995) note that comorbidity is problematically persistent across the life span, irrespective of childhood and adolescent onset depression. Given this finding, it may be reasonable to assume that current criteria requires further evaluation and revision for all age groups.

Thirdly, and finally, it remains possible that instruments and techniques commonly used to assess depressive symptoms simply lack sufficient divergent validity to differentiate depression from other forms of psychopathology. While this is believed to be an unlikely possibility, a number of limitations in contemporary assessment methodology have been researched.
Developmental Considerations

Symptom variability across the life span remains another unique and challenging problem in assessing depression, particularly with child and adolescent populations (Hammen & Rudolph, 1996). That is to say, compared to adults, children and adolescents seem to experience and express their depression more erratically and distinctly (APA, 1994; Cicchetti & Toth, 1998; Weiss, Weisz, Politano, Carey, Nelson, & Finch, 1991; Worchel, Hughes, Hall, Stanton, Stanton, & Little, 1990). Unfortunately, high symptom inconsistency has been found over numerous empirical studies involving depressed children and adolescents (please refer to Weiss et al., 1991 for a more thorough review); however, epidemiological studies often provide useful, albeit approximate, estimates of depressive symptoms over childhood and adolescence.

Miller, Bakalar, and Grinspoon (Eds., 2002) note that common depressive symptoms occurring during early childhood (i.e., up to age three) may include feeding difficulty, frequent tantrums, and blunted ability to express emotions. Increased physical complaints are often noted among depressed children between the ages of five and eight, as well as aggression and social withdrawal and/or excessive clinginess to parental figures. Studies have consistently found that anhedonia, hypersomnia, psychomotor retardation, and diurnal affective fluctuations increase after the onset of adolescence, while somatic-related complaints and reports of low self-esteem tend to decrease during adolescence (Carlson & Kashani; Mitchell, McCauley, Burke, & Moss; Ryan, Puig-Antich, Ambrosini, Rabinovich, Robinson, Nelson, Iyengar, & Twomey, cited in Hammen & Rudolph, 1996). Also, adolescents are more likely to exhibit depressive symptoms via externalizing behavior (e.g., delinquency, hostility, substance abuse, etc.).

From a cognitive perspective, manifestation of depressive symptoms also appears to vary in terms of development. Excessive self-reproach has been particularly noted among young children (i.e., 5-8 years old), as opposed to adolescents (Miller et al., Eds., 2002). Some researchers have also posited that depression in adolescents is differentiated from that of children by feelings/verbalizations of hopelessness, congruent with cognitive developmental theory (Kovacs & Paulauskas, as cited in Weiss et al., 1991). Unique manifestations of depressive symptoms also vary by gender. For example, research indicates that females demonstrate a greater tendency to endorse internalizing symptoms,
while males are more prone towards expressing externalizing ones (Edelbrock, as cited in Worchel, Hughes, Hall, Stanton, Stanton & Little, 1990; Hammen & Rudolph, 1996; Worchel et al., 1990).

Although these studies suggest that developmental factors greatly influence depression across the lifespan, children, adolescents and adults continue to be diagnosed using identical DSM-IV criteria (Hammen & Rudolph, 1996; Joiner, Catanzaro & Laurent, 1996; Reynolds & Johnston, 1994). This indiscrimination does not account for the conceptual and theoretical changes in clinical child and adolescent psychology within the past thirty years, and lacks support from empirically- and theoretically-validated research (F. Prevatt, personal communication, March, 2000). Despite imperfections, the DSM-IV acknowledges the developmental uniqueness of depressive symptom expression in children by allowing for the substitution of irritable, rather than depressed mood (APA, 1994; Sherak, et al., 1994). Unfortunately, however, this provision fails to differentiate adolescent and adult symptoms of depression.

Research in development- and gender-based symptom manifestation lends considerable support for an increasingly-specific diagnostic taxonomy, as well as greater consideration of related etiological variables. While this is not the focus of discussion at present, a growing number of theorists conceptualize depressive symptoms not as static, state-dependent characteristics of depressive disorders, but as more dynamic behavioral repertoires that interact with children’s natural maturation process (Cicchetti, Rogosch, & Toth, 1994; Cicchetti & Toth, 1998; Cole, 1991; Cole & Turner, 1993; Leichtman, 1996a, 1996b; Rudolph, Hammen, Burge, Lindberg, Herzberg, & Daley, 2000; Shafii & Shafii, 1992; Thapar & McGuffin, 1996). This perspective is also increasingly sensitive to the cognitive skills required of a subject by a particular assessment instrument (Leichtman, 1996b; Weiss et al., 1991). In other words, the ability of a child or adolescent to respond appropriately to the demands of an assessment medium (e.g., self-report, structured interview format, projective measures, or cognitive/perceptual types of tasks) depends on level of achieved cognitive development (Cicchetti, Rogosh, & Toth, 1994; Cicchetti & Toth, 1998; Kazdin, 1987; Leichtman, 1996b), which will be discussed in a later section.
Problems in the Assessment of Depression

Categorical classification systems for child and adolescent depressive disorders have been based primarily on clinical observation and interview-based reporting of symptoms; lacking in empirical validity, these systems are theoretically problematic. Culminating in the DSM-IV (APA, 1994), the most commonly accepted psychological classification system, such attempts continue to strive to categorically differentiate depression from other disorders; however, as dimensional assessments rely on categorical classification for validity, they remain constrained by the theoretical problems of the past. Also, as behaviors consistently found to be characteristic of depression have come to comprise DSM-IV diagnostic criterion, this criterion, itself, has often been utilized as a form of assessment, which remains theoretically problematic (Kazdin, 1987; Reynolds, 1990a; Weiner, 1989).

Despite the criteria established by the DSM-IV, many researchers and clinicians concur that the term “depression” suffers from a pervasive lack of definition (Kazdin, 1987; Nurcombe, 1994; Weiner, 1989; Worchel, Hughes, Hall, Stanton, Stanton, & Little, 1990). Thus, given the theoretical insufficiency of current depressive taxonomy, assessment of depression has been undertaken both for research purposes and for clinical/diagnostic outcomes. As a result, within both research and practice, depression is capable of being represented as any and/or all of the following: one or more isolated symptoms; a set or series of encapsulated symptoms; specified or generalized characteristics of unspecified consistency/stability; broad, non-clinical syndromes; and clinical, taxonomy-based diagnoses, among other variations. Given this confusion, it is reasonable to question whether separate instruments measure different constructs, as opposed to the same construct at differing levels of severity and specificity (Hammen and Rudolph, 1996). With this constituting the focal point, a discussion will now clarify some of the more prominent foci of research addressing this, and other issues germane to depression.

Internalizing vs. Externalizing Disorders

The prominent forms of assessment for depression have targeted two types of symptoms: a.) internalizing, and b.) externalizing. The ‘internalizing’ classification,
while not inclusive of a restrictive subset of disorders, comprises a useful typology of
definitive “inner-directed” (p. 138), and predominantly covert symptoms of depression
(e.g., feelings of anxiousness, sadness, hopelessness, etc.)(Kazdin, 1989). Because it is
particularly challenging to identify and assess for internalizing symptoms, they are often
confirmed by observing other, explicit behaviors (i.e., externalizing symptoms) found to
be associated with them (Kazdin, 1989; Patterson et al., 1997; Reynolds, 1990a, 1990b).
Externalizing symptoms include social skills deficits, social withdrawal/isolation,
tearfulness/crying, fatigue, irritability, excessive anger, erratic sleep and eating patterns,
psychomotor agitation/retardation, concentration problems, and somatic complaints

Due depression’s persistent lack of clarity, an enduring debate continues to
address the issue of whether many externalizing symptoms constitute inherent
characteristics of depressive disorders, as opposed to comorbid conditions or correlates of
clusters of depressive symptoms (Angold & Costello, 1995; Patterson, Greising, Hyland
& Burger, 1997; Renouf & Harter, 1990). As a result, despite challenges to their
accessibility, internalizing symptoms presently command particular attention due to their
accepted diagnostic utility (Kazdin, 1987, 1989; Kendall et al., 1989). As these
symptoms are subjective phenomena, self-report measures comprise the most widely
Reynolds and Johnston, 1994). Despite some evidence of social desirability among these
instruments (Joiner, Schmidt, & Schmidt, 1996), their acceptance and use remain
common among psychological professionals (Reynolds, 1994; Reynolds & Graves, 1989;
Reynolds & Mazza, 1998; Sherak et al., 1994). For one, their credibility is bolstered by
findings indicating that child and adolescent self-reports of depression tend to remain
highly reliable over short periods of time (i.e., approximately four weeks), excepting
extraneous influences (e.g., treatment, situational variables, etc.)(Reynolds & Mazza,
1998). Overall, extensive empirical research has validated the utility of self-report
measures with children and adolescents (see Kazdin, 1987 for a synopsis); however,
numerous methodological implications merit further discussion.
Categorical vs. Dimensional Methodologies

Assessment instruments are typically utilized for two types of diagnostic purposes: (a) Categorical: to verify the presence of a discrete depression-related diagnosis, based on specified criteria (i.e., DSM-IV), and (b) Dimensional: to determine the extent or severity of observed depressive symptoms on a pre-specified continuum (Kazdin, 1987; Reynolds, 1994). Particular selection of one of these two types of assessment practices, however, hinges implicitly on the purpose of the evaluator (Kazdin, 1987). Within exclusively clinical settings (e.g., inpatient facilities), for example, categorical methodologies represent the most commonly-selected forms of intake assessment. In this capacity, diagnoses of depression typically stem from informal interviews conducted by clinicians independently with both children/adolescents and with their parents, if possible. The intake interview, structured to evaluate the nature and extent of the child or adolescent’s symptoms, consists of comparing reported information with predetermined criteria (e.g., DSM-IV).

Interview-based assessment of depression is flawed, however, due to its lack of consistency, symptom specificity, and empirical support. Clinicians do not rely on a uniform interview methodology, and interview formats are likely to assess broad, rather than specific, symptoms of depression. Additionally, considerable subjective judgment is often required. For example, due to depression’s internalizing nature, clinicians are often forced to reconcile divergent reports of subjectively perceived symptoms. Research has shown that symptoms described by children and adolescents, as opposed to their parents’ descriptions, tend to vary greatly in presence and severity (Kendall, Cantwell & Kazdin, 1989; Reynolds, 1984). Specifically, children and adolescents tend to report lesser symptom severity without necessarily attempting to overtly appear in a non-symptomatic light (Kendall et al., 1989).

Dimensional measures are commonly employed to increase the validity of categorical criteria by providing an empirical normative basis against which to gauge symptom manifestation. Dimensional assessments typically comprise paper-and-pencil inventories, completed by children and adolescents and their parents, requiring participants to rate the severity and extent of their depression-related symptoms.
(Reynolds, 1984, 1990a, 1994; Reynolds & Graves, 1989). These inventories use standardized guidelines delineating appropriate “cutoff” scores at which to deem the presence of depression a statistically significant probability, respective of normative populations. As a result of this increased empirical support, dimensional assessment constitutes the most common mode of diagnosing depression in children and adolescents for research purposes (Kazdin 1987, 1989).

A common misuse of dimensional instruments, however, involves reliance on their results as exclusive categorical criteria for depression. Ideally, it is advocated that data derived from dimensional forms of assessment be compared post-hoc with specified categorical criteria (i.e., DSM-IV) to verify the existence of a depressive disorder, in the taxonomic sense; however, an inherent theoretical problem with this approach is noted. Dimensional and categorical assessments utilize substantially different criteria to ascertain depression. That is to say, the nature and severity of depressive symptoms sufficient to suggest the presence of depression on independent report measures (i.e., child, parent, etc.) differs considerably from that required by DSM-IV criteria, resulting in questionable diagnostic precision.

Diagnostic criteria among different dimensional assessments are also often found to conflict. In a study conducted by Kazdin (1989) involving both child and parent reports of depressive symptoms, mutually exclusive criteria were found to exist. In other words, for each assessment instrument, quantitatively different cutoff points were required to reach clinical significance. Additionally, each instrument identified discrepant proportions of clinically referred children (51.6% using child-report criteria vs. 28.1% using parent-report criteria) as being clinically depressed. As a result, while dimensional instruments hold the potential to provide invaluable diagnostic information, depression as a definitive diagnostic entity retains its validity based on DSM-IV criteria. Thus, it is concluded that obtainment of different informants’ ratings of depression constitutes necessary, yet not sufficient criteria.

Integration of Multiple Assessment Modalities and Respondents

Because of the many complexities and limitations of theory and methodology, no single instrument, administered independently, has been found capable of validly
indicating the presence of depression in children (Kazdin, 1987; Worcel et al., 1990). Reynolds and Johnston’s (1994) findings corroborate this conclusion, citing that a higher percentage of individuals receive diagnoses of depression as a result of being administered a full psychometric battery using numerous instruments (i.e., assessment of broad psychopathology assessed using several domains and multiple informants) than from one, sole measure exclusively assessing depression. As a result, these types of assessments are increasingly advocated, including assessments administered to the peer group members and teachers of children and adolescents (Crowley & Worcel, 1993; Crowley, Worcel & Ash, 1992; Kazdin, 1987; Reynolds, 1994; Worcel, Hughes, Hall, Stanton, Stanton, & Little, 1990).

Despite standardized formats, integrating data obtained from multiple informants remains challenging. Data collected among the various respondents tend to conflict, likely because multiple-rater assessment instruments tend to demonstrate a limited propensity to demonstrate a substantial relationship (Asarnow & Carlson; Doerfler, Felner, Rowlinson, Raley & Evans, in Kazdin, 1987). While relatively strong correlations tend to be consistently observed among different self-report measures, this is not the case between results derived from self- and other- (i.e., peers’ and teachers’) reports. A study conducted by Worcel et al. (1990) using three common measures used to gauge both internalizing and overall symptoms of depression (e.g., self, peer, and teacher formats) provided evidence of this pervasive and egregious problem. Correlations of 0.27 were found between self- and peer-report formats, of 0.13 between self- and teacher-reports for overall depression, and of 0.10 were found between self- and teacher-reports of internalizing symptoms. Additionally, an average correlation of 0.22 was derived from a meta-analytic study evaluating the relationship between parent- and self-report formats assessing depression (Achenbach, McConaughy, & Howell, 1987). Thus, a popular consensus is that third-party raters (i.e., peers, parents, and teachers) fail to accurately appreciate internalizing symptoms, presumably because their presence and extent must be inferred on the basis of external behavior (Crowley et al., 1992; Kazdin, 1987, 1989).

Modest correlation between divergent scales may also be attributable to a number of additional factors. For one, mothers’ ratings of their children’s behavior and
adjustment have been found to correlate with level of maternal psychopathology (Forehand, Lautenschlager, Faust & Graziano; Griest, Wells & Forehand; Mash & Johnston; Moretti, Fine, Haley & Marriage, as cited in Kazdin, 1987). This phenomenon may also extend to the psychopathology of peer groups. In one study, a statistically significant correlation of 0.26 was found between peer- and self-reported ratings of depressive symptoms (Crowley and Worchel, 1993). While this relationship is modest, the considerable implications of this study suggest a tendency of individuals exhibiting symptoms of depression, as opposed to non-depressed peers, to rate others as being more depressed. Overall, this research cautions against the use of unitary, single-informant sources of data in the diagnosis of depression, and also indicates a clear need to consider the emotional state of those who evaluate depressive symptoms in children and adolescents.

Psychometric test variation (a.k.a., measurement error) among divergent assessment instruments has also been implicated as a likely cause of lack of correlation (Cole, Hoffman, Tram, & Maxwell, 2000; Crowley & Worchel, 1993; Hammon & Rudolph, 1996; Joiner, et al., 1996; Worchel et al., 1990). One prominent hypothesis posits that assessments completed by different individuals utilize different formats (i.e., different test items, item organization, and item presentation), which results in low inter-rater correlation. One study controlled for this potentially attenuating influence by expanded the use of one common peer-rating instrument to include both self-reports and peer-reports, thus utilizing a consistent pool of items presented in a uniform manner (Crowley and Worchel, 1993). Results of this research revealed a significant correlation between self- and peer-reports, based on the one expanded scale format (r = 0.44), which was found to be significantly higher than the relationship between two, separately-completed report scales (r = 0.22). While these results are not found to be completely attributable to specific instrument test error variance, they implicate that such instruments’ psychometric structure is partially responsible.

Research has also explored the structural differences between other-rater scales in order to account for their lack of correlation. One such study simultaneously investigated the factor structure and item similarity of three commonly used inventories in order to verify their construct validity (Crowley, Worchel, & Ash, 1992). Although this study
was able to demonstrate differences across the constructs (i.e., factors) of these three instruments, it also revealed overall consistency among the three instruments at the item level. Despite this finding, a considerably greater relationship was found among same-test items than among combinations of similar items involving multiple tests. Thus, rather than attributing such lack of consistency to insufficient construct validity, the research results appeared to corroborate the aforementioned findings attributing informants’ emotional state to lack of inter-scale correlation.

Cole, Hoffman, Tram and Maxwell (2000) also undertook a factor analysis to determine the consistency of the underlying factor structures of parent and child forms of a popular measure of depression. In contrast to Crowley et al.’s findings (1992), these authors concluded that parent and child reports of the same instrument comprise different underlying factors that, as a result, typically assess different aspects of psychopathology in children. It is noted, however, that the research of Cole et al. (2000) may have failed to address the construct of depression as a categorical entity, rather than a dimensional phenomenon, which is theoretically problematic.

Overall, Kazdin (1987) notes that the overarching problem inherent in the use of multiple-informant responses consists of inherent rater bias. This could account for a number of factors, particularly the susceptibility of assessment instruments to response distortion (e.g., responding in a favorable manner) (Joiner, Schmidt, & Schmidt, 1996). Despite the numerous problems inherent in integrating report data, however, the practice has proven to be worthwhile in several ways (Kazdin, 1987, 1989). For one, dimensional data obtained from more than one source bolster the validity of clinical, interview-based diagnosis. Further, self- and other-rater (i.e., self, parent, peer, teacher) assessment data has demonstrated strong concurrent validity with specific internalizing depressive symptoms. Such internalizing symptoms lend themselves towards a cognitive conceptualization of depression, and include the following broad, non-exhaustive list: low self-esteem, negative attributional style, negative cognitions, and suicidal intent/ideation. In addition, factors such as decreased social involvement and maladaptive parenting style have been associated with both parent- and self-reports.

Cognitive Theoretical Perspectives on Child and Adolescent Depression
Cognitive theory has occupied a prominent place within the body of research on depressive disorders in recent years, and has held considerable influence on the manner in which depression is conceptualized and assessed (Beck, 1967; Dowd, Ed., 1993; Hammen & Goodman-Brown, 1990; Kaslow, Brown, & Mee, 1994; Kazdin, 1987, 1989; Reynolds, 1984, 1990; Schrodt, 1992). While a dominant intent of these models is to delineate etiological factors, this focus falls largely outside of the scope of this discussion. Similarly, an exclusively cognitive approach does not seek to refute the ample body of evidence supporting genetic/biological influences in depressive disorders (for further review, see Bemporad, 1992; Carlson, 1998; Dahl, Ryan, Puig-Antich, Nguyen, Al-Shabbout, Meyer, & Perel, 1991; Emile, Walkup, Pliszka, and Ernst, 1999; Geller, Reising, Leonard, Riddle & Walsh, 1999; Kazdin, 1987; Susman et al., Schmeelk, Worrall, Granger, Ponirakis and Chrousos, 1999). Cognitive conceptualizations of depression are generally inclusive of biological and affective characterizations (Dowd, Ed., 1993), and are considerably more amenable to a discussion of assessment applications.

**Beck’s Theory of Depression**

In particular, the pioneering research of Beck (1967, 1987) has provided a sound basis from which to enhance current diagnostic utility with children and adolescents by exploring cognitive symptoms of depression. Beck’s theory, representing an information-processing and diathesis-stress model of cognitive functioning, essentially views depression as related to and sustained by negatively biased cognitive processing (Beck, 1967; Beck & Weishaar, 1989; Dowd, Ed., 1993). Beck (as cited in Dowd, 1993) posits that, independent of many causal factors (i.e., biological, environmental, interpersonal), depression in both children and adults is mediated by and sustained by cognitive processes or, more specifically, one’s cognitive schemas.

Cognitive schemas are conceptualized as unified, related collections of consistently held cognitions/beliefs (Beck & Weishaar, 1993) that are formed as a result of one’s life experiences (e.g., interactions with others and environmental circumstances, as well as feedback from others and the environment). Beck’s (1967, 1987; Beck & Weishaar, 1993) theory asserts that targeted, negative thoughts lead to expression of
affective and physiological symptoms. More specifically, he proposes three types of distorted beliefs consistently identified in depressed individuals (regardless of age): (a) selective abstractions, (b) overgeneralizations, and (c) negative self-attributions (Beck, 1987). Selective abstraction occurs when an individual attends to the negative aspects of a situation/event to the exclusion of its positive aspects. Overgeneralization, in turn, represents one’s tendency to apply negative conclusions/cognitions to events/situations other than those in which they occurred. Finally, negative self-attribution, a form of selective abstraction, involves the individual’s application of negative cognitions to personal qualities and attributes.

Beck asserts that the negative schemas of clinically depressed individuals manifest in persistent cognitive biases that culminate in the formation of what is termed the “cognitive triad” (Beck, 1967, 1987; Dowd, 1993). This triad consists of negatively skewed attributions pertaining to oneself, the world, and the future. Although such negative views may not be consistent with others’ views of the depressed individual, the negative schemas serve as subjective filters for the depressed individual through which he or she negatively interprets events and perceptions (Beck, 1987; Dowd, 1993; Hammen & Goodman-Brown, 1990). Beck’s theory posits a cognitive diathesis-stress model to account for depression as a syndrome. In other words, negative beliefs associated with Beck’s cognitive triad are likely to predispose individuals towards clinically significant levels of depression upon exposure to stressful stimuli (Beck, 1987; Dowd, 1993; Hammon & Goodman-Brown, 1990; Hilsman & Garber, 1995).

Reynolds (1990b) notes that applications of Beck’s theory hold considerable assessment utility, as children and adolescents’ negative cognitions are more readily identifiable than many other internalizing symptoms. A number of promising studies have been performed that have targeted the elicitation of explicit cognitive symptoms from depressed children and adolescents (Epkins, 2000; Hammen & Compas, 1994; Hammen & Goodman-Brown, 1990; Kaslow, Stark, Printz, Livingston, & Tsai, 1992; Meyer, Dyck, and Petrinack, 1989; Stark, Schmidt, & Joiner, 1996). Research has found that both clinical and non-clinical samples of children diagnosed with depression according to DSM-IV criteria reported increasingly negative views pertaining to all three aspects of Beck’s cognitive triad (Kaslow, Stark, Printz, Livingston, & Tsai, 1992;
Meyer, Dyck, & Petrinack, 1989). Additionally, Epkins (2000) recently utilized Beck’s theoretical model to successfully distinguish between internalizing and externalizing symptoms on the basis of self-reported cognitive disturbance. These results are found to be comparable to those observed by Worchel, Little and Alcala (1990), in which children experiencing symptoms of depression demonstrated pervasively more negative expectations (i.e., negative view of self and world/outcome) relating to their performance on a task of cognitive ability, compared with their non-depressed counterparts, and irrespective of actual performance.

The Learned Helplessness Model

Cognitive theory is especially germane to the assessment of depression in its ability to gauge an individual’s attributional style and perceived locus of control, which can also be evaluated through psychometric means (Kazdin, 1987). The learned helplessness model of depression remains a major theoretical contributor to this position (Abramson, Seligman, & Teasdale, 1978). In essence, this model posits that feelings of helplessness result when individuals perceive personally relevant situations to be uncontrollable (i.e., inescapable, or unable to be successfully overcome). This derivation of Beck’s (1967; 1987) theory proposes that individuals develop a “depressogenic” attributional style in which negatively-perceived events are attributed to internal, global, and stable causes rather than to transitory, specific, and external ones (Abramson, Metalski, & Alloy, 1989; Kaslow et al., 1994; Seligman, Peterson, Kaslow, Tanenbaum, Alloy & Abramson, 1984). Thus, negative outcomes are attributed to subjectively-perceived personal defects/shortcomings over which individuals perceive a lack of control. This negative attributional style consists of feelings of helplessness (Seligman, 1975), and at times hopelessness (Abramson, Metalsky, & Alloy, 1989), both of which have been associated with more overt symptoms of depression.

A significant relationship has been found between negative attributional styles and the presence of categorically defined depressive disorders (Seligman et al., 1984). Additionally, attributional styles assessed in Seligman et al.’s (1984) study were found to consistently predict depressive symptoms six months after the study’s onset. Similar findings have been found, lending considerable support to the inclusion of assessments
that either directly or indirectly evaluate attributional style for diagnostic purposes (Asarnow & Bates; Bodiford, Eisenstadt, Johnson, & Bradlyn; Gladstone & Kaslow; Kaslow, Rehm & Siegel; Seligman, Peterson, Kaslow, Tanenbaum, Alloy & Abramson, as cited in Simpson, 1997; Joiner, 2000).

The Tripartite Model of Depression

As previously mentioned in relation to the problem of comorbidity, Clark and Watson’s (cited in Clark, Steer, & Beck, 1994) Tripartite theory posits the existence of several factors accounting for depression in childhood and adolescent populations. This theory differentiates itself from other cognitive models by proposing the following three constructs: (a) positive affect (PA) (i.e., lack of anhedonia or dysphoria) which is associated with non-depressed individuals; (b) physiological hyperarousal (PH), which is associated primarily with anxiety, and; (c) negative affect (NA), which is associated with both anxiety and depression. Several studies attest to the validity of this model in conceptualizing and differentiating depressive symptoms among both clinical and non-clinical populations of children and adolescents (Clark, Steer, & Beck, 1994; Clark & Watson, as cited in Clark, Steer, & Beck, 1994; Joiner, Beck, Rudd, Steer, Schmidt, & Catanzaro, 1999; Joiner, Catanzaro, & Laurent, 1996; Joiner, & Lonigan, 2000).

The Tripartite model is particularly appealing to the assessment of depression in children and adolescents due to the ability of its three distinct constructs to differentially diagnose depression as a separate disorder from its traditionally comorbid anxiety (Joiner & Lonigan, 2000). Research has implicated a higher prevalence of NA symptoms (e.g., emotional distress such as sadness, anger, fear, and guilt) (Watson & Clark, as cited in Clark, Steer, & Beck, 1994) associated with both depression and anxiety, assessed using self-report measures. Further, several factor analytic studies have found that depressive symptoms are distinguishable from anxious symptoms on the basis of their decreased loadings on the PH factor (Joiner, et al., 1999), as well as the PA factor (Joiner & Lonigan, 2000). In both studies, a relatively consistent degree of association with the NA factor was identified for both depressive and anxious disorders.
Cognitive Information Processing and Resource Allocation Models

Researchers have endeavored for quite some time to understand and conceptualize the complex cognitive mechanisms by which humans processes information (Baddeley, 1990). As a result, numerous information processing models have been posited, key assumptions of which are implicit in cognitive theories that account for depressive symptoms and disorders (Beck, 1967). One such assumption is that memory comprises the fundamental cornerstone of cognitive information processing (Baddeley, 1990). Defined as “records of percepts,” (p. 13), memories represent the product of humans’ ability to cognitively encode, or store, information as it is perceived by one of the five senses. Baddeley emphasizes the distinction between physiological perception (i.e., the reception of stimuli by the sense organs) and cognition, which he describes as primarily subjectively-based “mental activity” (p. 98). While the preponderance of information processing research lauds human visual capability as the dominant sensory modality, Baddeley (1990) notes that other senses (i.e., taste, touch, hearing, and olfactory) are likely encoded as equally salient memories.

Analogous to a digital computer metaphor, encoded memories are mediated by a lengthy array of factors (i.e., attention, the salience of the stimuli, context, etc.) (Baddeley, 1990). The model also conceptualizes key cognitive structures according to the type of information that is stored, as well as its ability to be processed and recalled by the individual (For ease of reference, a diagram of these structures is provided in Appendix A). These consist of short-term memory (STM), working memory (WM), and long-term memory (LTM), the first and last of which can be broken down further into more specific components/manifestations of information. Working memory is responsible for manipulation and temporary storage of information heeded in Short-Term Memory, and has been implicated in a variety of cognitive tasks instrumental to daily functioning and learning (Atkinson & Shiffrin, 1968, as cited in Baddeley, 1990).

While a number of studies have pragmatically applied cognitive information processing models directly to learning-based (i.e., academic performance-based) types of cognitive settings (see Schneider, 1996, for further information), fewer studies have evaluated the specific effects of early childhood neurological and physiological
development on cognitive information processing capabilities (Blatt, 1990; Exner, 1996; Leichtman, 1996a, 1996b). Relating to the construct of depression in children and adolescents, the Resource Allocation model of cognitive information processing proposed by Ellis and Ashbrook (1989) serves as a useful complement to contemporary theories of depression in understanding the relationship between emotional states and cognitive processes.

The way individuals process information may hold considerable potential as a marker for depression in children and adolescents. In fact, research has investigated the association between information processing capabilities and depression in adults as early as 1989 (Ingram & Reed; Ingram & Hollon, cited in Ellis & Ashbrook, 1989). Different from Beck’s (1967) theory, however, Ellis & Ashbrook’s (1989) resource allocation model more clearly delineates the direct mechanisms by which human cognition, as opposed to individuals’ internal verbalizations of this cognition, is mediated by emotional functioning.

The resource allocation model hypothesizes that any given individual is able to apply only a finite cache of cognitive-attentional resources to particular cognitive tasks that, according to their nature, demand varying amounts of these resources. Individuals exhibit varying ability to allocate their cognitive-attentional resources, and this ability can be reduced by the overwhelming preponderance of negative thoughts that accompany depression-related emotional instability. These depressive cognitions have been found to interfere with productive cognitive capabilities by limiting individuals’ ability both to process and encode information (Baddeley 1990). Serving as distracting influences (conceptualized as thoughts irrelevant to the task at hand), negative cognitions interfere in the depressed individual’s ability to process information in the most efficient manner. Thus, the ability to sustain the cognitive attention required to optimally complete a particular task is reduced (Ellis, Moore, Varner, Ottaway, & Becker, 1997).

Research has demonstrated that a number of factors relating to the nature of presented tasks also tend to mediate the influence of depression on cognitive performance, specifically the recollection and retrieval of information (Ellis et al., 1989). Increasingly ambiguous and unorganized visual stimuli tended to result in greater difficulty processing the information meaningfully among depressed, as opposed to non-
depressed individuals, presumably due to the cognitive resource limitations resulting from depression. (Ellis et al., 1989; Ellis et al., 1997). Depressed individuals in this study were identified on the basis of their verbally expressed depressive schema. Similarly, Ellis, Thomas, & Rodriguez (as cited in Ellis et al., 1989) found that the performance of depressed subjects was detrimentally affected to a significantly greater degree on increasingly difficult, as opposed to more simple cognitive tasks.

The resource allocation model also provides a theoretical basis attempting to account for the categorical distinction between depression and anxiety, which also exerts a detrimental influence on individuals’ cognitive-attentional resources. Specifically, processing capabilities are impeded by individuals’ negative views concerning the future in instances of anxiety, which is contrasted by negative views concerning the past or present in the case of depression (Baddeley, 1990). Roughly similar to the Tripartite model’s PH factor, the resource allocation model also emphasizes that physiological symptoms result from anxiety, but not depression.

**Commonly-Used Self-Report-Format Instruments**

Research integrating cognitive theory with the assessment of depression in children and adolescents has utilized a variety of instruments that have integrated specific cognitive manifestations of depressive symptoms into self-report questionnaires. Examples of these instruments include the Cognitive Triad Inventory for Children (Kaslow, Stark, Printz, Livingston, & Tsai, cited in Stark, Schmidt, & Joiner, Jr., 1996), the Positive and Negative Affect Scale for Children (Laurent et al., cited in Joiner & Lonigan, 2000), the Children’s Attributional Style Questionnaire (Seligman et al., 1984) and the Hopelessness Depression Symptom Questionnaire (Metalsky & Joiner, 1997); however, the presence of symptoms broadly associated with cognitive theory has also been validated using more commonly-used dimensional instruments.

While a large variety of dimensional instruments are frequently utilized to assess the extent and severity of depression in children and adolescents, fewer of these demonstrate superior psychometric properties supporting their use in clinical settings. The Children’s Depression Inventory (Kovacs, 1992), the Reynolds Children’s Depression Scale (RCDS)(Reynolds, 1989a) and the Reynolds Adolescent Depression
Scale (Reynolds, 1987) are three of the most popular of these instruments, and will be discussed in relation to their ability to assess depression in children and adolescents.

**The Children’s Depression Inventory**

The Children’s Depression Inventory (Kovacs, 1992) is the most commonly administered self-report measure assessing depression among children and adolescents between the ages of seven to seventeen years of age (Craighead, Craighead, Smucker, & Ilardi, 1998; Kazdin, 1987, 1989; Kendall, Cantwell, & Kazdin, 1989; Stark, Schmidt, & Joiner, 1996). This paper-and-pencil format inventory consists of twenty-seven subjectively endorsed items (completed independently by the child or adolescent), each of which is comprised of three statements of varying symptom severity. The examiner is instructed to read the items aloud to children between the third and sixth grades in order to compensate for any potential reading comprehension difficulties. In addition to addressing more traditional, overt symptoms of depression, the Children’s Depression Inventory items also consist of statements relating to the academic environment, aggression, and peer interaction. After completion of the Children’s Depression Inventory, items are given score values ranging from 0 and 2, and are summed to comprise a total score ranging from 0 to 54.

On the whole, Kovacs (1980/1981, as cited in Joiner, Katz, & Lew, 1997) has indicated that the Children’s Depression Inventory demonstrates acceptable internal consistency (i.e., over .80), and test-retest reliability coefficients of between .86 and .72 over a one month duration. Additionally, significant concurrent validity has been observed at the .55 level with categorical diagnoses of depression conducted by clinicians. Additionally, other studies have obtained adequate internal consistency alphas ranging from .81 to .91 (Cole, Hoffman, Tram, & Maxwell, 2000; Kendall, Kantwell, & Kazdin, 1989; Joiner, 2000), and .93 (Joiner, Katz, & Lew, 1997). Kovacs (1992) reports that the Children’s Depression Inventory adheres most closely to a five-factor structure that includes the following factors: negative mood, anhedonia, negative self-esteem, ineffectiveness, and interpersonal problems. Correspondingly, alpha coefficients for these scales are reported to range from .59 (interpersonal problems) to .88 (self-esteem). It is noted, however, that some support for a six-factor structure adding a Biological
Dysregulation factor has recently been proposed by Craighead, Craighead, Smucker, & Ilardi (1998) using responses from a community sampling of children and adolescents.

In relation to cognitive theory, the Children’s Depression Inventory has been found to correlate adequately with several cognitive-related constructs (i.e., cognitive attributions, hopelessness, and negative view of self) (Kazdin, French, Unis, Esveldt-Dawson, & Sherick; Kovacs, Saylor, Finch, Baskin, Furey, & Kelly; Saylor, Finch, Spirito, & Bennett; Seligman, et al., as cited in Kendall, Cantwell, & Kazdin, 1989). Additionally, the Children’s Depression Inventory has been positively correlated with constructs from Seligman, et al.’s (1984) attributional theory (Joiner, 2000), Clark and Watson’s (cited in Clark, Steer, and Beck, 1994) Tripartite model (Joiner, & Lonigan, 2000), and Beck’s (1967) cognitive theory (Stark, Schmidt, & Joiner, 1996), demonstrating this instrument’s versatility and adaptability to cognitive modalities of research.

The Reynolds Child Depression Scale (RCDS)

The Reynolds Child Depression Scale (RCDS; Reynolds, 1989a), revised from its predecessor the Child Depression Scale (CDS; Reynolds, 1990a; Reynolds & Graves, 1989), is a 30-item, self-report scale based on an ethnically and socioeconomically diverse normative sample designed for use with children between the third and sixth grades (Reynolds, 1989b). Corresponding with clinical, research-based standards, the RCDS reports to assess the prevalence of “most symptomatology specified by the DSM-III for major depression and dysthymic disorder” (p. 164)(Reynolds, 1990a). Thus, while congruent with current categorical standards, the RCDS does not presume its function to represent an exclusively valid indicator of depression (Carlson, 1992). Given the minimal changes accompanying the DSM’s most recent revisions, this instrument is felt to be adequately consistent with the established criteria for depressive disorders on which it was based (i.e., DSM-III). It is noted, however, that the RCDS does not include items assessing for more overt types of externalizing behavior, which have recently been given increasing research attention in the DSM-IV.

The RCDS gauges the extent, or severity, of children’s depressive symptoms on a four-point likert-type scale (from “almost never” to “all the time”) corresponding with 29
of the 30 items; the remaining item presents the subject with five simple, line-drawn faces with expressions ranging from happy to sad. The first 29 items are worded so as to represent responses indicative of current (i.e., present-time) depressive symptoms (Reynolds, 1994). Clinically-validated cutoff points accompany the RCDS’s technical manual for interpretive purposes (Reynolds, 1989b).

The RCDS standardization sample demonstrates excellent split-half internal consistency reliability at the .90 level (Carlson, 1992; Reynolds, 1990a;), and studies have demonstrated estimates ranging from 0.88 to 0.92 among additional samples of both clinical and non-clinical children (Reynolds, 1994). Reynolds and Graves (1989) obtained an overall test-retest reliability coefficient of 0.85 using a population of non-clinical school children between the third and sixth grades, which ranged from 0.81 to 0.92 among the respective grade levels. The validity of the RCDS has also been empirically verified through concurrent comparisons with other prominent, psychometrically valid depression screening instruments. For example, numerous research studies since the inception of the RCDS have derived correlation coefficients ranging from .70 to .79 when comparing the RCDS to the Children’s Depression Inventory (Reynolds, 1994). Additionally, high convergent validity has been evidenced at the 0.76 level in comparing the RCDS with the Children’s Depression Rating Scale: Revised (CDRS-R), which is frequently used in structured diagnostic interviews.

**The Reynolds Adolescent Depression Scale (RADS)**

The counterpart to the RCDS, the Reynolds Adolescent Depression Scale (RADS; Reynolds, 1986), consists of 30 items, all of which are endorsed according to a four-point likert-type scale (from “almost never” to “most of the time”). As with the RCDS, the RADS clearly differentiates itself from categorical measures of depression, to its benefit (Kaplan, 1992). This inventory, also a self-report scale, was created to evaluate the presence and severity of depressive symptoms endorsed by adolescents between the ages of 12 and 18 (Reynolds, 1987). Similar to the RCDS, Reynolds (1994) notes that the RADS was designed predominantly to measure internalizing types of symptoms such as somatic concerns, lack of motivation, negative cognitions and affect, and also vegetative symptoms typically indicative of endogenous forms of depression (Reynolds, 1994).
Kaplan (1992) asserts that the RADS lacks fundamental theoretical support to bolster its construction; however, compared with the RCDS, the RADS shows developmental sensitivity by including maturational factors associated with adolescent-, as opposed to childhood-onset depression (e.g., restlessness, inattention). The RADS was also based on a geographically, ethnically, and socioeconomically diverse normative sampling of adolescents.

As Reynolds (1994) notes, numerous research studies have demonstrated superior internal reliability coefficients pertaining to RADS, ranging from .91 to .96 among both clinical and non-clinical samples. A test-retest reliability coefficient of .80 is reported, based on a moderately large sample (n = 104) within a six-week period of time, and .79 based on a larger sample (n = 415) over a 12-week period of time. Additionally, numerous studies have compared RADS results with other prominent measures of depression in adolescents, adults, and children. Reynolds (1990a) asserts that coefficients ranging from .70 to .89 have been observed as a result of correlating results of the RADS with those obtained from the Beck Depression Inventory (BDI), the Center for Epidemiological Studies-Depression Scale (CED-D), and the Zung Self-Rating Depression Scale. Additionally, correlations of .75 and .70, respectively, were obtained from retrospective studies of RADS and RCDS scores among samples of non-clinical eighth-grade students. A correlation coefficient of .64 was observed in a similar type of clinical study relating RADS and RCDS scores.

A number of studies have been amply demonstrated the excellent psychometric properties of the RADS, most notably its considerable convergent validity. Reynolds and Mazza (1998) report a 90% concordance rate with diagnoses of depression derived using a DSM-IV-based criterion measure administered via a clinical interview format. This study was conducted with a population of 98 inner-city adolescents. Additionally, concordance rates of 78.1% and 89.2% have been observed among other studies of varying populations, both clinical and non-clinical (Reynolds, 1994). Of note, elevated RADS scores have also been found to discriminate between suicidal and non-suicidal inpatient adolescents in numerous studies cited by Reynolds (1994). No information pertaining to the factor structure of either the RADS or the RCDS is available to support these instruments’ construct validity, which is problematic and warrants further study;
however, both of these instruments demonstrate ample ability to detect the presence of depressive disorders using DSM-IV criteria.

**Alternative assessment: The Rorschach Comprehensive System (CS)**

Exner (1993) asserts that the imprecision of instruments designed to measure depression is a product of the ambiguity of current categorical taxonomy. New methodology for improving diagnostic criteria is warranted and has been advocated for some time (Kazdin, 1987). Various studies have been conducted to develop alternative forms of assessment with children and adolescents within clinic/laboratory contexts (Garber & Kaminski, 2000; Cole, Martin, Peeke, Seroczynski, & Hoffman, 1998). Specifically, researchers have investigated potential markers of depressive disorders by observing intricate, *en vivo* patterns of interactions/behavior between parents and children/adolescents. Also, performance on various cognitive information processing tasks (e.g., exercises involving schema-related self-attributions and expectations, and ability to sustain attention) have been evaluated (Garber & Kaminski, 2000). Unfortunately, findings have indicated that a) the construct validity of such alternative assessment measures is questionable at best, b) these types of assessment involve highly complex methodology, and c) they are highly labor intensive and time-consuming.

A more structured instrument, the Rorschach inkblot test, may show more promise relating to the assessment of depression with children and adolescents. Scored according to the Exner Comprehensive System (CS) (Exner, 1993), the Rorschach represents a valuable and oft-neglected complement to criterion-based assessment (L. Kubiak, personal communication, June 2000). Similar in some ways to the research of Garber and Kaminski (2000) and Cole et al. (1998), the Rorschach conceptualizes different types of shortcomings in individual cognitive-perceptual processing. Specifically, the CS evaluates markers of depressive symptomatology that manifest in specific ways of performing, and correspond with affective functioning, negative cognitions, and cognitive complexity/coping resources (Exner, 1993). These aspects of processing represent the “primary features in the psychological organization of the person” (p. 362).
**Historical and Present Use of the Rorschach**

The evolution of the Rorschach Inkblot Test as an indicator of depression has spanned multiple decades. Throughout this time, the Rorschach has been the target of tremendous controversy among clinicians and researchers, largely due to differences in its theoretical backing and selection of interpretive methods (Weiner, 1997). The Rorschach can be scored and interpreted according to two different types of interpretive approaches: nomothetic and idiographic. Nomothetic interpretations “address ways in which respondents resemble other kinds of people” (Weiner, 2000)(p. 170) and, hence, tend to compare their performance with what is typical of normative populations. Idiographic interpretation, in contrast, does not directly address individual differences using normative data. Instead, the emphasis of interpretation falls on aspects of interest “that are at least somewhat unique to individuals’ particular psychological makeup and current circumstances” (p. 171). Nomothetic interpretations follow empirical guidelines, while idiographic interpretations typically conform to the tenets of particular psychological theories (i.e., usually psychoanalytic or psychodynamic). At present, clinicians using the Rorschach advocate for either exclusively idiographic or nomothetic interpretations, or conjoint interpretation using both approaches. Historically, this disparity in approaches has greatly confounded the issue of interpretation.

Created by Hermann Rorschach in Switzerland in 1921, the Rorschach Inkblot Test (RIT) was originally designed as a nomothetic instrument, and was intended primarily to differentiate psychopathological individuals from “normals” (Weiner, 1986); however, the Rorschach gradually came to be used by its creator as a psychoanalytic, content-oriented (i.e., idiographic) method of projectively analyzing personality (Aronow, Reznikoff & Moreland, 1995). From a historical perspective, this use is generally attributed to the dominant psychoanalytic zeitgeist of Hermann Rorschach’s time; however, it is ironic that a nomothetic approach would have better suited Rorschach’s attempt to assess psychopathology (Blatt, 1990).

Hermann Rorschach continued to administer and interpret his inkblots idiographically until his early death at the age of 37 (Aronow et al., 1995). After that
time, due to the ascendancy of psychoanalytic theory in the early 1900’s, clinicians continued to interpret the Rorschach in almost exclusively idiographic ways; however, acceptance and practice of nomothetic interpretation grew in the United States throughout the mid-1900’s, as clinicians rediscovered Hermann Rorschach’s original 1921 manuscript *Psychodiagnostik* (Aronow et al., 1995). The Rorschach’s utility as an indicator of specific psychopathology began to be appreciated, and its results could finally be interpreted against the template of a set of uniform categorical diagnostic criteria. Although idiographic approaches persisted and often were integrated with nomothetic interpretation (Lerner, 1996b), American clinicians began to show greater interest in cognition and perception in attempts to better predict and comprehend symptoms (Weiner, 1986).

Clinicians presently interpret individuals’ responses to Rorschach inkblot stimuli in both nomothetic and idiographic ways, largely driven by their theoretical beliefs. The Object Relations (Lerner & St. Peter, 1984; Priel, Myodovnik & Rivlin-Beniaminy, 1995; Smith, 1990, among others) and Psychodynamic (Kleiger, 1997; Meloy & Singer, 1991; Murray, 1994; Viglione, Brager, & Haller, 1991, among others) schools continue to embrace idiographic interpretations of the Rorschach, while clinicians advocating its use as a cognitive-perceptual instrument embrace nomothetic interpretations (Exner, 1993).

**Overview of the Exner Comprehensive System (CS)**

The Comprehensive System (CS) is used predominantly to assess broad personality characteristics and their resulting behavioral proclivities (Blatt, 1990; Exner, 1988; Weiner, 1997). This is accomplished by evaluating cognitive processing *via* examination of the meaning individuals assign to their perceptual experiences (Exner, Arbruster, & Viglione, 1978; Weiner, 1997). The CS distinguishes itself from idiographic approaches by empirically comparing responses against age-based normative data (Exner, 1997). In addition to assessing for psychopathology, this systematic, comprehensive process evaluates numerous aspects of personality, and involves multiple sets of complex scoring criteria (Exner, 1993, 1997).

**Administration of the CS**
Examinees are exposed to the ambiguous Rorschach inkblot stimuli and are prompted to elicit their perceptions, based on their viewing of the cards. Each of the ten total inkblot cards are initially shown to the examinee accompanied by the question, “What might this be?” (Exner, Colligan, Hillman, Ritzler, Sciara, and Viglione, 1995)(p. 5). Examinees’ responses are recorded verbatim. To improve the scoring accuracy of the CS, administration involves both an initial presentation of the inkblot forms, and a subsequent “inquiry” phase to clarify aspects of the subject’s responses. During this phase, all ten cards are shown to the examinee a second time after the following instructions are stated:

Now we are going to go back through the cards again. It won’t take very long. I want to see the things that you said you saw and make sure that I see them like you do. We’ll do them one at a time. I’ll read what you said and then I want you to show me where it is on the blot and then tell me what there is there that makes it look like that so that I can see it too, just like you did. Is that clear? (p. 13)

CS standardization guidelines include reading instructions verbatim and avoiding nonverbal gestures and irrelevant verbalizations that could influence individuals’ responses (Exner, 1993). It is also recommended that the examiner sit beside the examinee during test administration to avoid unintentional nonverbal communication.

**Scoring the CS**

Exner (1993) refers to the content of examinees’ verbalized responses as “percepts” (e.g., “That looks like a black cat”). The following three components of each percept are evaluated by completing eight or nine distinct steps per response: (a) the inkblot location, or particular area(s) of an inkblot the percept was identified; (b) the determinants, or characteristics of the inkblot that contributed to the subject perceiving the ink as a particular percept (i.e., color, shape, texture, etc.); and (c) the specific content, or identity/nature of the percept, organized into discrete categories (i.e., humans, animals, nature, etc.). Numerous variables may or may not be applicable/scored, based on the percept, its attributes, and extraneous content verbalized during the response. By examining the aforementioned components of verbalized percepts, the CS evaluates the
following: (a) the quality of examinees’ cognitive processing (i.e., Developmental Quality); (b) examinees’ perceptual accuracy (i.e., Form Quality), relative to normative, age-based populations; (c) the commonness of examinees’ responses, relative to a normative sampling (i.e., Popular responses); and (d) the examinee’s level of cognitive organization (i.e., Organizational Activity). Finally, the Special Scores section of the CS evaluates uncharacteristic verbalizations indicative of cognitive dysfunction. The subject is required to provide at least 14 responses in order to derive a valid interpretive profile (Exner, 1993). A more specific outline of the steps involved in scoring the CS is included in Appendix B.

Interpreting the CS: the structural summary

Once scored, the numerous CS variables applicable to examinees’ responses are quantified, re-arranged, and integrated into numerical formulas to derive interpretable data presented in the Structural Summary (Exner, 1993). These formulas represent ratios, based on the frequency of variables and variable combinations, relative to an empirically based continuum of functioning. The Structural Summary contains the following three types of interpretive data: (a) frequency statistics for the numerous individual variables assessed by the CS; (b) seven subsections comprised of combinations and ratios of variables that evaluate various types of cognitive and affective processing; and (c) six special Indices, or Constellations, that assess cognitive-perceptual proclivities related to specific types of categorical psychopathology. For convenience, a listing of variables and variable formulas relevant to depressive symptoms is included in Appendix C.

Theoretical Basis and Applications

The Comprehensive System’s creators have been criticized for failing to specify an underlying theoretical basis for their instrument (Exner, 1997; Weiner, 1989, 1997). Responding to this criticism, Exner (1997) unfortunately makes no acknowledgement of any theory supporting the use of the CS. To the contrary, he refutes the significance of the criticism by citing the common use and wide clinical acceptance of the MMPI-2, another purportedly atheoretical instrument. Ostensibly, Exner claims to endorse a
statistical theoretical backing; however, this is insufficient as statistical principles do not account for the way in which the CS is interpreted.

Despite its lack of published theoretical rationale, the Comprehensive System is clearly supported by a rich tradition that shares multiple similarities with established cognitive theories and research. Exner (1993) openly acknowledges that the CS takes into consideration the cognitive processes individuals utilize while engaged in active perception of Rorschach stimuli, which is congruent with both information processing and clinical cognitive models. Moreover, the CS has been implicated in a number of cognitively-oriented studies (see Blatt, 1990 for further details).

**Perception vs. cognition**

Discussions regarding the Rorschach task tend to focus on the interrelated terms perception and cognition (Exner, 1993). Blatt (1990) defines perception as “the relatively clear-cut recognition and relatively immediate reading of primarily figurative aspects of available stimuli” (p. 398). Leichtman (1996b) asserts that perceptual processing remains the cornerstone of the Comprehensive System’s success, as clinicians throughout the past have invariably assumed that the Rorschach is capable of extrapolating broad patterns of routine perception from the brief perceptual samples exhibited during a given testing session. This assumption has come to be known as the “perception hypothesis” (p. 398). A sizable body of research leading up to the development of the CS has investigated the projective qualities of perceptual processing (Rapaport, Gill, & Schafer; Schachtel; Exner & Weiner, as cited in Leichtman, 1996b).

Cognition refers to the mental activity individuals utilize to process their perceptions of stimuli (Leichman, 1996a, 1996b; Exner, 1993). Baddeley (1990) notes that cognition is premised on memory, which is a cumulative product of perception-based learning experiences. Although they are conceptualized as two distinct processes, cognition and perception are inextricably intertwined, and occur virtually simultaneously. Blatt (1990) notes that successful processing of a stimulus depends on a subject’s construction of meaning from perception. In other words, perception must be subjectively transformed into mental representations (i.e., cognition, or imagery), which Baddeley (1990) notes are qualitatively different than actual sensory perception (i.e.,
visualizing the inkblot stimuli). Thus, the perception hypothesis presumes the acceptance of an implicitly constructivistic theoretical approach to cognition (Exner, 1996; Exner, Thomas, & Mason, 1985).

**The limits of constructivism**

Despite individual differences in cognitive processing, a functional society must maintain a consensus regarding consistent, commonly-experienced aspects of reality (Mahoney, 1995). This assertion is consistent with the considerable body of research leading to the development of the CS (Attneave; Hochberg & McAllister; Berkeley; Koffka; Mach, as cited in Exner, 1996). In order to establish the parameters of atypicality, the Rorschach must first provide an estimate of the degree to which individuals’ typically perceive and verbalize aspects of reality in consistent, agreed-upon ways (Exner, 1993, 1997; Lichtman, 1996b). Stated differently, by establishing a normative baseline of cognitive-perceptual consensus, the CS is capable of discriminating commonly shared responses from uniquely idiosyncratic and psychopathological ones. For example, Exner (1996) notes that approximately 85% of individuals exposed to card V of the Rorschach inkblot test report either the percept of a bat or a butterfly, as opposed to an airplane or a banana. Similarly, individuals tend to report particular inkblot features (i.e., color, shading, form, etc.) in predictable proportions. On the basis of cognitive-perceptual deviations from these typical modes of responding, the CS proposes to be able to identify important markers of depressive symptomatology, among other markers. The CS contains two Indices specifically designed to assess depressive symptomatology: (a) the Depression Index (DEPI), and (b) the Coping Deficit Index (CDI) (Exner, 1993).

**The Depression Index (DEPI)**

A revised version of the DEPI was implemented in 1990 due to the diagnostic imprecision of the former version (i.e., 10%-30% false positive rates and >60% false negative rates among adult samples)(Exner, 1990). Exner (1990, 1993) notes that the current, revised DEPI demonstrates higher discriminant validity. In 1990, preliminary studies using the DEPI revealed that 69 out of 100 adolescents previously identified as being clinically depressed, using DSM-III criteria, demonstrated significantly elevated
scores (Exner, 1990). Also, a reported false positive rate of less than 2% has been reported among a large sample of non-clinical children (Exner, 1993).

The DEPI is comprised of 15 variables, and consists of an actuarial measure of depression ranging in value from 0 to 7. Scores ranging from 0-4 are considered non-significant from an actuarial perspective, while scores of 5 indicate the presence of some depressive features without the likelihood of corroborating a definitive depressive disorder. Exner (1993) notes that scores of 5 on the DEPI Index may be indicative of cognitive-based susceptibility to affective fluctuations. Scores of 6 and 7, however, are deemed “much more definitive” (p. 360), and thus considerably more suggestive of symptoms consistent with a categorically validated depressive disorder.

Exner (1993) describes the DEPI as being comprised of “a variety of affective, cognitive, and interpersonal variables” (p. 489) represented by the following five derivations (please refer to Appendix B for a more detailed description of CS variables): vista responses (V), color-shading blends (C.Sh), the egocentricity index (EI)(3r + (2)/R), achromatic color responses (C’), and morbid content scores (MOR) (Exner, 1993; Viglione, Brager & Haller, 1988). Additionally, Exner and Weiner (1995) note that comparative analyses of certain Rorschach markers also hold diagnostic utility. For example, a greater proportion of shading responses (SumSh) than combined FM and m responses has been found to be indicative of “dejection or distress” (p. 191), and combinations of two or more of the aforementioned variables are suggestive of “considerable dysphoria” (pp. 191-192). Past research has identified correlations between these constructs and such symptoms as dysphoric affect, negative self-image, lack of orientation/responsiveness to environmental demands, and emotional instability/distress (Exner, 1986; Isen, 1984, Johnson & Magaro, 1987; Martin, 1985; Watson & Clark, 1984, as cited in Meyer, 1992).

In addition to empirical studies, many DEPI variables have intuitive similarities to clinical cognitive and cognitive information processing theories, particularly relating to subjective depressive symptoms (please refer to Appendix C for a summary of these similarities). For example, morbid content scores (MOR), which are assigned to verbalizations containing negative themes, are found to be consistent with numerous cognitive constructs such as Beck’s (1967) cognitive triad and negative self-talk. Low
proportions of reflection responses (Fr or rF, relative to overall R), such as those evaluated using the egocentricity index, are also intuitively consistent with Exner’s (1993) interpretation of pervasively negative self-appraisal; it appears reasonable to assume that individuals who think little of themselves will generalize that negative perception to others and the environment (e.g., Beck’s cognitive triad), and will minimize the influence (i.e., reflections are manifestations of presence, and perhaps importance) of percepts.

As previously mentioned, subjects’ Rorschach responses are based on meanings that have been encoded into their cache of long-term and working memory as the result of repeated cognitive-perceptual experiences (Baddeley, 1990; Blatt, 1990; Exner, 1996; Exner, Thomas, & Mason, 1985). Among cognitive theories, the valence of individuals’ cognitions rather than their specific content has been given precedent in empirical research. From a non-empirical perspective, however, meanings associated with particular verbal statements are virtually taken for granted among modern society, and have come to pervade the English language. Persons reference chromatic and achromatic color by stating they feel ‘blue’ and by speaking of unfortunate times as ‘dark days,’ and the literary concepts of light and shadow have been used throughout the ages to portray concepts of good and evil. Additionally, diminutive words have often come to connote unfavorable mood states. For example, a remorseful or sad person might state that he or she feels ‘low,’ ‘down,’ or ‘small.’ Among society, these terms are clearly not definitive indications of depressive symptoms independent of numerous other considerations; however, their meaning within the context of a purportedly neutral and ambiguous test situation is likely to be much more significant. Thus, it is not unreasonable to posit that depressive symptoms coincide with a greater number of responses involving achromatic color (C’) (e.g., excessive attention to the dark or drab features rather than form) or number of vista responses (V), given that, by nature, individuals’ subjective responses are self-referential (Blatt, 1990).

The Coping Deficit Index (CDI)
The Coping Deficit Index (CDI) is reported to be a direct result of research conducted to improve the diagnostic precision of the DEPI (Exner, 1993). The CDI range extends from 0 to 5. Using factor analysis, Exner and Weiner (1995) conceptually delineated depression into the following three typologies: (a) emotionally distraught individuals; (b) cognitively pessimistic, lethargic, and self-defeating; and, (c) “helplessness depression” (p. 221), in which individuals are found to be ill-equipped to address and successfully resolve environmental demands. The three typologies were found to overlap considerably; however, the DEPI was found to best assess typologies 1 and 2, while elevated scores on the CDI (i.e., 4 or 5) were more closely associated with typology 3 (i.e., helplessness depression). As such, Exner (1993) and Exner and Weiner (1995) advocate that the DEPI and the CDI be considered co-jointly in order to order to increase the probability of accurately assessing the presence of a clinically significant depressive symptoms.

Although not designed to be a measure of depression per se, the addition of the CDI has been found to account for considerably greater numbers of depressed individuals than those identified by the DEPI alone (Exner, 1993). Exner & Weiner (1995) note that in 79% of studies in which individuals diagnosed with a depressive disorder were not correctly identified using the DEPI, their responses yielded scores of 4 or 5 on the CDI. Unfortunately, there is a dearth of research corroborating these claims; the most current research evaluating both the DEPI and CDI co-jointly remains Exner’s (1993) standardization sample of 315 inpatient subjects utilized in the formation of the revised DEPI and CDI. Among this sample, 50% demonstrated clinically significant score elevations on the DEPI, 18% on the CDI, and 25% for both of these scales.

The CDI is calculated using the following five variable formulas (please refer to Appendix B for a more detailed description of CS variables)(Exner, 1993):

1. Experience Actual (EA) (i.e., M + the sum of FC, CF, and C responses) < 6, or Adj. D (i.e., EA – [the sum of FM + m + C’ + T + Y + V])< 0

2. Cooperative (COP) responses < 2 and Aggressive (AG) responses < 2
3. The weighted Sum C < 2.5, or Afr (i.e., the number of responses on cards VIII through X / the number of responses on cards I through VII) < .46

4. The number of passive responses > Active + 1 (Active:Passive Ratio, or a:p), or the number of Pure H responses < 2

5. The Sum T > 1, or the Isolation Index (Isolate/R, or the ratio of Sum Botany, Clouds, Geography, Landscape, and Nature responses: R) > .24, or the number of Food responses > 0

Exner and Weiner (1995) posit that symptoms of ‘helplessness depression’ result in impairments in social functioning, as well as a preponderance of interpersonal problems compared to more overt depressive symptoms. Exner (1993) proposes that such coping deficits contribute to individuals’ increased vulnerability to various, sundry problems that may superficially appear unrelated to the concept of depression, but in actuality are likely to result in symptoms of depression. Individuals exhibiting elevated scores on this index are likely to exhibit despondency and reluctance to deal with affect, and are more likely to exhibit substance abuse (L. Kubiak, personal communication, June, 2000).

The theoretical foundation of the CDI corresponds best with the learned helplessness (Abramson, Metalsky, & Alloy, 1989) and cognitive resource allocation (Ellis et al., 1989) models. The former model bears striking similarities to Exner and Weiner’s (1995) conceptualization of helplessness depression, while the latter model conceptualizes the outcome of limited cognitive resources with which to cope with life’s demands. According to the cognitive resource allocation model, it is reasonable to assume that individuals overwhelmed by symptoms of depression should produce low scores on the Experience Actual (EA) cluster, which assesses the degree of allocated cognitive resources uninhibited by affective distractions (Exner, 1993). The EA cluster’s Human Movement (M) variable has been associated with “reasoning, imagination, and a higher form of conceptualization” (p. 419), while the prevalence of Chromatic Color (C) responses has been linked in numerous instances with affect and affective modulation.
Low Adj. D involves multiple variables that are related to limited stress tolerance (see Appendix C). Low Adj. D., and hence limited coping ability on the CS, could be attributable to limited allocation of cognitive resources. Many CDI variable clusters also appear to correspond with cognitive outcomes predicted by the Learned Helplessness model (Abramson, Metalsky, & Alloy, 1989). For example, minimal cooperative (COP) and aggressive (AG) types of responses, as well as a higher proportion of passive responses, appear characteristic of hopelessness depression.

Current Controversy: An Ongoing Critique of the Rorschach CS

For some time, researchers have been embroiled in an unresolved controversy over a number of issues relating to the CS, including its ability to accurately assess depression. These two positions have voiced diametrically opposed views on the basis of identical data, and have advocated both the banishment of the CS (Garb, Wood, Nezworski, & Grove, 2001; Wood, Lilienfeld, Garb, & Nezworski, 2000; Wood, Nezworski, & Stejskal, 1997) as well as its broad use in a number of clinical applications (Exner, 1997; Weiner, 2001; Weiner, Exner, & Sciara, 1996). Hunsley and Bailey (1999, as cited in Garb et al., 2001) state:

“The Rorschach has the dubious distinction of being, simultaneously, the most cherished and the most reviled of all psychological assessment tools...The Rorschach is held in great esteem by many psychologists for its ability to access intrapsychic material, whereas others point to the Rorschach as a prime example of unscientific psychological assessment” (p. 266).

Critiques of CS Psychometric Properties

The psychometric stability of the CS has been widely evaluated by numerous research studies throughout the years (including meta analyses), recent publications expressly conducted to summarize and reconcile the issue of CS clinical use (Meyer, 2001; Meyer & Archer, 2001), and judicial authorities (McCann, 1998; Meloy, Hansen, & Weiner, 1997; Weiner, Exner, & Sciara, 1996). From a clinical perspective, the use of idiographic and nomothetic interpretation, either co-jointly or respectively, has been
supported for years (Acklin, 1995; Aronow, Reznikoff & Moreland, 1995; Exner & Weiner, 1995; Moreland, Reznikoff & Aronow, 1995; Weiner, 1986); however, legally and psychometrically, idiographic interpretation may be unacceptable in cases involving human placement and appropriate treatment (Blatt, 1990; McCann, 1998; Meloy, Hansen, & Weiner, 1997; Weiner, 1996; Weiner, Exner, & Sciara, 1996).

Legal/judicial standards assert that the admissibility of a given psychometric instrument is dictated by (a) its helpfulness to the jury, as indicated by its endorsement by a substantial proportion of psychological practitioners (Federal Rules of Evidence, 1992, as cited in McCann, 1998), and (b) its ability to withstand empirical scrutiny, including ascertainment of its degree of test error, by peer review processes (the Daubert standard) (Daubert v. Merrell Dow Pharmaceuticals, Inc, 1993, as cited in McCann, 1998). Thus, from a legal perspective, the Rorschach’s acceptable empirical standards embody the cornerstone of its psychometric validity. Several studies conducted in recent years have concluded that the Rorschach, administered and scored according to the guidelines specified by the CS (Exner, 1993), fully satisfies standards for admissibility established by American legal precedents (McCann, 1998; Meloy et al., 1997; Weiner, Exner, Jr., & Sciara, 1996).

Weiner (1997) notes that the Rorschach, typically scored using the CS, remains the second-most commonly researched instrument among prominent psychological journals, endorsed as ‘routinely used’ in assessment by 82% of a sample of 412 clinical psychologists. Additionally, demographic studies indicate that 85% of psychological programs accredited by the American Psychological Association in 1993 affirmed having taught the Rorschach to matriculated students (Piotrowski & Zalewski, 1993, cited in McCann, 1998). The CS is credited for its inclusion of a diverse normative sample generalizable to diverse racial, geographic, and socioeconomic populations (Krall, Sachs, Lazar, Rayson, Growe, Novar, & O’Connell, 1983); however, as acknowledged by Exner and Weiner (1995), the CS standardization sample was based on modest samples for respective age groups of children and adolescents. Also, it is unfortunate that the majority of research evaluating the CS depression indicators was conducted exclusively using the CS standardization sample (Meyer & Archer, 2001). Overall, however, the
most dominant and recurrent criticisms stem from a lack of consensus on what constitutes acceptable psychometric stability for the CS, largely due to its unconventional nature.

**Variable R**

Lipgar (1992) notes that R, or total number of responses, has been conceptually and empirically problematic to measuring internal consistency coefficients, as the CS allows subjects to give a variable number of responses, overall. While this presents the greatest challenge to research comparing the CS with other depression instruments, inadequate R is also detrimental to the Comprehensive System’s psychometric validity, and possibly related to resistance (Blatt & Berman, 1984; Exner, 1997, 1990; Exner & Weiner, 1982; Finch, Imm & Belter, 1990; Lipgar, 1992). As a result, profiles must be based on at least 14 responses in order to be considered valid and interpretable (Exner, 1993). With this exception, McCann (1998) concludes that variable R does not detract from CS psychometric validity. Rather, test-retest and inter-rater reliability remain more amenable to research, compared to the difficulty of obtaining internal consistency ratings.

**Test-retest reliability**

Overall, Exner, Thomas and Mason (1985) report “extremely high” test-retest reliability coefficients for CS variables over relatively short periods of time (i.e., less than nine months) for non-clinical children and adolescents ranging in age from 8 to 16 years over a two-year interval. In general, McCann (1998) notes that these short-term test-retest coefficients exceeded .80 for children, as well as for adults, establishing acceptable test-retest reliability for legal and clinical applications; however, Exner and Weiner (1995) note that test-retest coefficients fluctuate to a greater degree for child and adolescent populations over longer intervals. This is to be expected, given the dynamic and rapid cognitive changes underway throughout normal development. It is noted that such low retest correlations do not signify a lack of validity inasmuch as they suggest that the CS may not predict children’s and adolescents’ responses over longer intervals as accurately as adult responses. For this reason, many researchers have emphasized the importance of carefully considering developmental factors when interpreting the CS
protocols of children and adolescents (Exner & Weiner, 1995; Ornberg & Zalewski, 1994).

Some CS variables predictably exhibit very low test-retest stability for subjects of all ages due to their transient nature (i.e., state vs. trait). For example, the m and Y variables both involve subjects’ perception of situational stress, which would not be expected to endure over a substantial period of time. For the m variable, Exner and Weiner (1995) cite consistently low test-retest coefficients (ranging from .20 to .49) for children of varying ages, as well as adults, after seven-day and three-week intervals, respectively. Individuals of a broad range of ages also unilaterally exhibited low test-retest coefficients (i.e., ranging from .17 to .42) for the Y variable, which measures tolerance of stressful situations.

**Inter-rater reliability**

The Rorschach CS has been criticized for its complex administration and scoring procedures (Viglione, Brager & Holler, 1988; Weiner, 1997). As a result, its inter-rater reliability (IRR) has undergone considerable empirical scrutiny. Of note, these studies have primarily utilized adult subjects. McDowell and Acklin (1996) estimated inter-rater coefficients using Kappa coefficients for each of the scoring components/variables associated with the CS. These authors concluded that scores ranging from .41 to .60 indicated moderate agreement, and scores ranging from .61 to .80 indicated substantial agreement. The majority of coefficients fell within substantial-to-near-perfect agreement, indicating consistently high inter-rater reliability, overall. Results yielded moderate agreement of .59 relating to the scoring of Rorschach specific scores, and agreement of .68 at the lower end of the substantial range for scoring of Form Quality (FQ) scores.

Archer and Krishnamurthy (1997) also concluded that the CS yielded acceptable Kappa coefficients (between .66 to .98) for all CS variables, and McCann (1998) cites the ample test-retest reliability estimates provided by Exner (1993) as a testament to the CS’s inter-rater reliability. In the most recent study to date, Acklin et al. (2000) calculated Intra Class Coefficients (ICC), as well as Kappa coefficients, to evaluate IRR using both clinical and non-clinical samples of adults. Overall, based on the aforementioned
standard established by McDowell and Acklin (1996), the majority of CS variables (including depression indicators) yield substantial inter-rater consistency. With respect to CS depression indicators, exceptions include low agreement with regard to form vs. color dominance (i.e., FC: CF + C of .543 and .165), and Vista (Sum V) responses (i.e., .576).

Disparate conclusions have been reached in various studies addressing both the IRR of numerous CS variables, as well as the optimal statistical procedures with which to derive IRR coefficients (Acklin, McDowell, Verschell, & Chan, 2000; Exner, 1997; McDowell & Acklin, 1996; 1993; Weiner, 1986, 1997). For example, Garb, Wood, Nezworski, Growe, and Stejskal (2001) globally criticize the CS on the basis of Acklin et al.’s (2001) IRR study, although they adhered to a high standard of comparison (i.e., ICC coefficient values of .85). Viglione and Hilsenroth (2001), however, cite seven separate studies demonstrating “conclusive empirical evidence of strong interrater reliability for the great majority (95%) of CS variables.” Overall, the clear majority of studies concurs that the CS demonstrates adequate IRR. As adequate inter-rater reliability coefficients have been consistently required for some time by journal publications in which the Rorschach is routinely featured, Weiner (1997) notes that lack of rater agreement among competent professionals does not appear to be problematic.

Related to IRR, one study investigated the scoring accuracy of the CS (Guarnaccia, Dill, Sabatino, & Southwick, 2001). In contrast to IRR, which assesses the degree to which examiners agree on how CS variables should be coded, scoring accuracy evaluates whether or not such agreement concurs with CS scoring criteria. Thus, hypothetically, it is possible for examiners to demonstrate high inter-rater agreement while simultaneously exhibiting low scoring accuracy. Guarnaccia et al. (2001) assert that greater instances of scoring errors occurred using patient, as opposed to non-patient, responses. Unfortunately, the authors did not provide specific variables found to be problematic for scoring accuracy; however, they stated that higher instances of errors occurred for Determinants, Form Quality (FQ), Z scores, and Special Scores. Overall, this does not appear to be exceedingly problematic to the validity of the CS, although both Guarnaccia et al. (2001) and Viglione and Hilsenroth (2001) advocate the need for adequate examiner training in the use of the CS.
Convergent Validity

The CS depression indices have been criticized for demonstrating low convergent validity with other forms of dimensional assessment (i.e., self-report forms); however, results are somewhat mixed. Criticisms have particularly targeted the DEPI, as the CDI represents a relatively unique measure of depression that has not been addressed extensively to date. In one study, Gordon & Tegtmeyer (1982) demonstrated low (unspecified by the researchers), yet significant, correlations between the Child Behavior Check List (CBCL) and the DEPI index. Using a sample population of adolescents receiving either inpatient or outpatient psychological services, Archer and Krishnamurthy (1993) demonstrated no significant relationship between any of the MMPI-A variables (including the 2, or Depression scale) and the DEPI. While a similar lack of correlation was observed in a subsequent study addressing the CS DEPI and the MMPI-A (Archer & Krishnamurthy, 1997), significant relationships were noted in the latter study between the MOR variable and scale 2 (r=.22), scale 9 (r=.23), and the A-DEP scale (r=.27). Overall, the authors noted that correlation coefficients of equal to or less than .10 were obtained for all other CS and MMPI-A depression indices.

Other studies have investigated the DEPI’s convergent validity using adult populations. One such study attempted to derive correlations between the CS DEPI and several self-report measures (i.e., the BDI, the Multiple Affect Adjective Checklist-Revised Depression scale, and the Personal Feelings Questionnaire) (Greenwald, 1997). Despite the slight over-representation of apparently depressed individuals among an allegedly non-clinical population of college students (i.e., nearly 15% across combined instruments), consistent lack of correlation was found between these measures. As such, it was concluded that the convergent validity of the Rorschach DEPI may be somewhat lacking among non-clinical populations; however, a substantial methodological shortcoming was noted in this research due to the fact that the author failed to specify any means, other than the Rorschach CS itself, by which diagnoses of depression were verified in the study.

Another adult study comparing the Rorschach and the MMPI-2 Harris-Lingoes Scales utilizing an analysis of variance (ANOVA) technique found that individuals obtaining significant elevations on the DEPI also scored significantly higher on the
MMPI-2 Mental Dullness scale (Khouri & Greenway, 1996). Of the individuals obtaining significant DEPI elevations, those exhibiting greater than four affective blends scored significantly higher on the Subjective Depression and Mental Dullness subscales than all other depressed and non-depressed groups. Subjects exhibiting significantly elevated scores on the Isolation Index also scored significantly higher on the aforementioned MMPI-2 subscales. Overall, subjects obtaining significant DEPI scores (≥5) scored significantly higher on the Mental Dullness subscale than non-elevated DEPI scorers.

Overall, it is noted that findings of low convergent validity, while puzzling, are not necessarily problematic. While lacking a satisfactory explanation, this conundrum has also been noted to commonly occur among different types of self-report forms designed to measure the same constructs (Kazdin, 1987; Worchel et al., 1990). Ganellen (2001) states that the CS may measure similar, yet different constructs than its counterparts, which will be discussed in the following section.

**Discriminant Validity**

CS results pertaining to discriminant validity have been mixed. Wood, Lilienfeld, Garb, and Nezworski (2000) and Wood, Nezworski, and Stejskal (1997) cited overwhelmingly negative findings pertaining to the Rorschach’s ability to discriminate between criterion-based depressed vs. non-depressed individuals, concluding that the CS DEPI constitutes an unreliable and invalid assessment of depression among individuals of all ages. It is noted that the aforementioned authors have drawn negative conclusions regarding the validity of the CS on numerous occasions, whereas other authors have rejected such conclusions based on the same empirical data (Ganellen, 2001; Meyer & Archer, 2001; Viglione & Hilsenroth, 2001). Additionally, subsequent research has refuted these claims to some degree on the basis that coefficients were derived using the older, unrevised version of the DEPI (Ganellen, 2001; Weiner, 2001). Additionally, the majority of Wood et al.’s (1997; 2000) research appears to pertain exclusively to adult populations.

Recently, more impartial reviews of the literature have been published with regard to the DEPI’s discriminant validity. Ganellen (1996) demonstrated by evaluating Exner’s
standardization data that the Rorschach depression indices appear to offer promising discriminant validity with regard to all aspects of psychometric efficiency, based on approximately equal populations of depressed and non-depressed inpatient adults (N=315 depressed, 320 non-depressed). Rorschach specificity coefficients (i.e., ability to correctly identify other DSM-IV-verified disorders from depression) were obtained at the .81 level, with a low false positive rate of .23. Further, Ganellen’s (1996) study yielded a specificity coefficient of .75, with a false negative rate of .25. The depression indices’ positive predictive power was assessed to be .79, and its negative predictive power at .76. Although these results were drawn using adult populations, they hold promising utility for younger populations. Partial, yet promising discriminant validity has also been demonstrated by Spigelman & Spigelman (1991), as evidenced by statistically significant increases in DEPI scores with children in a non-clinical setting who had experienced parental divorce, relative to their non-divorce counterparts.

Exner and Weiner (1995) acknowledge recurrent problems relating to the discriminant validity of the DEPI, and suggest a need for further studies corroborating this index’s utility with younger individuals. They note that, while a substantial proportion of children and adolescents exhibiting elevated DEPI scores tend to have corroborating depressive diagnoses, a substantial number of false negative results also occur. Ball, Archer, Gordon, and French (1991) have corroborated this conclusion. Although these findings tend to be somewhat discouraging, they do not dismiss the DEPI as an invalid measure of depression; rather, they have prompted researchers to consider the way in which DEPI results are interpreted. Finally, with the exception of Exner’s (1993) studies involving the initial CS standardization sample, empirical research examining the effectiveness of the CDI in discriminating depressed vs. non-depressed individuals are conspicuously absent from the literature (Weiner, 2001; Meyer & Archer, 2001).

**Changes as a Result of Treatment**

While extremely difficulty to predict, the course and recurrence of depressive syndromes in child and adolescent populations clearly affect the stability of CS depression indices over time. Miller, Bakalar, & Grinspoon (2002) note that single-
episodes of major depression are estimated to last approximately two weeks, and for considerably longer periods for dysthymia. Another review of epidemiological studies estimated the average length of a Major Depressive episode to last between seven to nine months, and the average length of a Dysthymic episode being four years (Birmaher, Ryan, Williamson, Brent, Kaufman, & Dahl, as cited in Cicchetti and Toth, 1998). Additionally, diagnoses are substantially more difficult to substantiate for children and adolescents, and the effects of treatment interventions confound longitudinal studies addressing the stability of depression-related constructs.

It is reasonable to assume that clinically significant elevations in CS depression indices will decrease subsequent to treatment interventions. Although a dearth of research studies has addressed this issue, promising results have been found in one study (L. Kubiak, personal communication, July 2000). Among a population of 50 depressed adolescents referred to a residential treatment facility, significant changes in the prevalence of CS depression-related variables were noted based on a two-year post-test. Among these changes, none of the four adolescents initially obtaining a DEPI score of 5 of greater exhibited significant post-test elevations. A more dramatic decrease was noted with respect to CDI scores, as only four of 35 adolescents exhibiting pre-treatment CDI elevations greater than three demonstrated significant post-test CDI scores.

Significant decreases were also noted to some degree on the variable level. Approximately 35 of the 50 individuals exhibited Experience Actual (EA) scores of less than seven, in contrast to approximately 29 post-treatment individuals. Low Egocentricity Index scores [i.e., 3r + (2)/R > .33] also decreased significantly from approximately 25 of the 50 initially referred adolescents to only one individual in the post-treatment group. Low affective ratios (Afr>.50) also decreased in approximately 26 of the 29 adolescents.

**Conceptual and Practical Considerations**

A number of conceptual and practical considerations merit mentioning that may help to explain the aforementioned critiques, and which serve to differentiate the CS from more traditional measures of depression. For one, Meyer and Archer (2001) have noted that conscious awareness of depressive symptomatology is not required in order for a
subject to yield elevated DEPI or CDI scores. Proponents of the Rorschach throughout the past have investigated relationships between particular types of Rorschach responses and depressive proclivities (Blatt, 1990; Exner, 1993; Weiner, 1997). As such, the Rorschach’s unique approach to assessing depression does not share the face validity of checklist format inventories. It also requires a greater degree of inference on the part of the interpreter, and is not characteristic of all of the CS indices of psychopathology. For example, the CS Schizophrenia Index (SCZI) evaluates overt behavior such as verbalized cognitive slippage or abject distortion of reality. As such, the methodology of the CS depression indices breaks significantly with categorical classification (i.e., the DSM-IV), which requires subjects to endorse their tangible cognitive and affective symptoms, as well as to exhibit observable behavior consistent with these symptoms.

Viglione (1996) notes that the present-focused approach of the CS also differentiates itself from its retrospective self-report counterparts. That is to say, in contrast to the CS, many self- and other-report inventories require subjects to estimate the frequency of their depressive symptoms within a past period of time. This discrepancy in format may account for findings indicating low CS convergent validity with pencil-and-paper types of assessment; however, it is noted that low convergent validity is also commonly exhibited among self- and other-report formats sharing similar retrospective approaches (Kazdin, 1987; Worchel et al., 1990).

Researchers have debated the CS’s applicability towards assessing depression with children and adolescents due to their premature level of development (Blatt, 1990; Exner, Thomas, & Mason, 1985; Leichtman, 1996b). Leichtman (1996a) questions whether younger children possess the cognitive, social, and emotional sophistication needed to respond to Rorschach inkblots in diagnostically revealing ways. For one, the ability to specify particular details relating to inkblot stimuli depends ultimately on verbal expressiveness, which may be considerably lacking in early childhood. In addition, Baddeley (1990) raises the possibility that very young children lack the opportunity to encode and integrate sufficiently sophisticated information relevant to the Rorschach task. In other words, these individuals would possess a limited cache of information within long-term memory due to lack of formation of meaning structures by association of these stimuli with predictable outcomes. Perhaps this insufficient
development/paucity of learning experiences accounts for Finch et al.’s (1990) observations that younger children provide fewer responses to Rorschach stimuli.

While there is no established, pre-requisite age for administering the Rorschach, research demonstrates that most children five years of age possess the cognitive resources necessary to participate meaningfully in the CS administration activity (Exner, 1993; Exner, Thomas, & Mason, 1985; Exner & Weiner, 1995); however, Leichtman (1996b) notes that it is often necessary for clinicians to modify CS standardized guidelines when administering the Rorschach to young children in order to enhance their comprehension of the task demands. It is admittedly difficult to predict the performance of young children on the CS, yet equally difficult to ascertain a child’s level of accomplished cognitive and emotional development at any given period of time.

As the majority of CS research has been based exclusively on adult samples, further study is warranted to investigate CS validity and diagnostic utility. Exner (1997) has speculated that the paucity of adolescent and pre-adolescent samples is likely due to the relatively lengthy amount of time required to administer the Rorschach, and the particular difficulty children experience sustaining their attention throughout lengthy series of tests; however, a vast quantity of research overwhelmingly accepts the CS’s compatibility with these populations when considered in tangent with developmental and situational (i.e., test-taking behavior, etc.) factors (Exner, 1993; Exner & Weiner, 1995; Meyer & Archer, 2001; Viglione & Hilsonroth, 2001, among others). Additionally, it is noted that the lexical format of more traditional (i.e., self-report) depression instruments is considerably more cognitively taxing to young children than that involved in administering the Rorschach. Younger children are more likely to lack the reading skills necessary to comprehend these tasks, or to fail to subjectively appreciate their overt symptoms of depression.

Overall, the Rorschach’s qualitative uniqueness both complements and challenges its clinical utility. First, the Rorschach does not exhibit the same degree of demand characteristics as paper-and-pencil types of tests. As a result, particularly when administered to resistant, guarded, and/or oppositional children and adolescents, the task may be perceived as more enjoyable and/or less personally threatening or revealing (Exner, 1993; Leichtman, 1996a; Dr. L. Kubiak, personal communication, June 2000;
Weiner, 1997). Pertaining to interpretability, it is important to hold a conceptual and practical understanding of the utility of the CS, overall. In assessing depression, the Rorschach is similar to dimensional types of assessment instruments/scales in its ability to calculate the probability of a depressive disorder on an individual case basis (Exner, 1993). In this regard, numerous researchers have advocated that a primarily pragmatic, rather than exclusively conceptual, utilization of Rorschach’s psychometric capabilities be undertaken (Ganellen, 1996; Meyer & Archer, 2001; Viglione, 1996; Viglione & Hilsenroth, 2001).

Reconciling the Controversy

The aforementioned research has demonstrated the significance of developmental, as well as psychometric factors affecting the assessment of depression with the CS. As previously mentioned, children and adolescents’ CS responses tend to be more consistent/predictable over shorter periods of time than over longer intervals. As this phenomenon is the product of dynamic developmental maturation (Exner & Weiner, 1995), it is predicted that greater test-retest stability over longer periods of time should occur for many CS depression-related variables during or after the onset of adolescence. Moreover, given that overt behavioral symptoms fluctuate over development, notable changes in the prevalence and nature of children’s vs. adolescents’ manifestation of depression, or lack thereof, should be reflected in their CS responses. Reflected by these manifestations, differences between the psychometric structure (i.e., underlying constructs) of the CS and other common measures of depression may also account to some degree for their lack of convergent validity.

A Comparison of the Factor Structures of Depressive Assessments

Two prominent studies to date have addressed the factor structure of the Rorschach CS, one of which was based on adolescent responses. Using a sizable sample of inpatient adolescents (N=268), Anderson and Dixon (1993) derived an acceptable
three-factor structure solution accounting for the overall variance associated with all Rorschach variables. The authors labeled these factors (a) struggles with affective expression, (b) general psychological health, and (c) perceptual organization. The second factor analytic study was performed using only depression-related variables, and was conducted using responses from adults from the current CS standardization sample (Exner, 1993). That analysis also resulted in a three-factor solution, which was given the following labels: (a) emotionally depressed, (b) generally/cognitively depressed, and (c) helpless; however, it was noted that the emotional and cognitive groups overlapped considerably.

Several studies have investigated the factor structure of the Children’s Depression Inventory, based on populations of clinical and non-clinical children and adolescents. One study utilized Children’s Depression Inventory responses obtained from both clinical and non-clinical children aged seven through 12 (Hodges, Siegel, Mullins and Griffin, 1983). Two separate factor structures were considered optimal for the two, respective groups. Based on the clinical children group, four factors were identified: (a) cognitive-related symptoms, (b) lack of motivation/withdrawal, (c) poor social integration, and (d) somatic components. For the non-clinical group, a two-factor structure resulted that was labeled (a) general (i.e., inclusive of cognitive, affective, motivational, and social), and (b) non-compliance.

Exploratory factor analysis was also performed on the Children’s Depression Inventory by Carey, Faulstich, Gresham, Ruggiero, and Enyart (1987) using equivalent populations of clinical and non-clinical, evenly age-distributed children and adolescents. Overall, the authors report an optimal three-factor solution, which they labeled (a) Depressive Affect, (b) Oppositional Behavior, and (c) Personal Adjustment; however, only the first two factors discriminated the non-clinical from clinical subjects. A more recent study of the Children’s Depression Inventory differentiated between non-clinical populations of children and adolescents, and resulted in a greater number of factors (Craighead et al., 1998). The child group yielded the following results: (a) Externalizing, (b) Dysphoria, (c) Self-Deprecation, (d) School Problems, and (e) Social Problems. The adolescent group yielded the same five factors, with the addition of a sixth: Biological Dysregulation. Acceptable results yielded by the most recent CDI factor analysis study
Cole et al., (2000) suggested a three-factor solution comprised of the following: (a) Social Self-Esteem (including loneliness, poor self-image, and anhedonia), (b) Oppositional-Misbehavior (including misconduct, aggressive behavior, lack of motivation, and guilt), and (c) Dysphoria (including sadness, irritability, and self-loathing). This study involved non-clinical sixth-grade students.

While no studies were found relating to the factor structure of the Reynolds Child Depression Scale (RCDS), an exploratory study was conducted using a translated version of the Reynolds Adolescent Depression Scale (RADS) with a non-clinical population of French-speaking adolescents (Campbell et al., 1994). Although the authors acknowledge the difficulty of applying descriptive labels to their resulting five-factor structure, the following solution was derived: (a) Somatic, (b) Anhedonia, (c) Cognitive, (d) Negative View of Self, and (e) Loneliness.

It is evident that considerable disparity exists among factor analytic results involving self-report measures of children’s depression. Compared to the Children’s Depression Inventory and the RADS, the Rorschach CS appears to more clearly differentiate depression by cognition and affect, further specified by a hopelessness typology (Exner, 1993). In contrast, the factor structure of the Children’s Depression Inventory has been labeled by symptom variation to include such subsidiary (or perhaps comorbid) aspects of depression as social skills deficits (Cole et al., 2000; Craighead et al., 1998; Hodges et al., 1983), school problems (Craighead et al., 1998), somatic problems (Hodges et al., 1983), and oppositional behavior (Carey et al., 1987; Cole et al., 2000; Hodges et al., 1983). This is also noted to occur for the RADS; although the factor solution provided by Campbell et al. (1994) includes a cognitive factor, it is highly similar to the Negative View of Self factor. The five-factor RADS structure likely lacks convergence with the CS’s three factor solution, which may be similar to the factor structure of the RCDS; however, it is noted that all of the aforementioned factor analytic studies (including the CS) include an affective component similar to dysphoric mood.

**Developmental Symptom Manifestation and CS Scoring Patterns**

Exner and Weiner’s (1995) nonpatient CS data cite descriptive statistics for ages five through 16 pertaining to all CS variables. Using this data, Wenar and Curtis (1991)
demonstrated that CS responses from children and adolescents are consistent with developmental psychological theory with respect to cognitive complexity and precision, and social conformity; however, studies have failed to account for developmental changes in depressive symptomatology using CS response prevalence and consistency. Based on Exner’s (1993) factor analytic study, it is anticipated that variables comprising the DEPI would relate more closely to observed symptoms consistent with cognitive and affective symptoms, while those comprising the CDI would encompass symptoms corresponding to hopelessness depression. Ideally, it would be beneficial to compare developmental differences across descriptive data derived from clinically depressed, as opposed to non-patient populations of children and adolescents; however, such data remains unavailable to date (Exner, 1993; 1997; Exner & Weiner, 1995; Weiner, 2001). Additionally, Exner and Weiner’s (1995) normative samples lacked consistency with respect to sample sizes, varying from an N of 80 for the six-year-old group to an N of 140 for the nine- and 16-year-old groups. Nevertheless, broad data from Exner’s (1993) non-clinical adolescent normative population will be examined to elucidate possible age-related differences in CS depression-related variables.

DEPI variables

The Egocentricity Index [3r+(2)/R], a component of the DEPI, is conceptualized as an indicator of sense of self-worth and/or self-absorption (Exner, 1993; Groth-Marnat, 1997). Exner (1993) concluded that EI scores become increasingly stable during adolescence, based on test-retest reliability studies. Test-retest coefficients of .73 were derived for 14-year-olds over a two-year retest interval, and .86 for 15-year-olds tested over a nine-month retest period. Correspondingly, coefficients of .61 were derived for eight-year-olds, .43 for 10-year-olds, .21 for 11-year-olds, and .54 for 12-year-olds over a two-year retest period. In contrast to what would be expected, however, surprisingly high coefficients were derived for five-year-olds (.69), six-year-olds (i.e., .78), and nine-year-olds (i.e., .74) across a two-year period, and for seven-year-olds (.70) over a nine-month retest period. Further, the research yielded surprisingly a low test-retest coefficient for 16-year-olds (.48) over a two-year retest interval (Exner, 1993; Groth-Marnat, 1997).
As lower EI scores would purportedly suggest a decrease in self-esteem or, at the least, self-focus these findings are somewhat curious. It is felt that Erikson’s theory of development (1963) would predict a spike in EI scores among young children (i.e., ages 5-6) due to the challenge of autonomy, and among adolescents (i.e., ages 14-16) due to the challenge of identity development (Erikson, 1963), both of which would presumably predict a greater self-focus. This greater self-focus would correspond with Exner’s (1993) somewhat broad definition of the EI as an indicator of both self-absorption and self-worth. The current test-retest results, with the exception of 9-year-olds and 16-year-olds, also appear to be supportive of epidemiological studies that tend to suggest a higher prevalence of depressive symptoms with age (Angold & Costello, 1995; Harvard Medical School, 2002; Maier, Gansicke, Gater, Rezaki, Tiemens & Urzua, 1999; Robbins, Alessi, Cook, Pooznanski, & Yanchysyn; Strober, Green, & Carlson, cited in Poznanski & Mokros, 1994). As an indicator of self-worth, the EI may reflect a distinct, age-related phenotype of depressive symptoms stemming from greater identity stability (e.g., older adolescents may be more aware of others’ perception of themselves, and hence more self-conscious/reticent to express depressive symptoms than their younger counterparts).

Exner (1993) also predicted that a higher prevalence of achromatic color (C’) responses signify affective, as well as cognitive manifestations of depression. Based on observations of children and adolescents’ overt symptom manifestation, it is difficult to anticipate the developmental course of C’ responses. The CS normative data indicates that these responses remain relatively stable for non-clinical populations of children and adolescents aged seven through 16 (i.e., mean response scores range from 1.25 to 1.15), although fewer of these responses are reported by children aged six and five (i.e., .58 and .63, respectively). Additionally, while this may be accounted for by the sample population, the 15-year-old group demonstrated a slightly greater proportion of these responses (i.e., 1.63). Mean Morbid (MOR) responses were also found to be relatively stable across development, based on the descriptive statistics, ranging from .78 for five-year-olds to .58 for 16-year-olds. Once, again, while not easily explained, seven- and eight-year-olds produced slightly higher MOR responses (i.e., 1.64 and 1.13, respectively).
Finally, while exhibiting high test-retest stability across all ages, Sum V responses are highly atypical of very young children (i.e., prior to age ten). As V is associated with self-critical thoughts (Exner, 1993; Groth-Marnat, Ed., 1997), this variable corresponds with a cognitive conceptualization of depression. Correspondingly, the absence of V responses among younger children (i.e., 10 years old and younger) appears consistent with developmental research in that premature individuals do not seem to possess the cognitive capacity to engage in and verbalize this level of symbolic content.

**CDI variables**

Given Exner’s (1993) findings that the CDI is more consistent with hopelessness typologies of depression, it would be anticipated that adolescents’ responses on this Index would demonstrate greater stability than those of children, given that depressed adolescents have been found to verbalize hopeless symptoms (i.e., feelings/ideation) with greater frequencies. Exner and Weiner’s (1995) normative data have verified this to some extent. Human Movement (M) responses comprising part of the CDI appear to become increasingly stable over chronological age progression, particularly during adolescence. In this regard, adolescents are noted to report progressively more M responses with age. Over a nine-month test-retest period, for example, test-retest coefficients were considerably higher for a group of 15-year-olds (.82) than for a group of seven-year-olds (.46), and essentially rose incrementally over a two-year period (i.e., .48 for six-year-olds, .53 for eight-year-olds, .62 for 9-year-olds, .47 for 10-year-olds, .69 for 12-year-olds, .77 for 14-year-olds, and .87 for adults). Along a continuum, the prevalence of reporting for this variable has been associated with a number of different interpretations consistent with the developmental manifestation of depressive symptoms in adolescence. Adolescents tend to provide greater numbers of human movement (M) responses (Exner, 1993; Groth-Marnat, Ed, 1997), increasing incrementally from a mean of 1.70 for five-year-olds to a mean of 4.31 for 16-year-olds (Exner & Weiner, 1995). Consistent with developmental theory, this increase suggests a greater propensity towards active fantasy life congruent with increasing cognitive sophistication, which could account for the increased prevalence of hopeless ideation/depressive symptoms (Miller et al., Eds., 2001). In other words, individuals during adolescence may become more adept
at cognitively interpreting and conceptualizing their depressive feeling/symptoms in tangible ways, which is congruent with Exner’s (1993) factor analytic research.

Quality of affective modulation, as indicated by increasing proportions of CF+ C and FC responses also seems to become more stable during adolescence, based on Exner and Weiner’s (1995) test-retest coefficients (i.e., FC: .76 and .73 for 12-14 y/o’s and 14-16 y/o’s tested five times over the course of two years, vs. coefficients ranging from .39-.63 for 10-12 y/o’s and 8-10 y/o’s, respectively; CF+C: .72 for the 14-16 age range, and ranging from .70, .47, and .26, respectively, for age ranges 8-10, 10-12, and 12-14). While adolescents appear to verbalize slightly higher instances of color responses than younger children, overall, form begins to gain dominance over color (control over affect) with age (i.e., means steadily increase from .71 for five-year-olds, to 3.43 for 16-year-olds). These findings suggest that the ability to modulate affect increases with age, which has also been corroborated in other studies (Exner, 1972, 1986 as cited in Ornberg & Zalewski, 1994). This finding also corresponds with developmental predictions concerning an increasingly cognitive presentation of depressive symptoms.

Exner and Weiner (1995) provide further evidence that greater control over affect increases with age. The CS normative sample data indicate that Affective Ratio (Afr) responses begin to become more temporally stable at the age of nine, as evidenced by steadily increasing test-retest coefficients (i.e., .51 for children aged 6 and 8, and .79 for those aged 9 and 11, tested over the course of two years). A series of five re-tests, also conducted with the normative sample over the course of two years, yielded a higher overall test-retest coefficient (i.e., .76) for a 14-16-year-old group, as opposed to three discrete groups involving children/adolescents ranging in age from eight years to fourteen years (i.e., coefficients ranging from .33 to .69). Interestingly, however, mean descriptive statistics did not increase dramatically throughout the normative data for five-to 16-year-olds, ranging from a mean of .88 at age five, to a mean of .65 at age 16.

The frequency of Texture (Sum T) responses stabilized somewhat for the 9 year/11 year-old population (i.e., test-retest coefficient of .84), as opposed to a 6 year/8 year-old population (i.e., .68). From a developmental perspective, as T and TF response are found to be associated with feelings of interpersonal dependence, as well as a potentially negative view of the self (Exner, 1993; Exner & Weiner, 1995), it is
reasonable to assume that response frequency would be higher among younger children, given their propensity to exhibit dependent and parentally-clingy symptoms of depression. Correspondingly, it is anticipated that T and TF responses (Sum T) would decrease around the time of adolescence. Based on Exner and Weiner’s (1995) normative data, this is not that case. Rather, the prevalence of Sum T responses increased slightly from a mean of .83 for five-year-olds to 1.02 for 16-year-olds, becoming progressively higher with age.

Overall, responses to the remaining CDI variables (i.e., COP, AG, Isolate/R, and active/passive response) remained predominantly stable across the five-to-16-year-old age range; however, as previously mentioned, clinically suggestive age-related changes in these variables were largely not expected for these non-clinical populations.

**Summary of the Literature**

In summary, challenges related to the categorical taxonomy, symptomatology, and assessment of depression have been discussed in conjunction with a number of prominent cognitive theories. An attempt was made to explore the relationship between these topics and the assessment of depression across childhood and adolescent development. A review of depressive symptom manifestation across children’s development into adolescence suggests that both cognitive and affective components play a role in overt depressive behavior; however, depressed children and adolescents also exhibit dynamic manifestations of social, externalizing, and somatic behavior that is often difficult to account for using either developmental and/or psychological principles, or standardized assessments.

Several standardized measures used to assess depression with children and adolescents have been discussed. The Rorschach Inkblot Test, administered and scored according to the standardized guidelines of the Exner Comprehensive System (1993), remains one controversial, yet promising, complement to commonly-used paper-and-pencil assessments. Several aspects of the CS were examined, however, that warrant further consideration and study. First, a surprising dearth of research exists pertaining to the CS depression indices and their integral variables, particularly among clinical populations of children and adolescents (Exner, 1993; 1997; Exner & Weiner, 1995;
Weiner, 2001). Considering this lack, it is not surprising that research has also failed to address developmental differences in CS response patterns exhibited by depressed children vs. depressed adolescents. Moreover, where few studies have examined the psychometric validity of the DEPI, virtually none have addressed the CDI. In particular, the crucial issues of incremental and discriminant validity, pertaining to both the DEPI and the CDI, as well as the Rorschach CS’s ability to assess depression overall, remains an issue in need of further study.

Research designed to replicate Exner’s (1993) normative sample, as well as the construct validity of the depression indicators, is also called for. Surprisingly, no known publications have provided correlation coefficients for CS variables to determine their degree of interrelationship (particularly the DEPI and CDI and their variables). Further research is also clearly warranted to evaluate controversial findings regarding the discriminant and convergent validity of the CS indices and variables. Specifically, the incremental validity of the CDI, in conjunction with the DEPI, to discriminate depressed from non-depressed children and adolescents warrants further investigation. Further research evaluating the constructs underlying CS indices also merits exploration, in light of factor analytic results suggesting that the DEPI and CDI may assess qualitatively distinct (perhaps developmentally sensitive) typologies of depression. Such results could be attempted by comparing the CS with more theoretically sensitive/specific external measures (i.e., measures of either cognitive and/or affective typologies of depression, measures of hopelessness ideation, coping skills deficits, etc.).

Weiner (2001) recommends that future research attempt to identify similarities between DSM-IV-based assessment protocols, rather than delineating differences between ‘psychopathic’ individuals and ‘normals.’ In this regard, ramifications of Ellis and Ashbrook’s (1989) resource allocation model towards assessing depression may also warrant further study. For example, investigation of the degree to which negative affect (and hypothesized, resulting irrelevant cognitions) affects cognitive processing could elucidate useful response patterns consistent with cognitively conceptualized typologies of depression. In other words, particular variables could hold diagnostic utility, and could be used potentially to bolster the discriminant validity of the current CS depression indicators. In particular, the relationship between categorically-validated depression and
the following variables is of keen interest due to their sensitivity to cognitive resource allocation: (a) u or – responses, (b) number of Zf responses, (c) number of W responses, (d) EA and EB ratio, and (e) Adj. D.

A paucity of research exists to evaluate the relationship between the Rorschach and more commonly-used (i.e., self-report) measures of depression. Such studies are advocated to address problematic issues such as the convergent and incremental validity of the CS, and could more thoroughly investigate CS construct validity in relation to other, prominent assessments. Recurrent estimates of the DEPI’s low discriminant validity also raise significant considerations and challenges regarding the manner in which this index should be interpreted and applied (Weiner, 1996; 1997; 2001). Although further studies are warranted to either corroborate or refute these findings, evaluation of appropriate ways of interpreting the DEPI and CDI have already been advocated. Meyer & Archer (2001) note, “it is now logical to reconceptualize the DEPI as a measure of implicit depressive propensities” (p. 489). In other words, less clinical emphasis on diagnosis-based criteria (i.e., DSM-IV) and more emphasis on convergence of CS scores with observable behavior (i.e., depressive symptoms) are called for. As such, studies should further investigate commonalities/differences between depressive symptoms typically exhibited throughout childhood and adolescent development, and as yet unavailable CS demographic data obtained from clinically depressed children and adolescents.

Finally, further research may be advocated regarding the appropriateness of Rorschach administration to young children. For example, further investigation is warranted to address Finch et al.’s (1990) observation that children tend to provide fewer responses than older individuals. Establishment of specific age parameters for the CS is recommended, based on empirical developmental evidence, and further evaluation of CS demographic data similar to the research of Wenar and Curtis (1991) is advocated to observe the influence of developmental variables on specific CS variables. To this end, Exner (1997) has noted that future research may support the creation of a new scale to better address the needs of these younger children.
CHAPTER 3

METHODOLOGY

Chapter 3 presents the methods and procedures utilized in the study. For the purposes of presentation the chapter has been divided into the following eight sections: statement of hypotheses, population, sample, research design, variables, instrumentation, procedures, and data analysis.

Statement of the Hypotheses

1. For the depressed psychiatric population of adolescents, there will be significant differences in both DEPI and CDI total scores compared with the population of non-depressed “normal” controls. The prediction is that significantly higher total scores on both the DEPI and the CDI will be observed for the population of depressed psychiatric adolescents than will be observed for the non-depressed “normal” control population.

2. For the depressed psychiatric population of adolescents, there will be significant differences in both DEPI and CDI total scores compared with the population of non-depressed psychiatric adolescents. The prediction is that significantly higher total scores on both the DEPI and the CDI will be observed for the population of depressed psychiatric adolescents than will be observed for the non-depressed psychiatric population.

3. For the non-depressed psychiatric population of adolescents, there will be no significant differences in DEPI total scores compared with the population of non-depressed “normal” controls.

4. There will be significant correlations between DEPI and CDI total scores, regardless of group membership (i.e., depressed-only, non-depressed clinical only.)

5. For both depressed and non-depressed adolescent psychiatric populations, there will be no significant mean differences in DEPI and CDI total scores attributable
to age. The prediction is that, for both psychiatric adolescent populations ages 13-17, there will be no significant main effect of age on DEPI or CDI total scores, respectively.

6. It is hypothesized that for the combined group of depressed and non-depressed psychiatric adolescents, the diagnostic efficiency of the DEPI will approximate the results derived by Ganellen (1996), based on Exner’s (1993) DEPI standardization sample of both depressed and non-depressed adults. It is further hypothesized that the contribution of the CDI, according to Exner’s (1993) recommended clinical score cutoff points, will improve DEPI diagnostic efficiency. The prediction is that the diagnostic efficiency of DEPI and CDI results produced by the study’s adolescent sample will be approximately similar to that produced by Exner’s adult sample with regard to sensitivity, specificity, positive and negative predictive power, false positive and negative rates, and overall correct classification rate, and that an optimal permutation of both DEPI and CDI scores will result in increased diagnostic efficiency/diagnostic discrimination.

7. Within the clinical adolescent group, there will be a significant incremental relationship between Rorschach DEPI and CDI total scores, MMPI-A Basic Scale 2 Scores, Reynold’s Adolescent Depression Scale Total Scores, and diagnoses of depression administered using the DSM-IV. The prediction is that, among the clinical population of adolescents, there will be a significant relationship between independent DEPI total scores, MMPI-A Scale 2 T-scores, and membership in the “depressed” group of adolescents. Further, a significant relationship is predicted between independent CDI index scores, MMPI-A Scale 2 scores, and categorical diagnoses of depression. Moreover, it is predicted that the greatest degree of significance will be identified pertaining to the relationship between DEPI and CDI index scores in combination, MMPI-A Scale 2 scores, and categorical diagnoses of depression.
Population

The population of interest consists of adolescents who were admitted within the past seven years to a small (approximately 15-bed) juvenile inpatient unit of a psychiatric hospital in the South East for short-term crisis stabilization. The average length of stay was approximately four days, during which medication administration/adjustment, brief therapeutic intervention, psychological testing, and discharge planning was accomplished. Subjects were either self-referred by their parents or legal guardians, or were referred by specific institutions (i.e., mental health or law enforcement agencies, schools, etc.) and/or community-based mental and physical health professionals, (i.e., social workers, counselors/therapists, physicians, etc.). Participants were referred from the broad tri-state geographical region surrounding the psychiatric hospital, which is affiliated with a prominent South Eastern medical hospital.

Sample

The clinical sample consists of a criterion group of roughly 340 adolescents, ranging in age from 13- to 17 years, who produced valid protocols on the psychometric tests of interest administered to them during their inpatient stay. All subjects included in the study were administered various Axis I disorders, according to DSM-IV criteria, upon admission to the psychiatric hospital. The most prominent Axis I disorders among the sample included Major Depressive Disorder, Bipolar I Disorder, Attention Deficit Hyperactivity Disorder, Adjustment Disorder, Conduct Disorder, and Oppositional Defiant Disorder. In order to rule out the confounding influence of intellectual deficits or developmental disabilities, all subjects identified as meeting DSM-IV criteria for Borderline Intellectual Functioning, Mental Retardation, or a Pervasive Developmental Disorder (i.e., Autistic Disorder, Asperger’s Disorder, etc.) were excluded from the study.

The most common reasons for referral of subjects for inpatient treatment included manifestation of specific symptoms of psychopathology, severe family/school/social maladjustment or trauma (i.e., physical and/or sexual abuse, domestic violence, interpersonal conflict and/or physical aggression/acting out), and delinquency. The proportion of males and females in the overall sample was approximately equivalent. The racial composition of the sample subjects was representative of the surrounding
geographical region, consisted predominantly of Caucasians (~79%) and African Americans (~17%), and comprised a small proportion of Hispanics (~3%) and other (~1%) minority groups. Although socioeconomic level varied somewhat, subjects represented predominantly middle- to low economic status.

All clinical data was obtained by review of archival, inpatient files. All clinical participants included within the study completed a psychometric test battery comprised of the Rorschach, the MMPI-A, and the RADS, and provided valid response sets on all instruments. For the purposes of the study, cases pertaining to the clinical data obtained were divided into two separate criterion-based groups, based on presenting symptoms/diagnoses administered at the time of each individual admission: depressed psychiatric adolescents, and non-depressed psychiatric adolescents. The depressed psychiatric group was comprised of roughly 160 inpatient adolescents (i.e., sample divided between ages 13-17) whose symptoms were consistent with categorical criteria for Major Depressive Disorder upon admission, and who were not diagnosed with other Axis I psychological disorders at the time of their inpatient visit.

Participants selected for membership within the non-depressed psychiatric group consisted of roughly 180 adolescents between the ages of 13 and 17. Non-depressed psychiatric subjects were selected on the basis of having been administered one or more externalizing or internalizing Axis I disorders, to the exclusion of a depressive disorder. Diagnosed internalizing disorders included in the study consisted exclusively of Post Traumatic Stress Disorder, Anxiety Disorder Not Otherwise Specified, Social Phobia, and Obsessive Compulsive Disorder (Note: Due to phenotypic similarities between Generalized Anxiety Disorder and depression, no subjects having received diagnoses of GAD will be included). Diagnosed externalizing disorders included in the study consisted of Conduct Disorder, Oppositional Defiant Disorder, Attention Deficit Hyperactivity Disorder, Impulse Control Disorder, Disruptive Behavior Disorder Not Otherwise Specified, Intermittent Explosive Disorder, and Adjustment Disorder with disturbance of Conduct. The sample consisted of roughly 30 adolescents diagnosed only with one or more internalizing disorders, roughly 100 adolescents diagnosed exclusively with one or more externalizing disorder, and roughly 50 adolescents diagnosed with more than one internalizing and externalizing comorbid disorders.
In addition to the two aforementioned groups, the first research question utilized a third control group of non-depressed/non-psychiatric “normal” adolescents. This sample was obtained from Exner and Weiner’s (1995) published study involving approximately 1400 non-patient adolescents (comprising part of the CS normative adolescent sample), and was comprised of frequency data pertaining to DEPI and CDI total scores. Specific demographic information pertaining to this sample was not provided by the authors, and the relationship between DEPI and CDI total scores and specific subjects was not specified.

**Research Design**

The study will include a series of 4 co-relational research design sets to investigate the relationships between DEPI and CDI total scores and diagnostic classification. Justification in utilizing a co-relational design is taken from Smith and Glass (1987), who advocate for the use of correlational analyses for two general purposes: 1.) to better understand the nature of specific constructs and their relationship to other constructs, and 2.) to be able to predict a certain variable on the basis of another.

**Variables**

This study will utilize Rorschach DEPI and CDI total scores as the primary variables to test all co-relational research hypotheses, based on the aforementioned literature review of Rorschach Comprehensive System (CS) research. In set 1 of the research analyses (i.e., research hypotheses 1 through 3), DEPI and CDI total scores will comprise the dependent variables with diagnostic group membership as the independent variable. In set 2 of the analyses (i.e., research hypotheses 4 and 5), DEPI and CDI total scores will comprise co-relational variables to test hypothesis 4, and in hypothesis 5 they will represent the dependent variables with age as the independent variable of interest.

Set 3 of the research analyses will address research hypothesis 6 by evaluating the diagnostic efficiency (i.e., sensitivity, specificity, predicative power, rate of classification, and Kappa) of the DEPI and CDI using total scores as independent variables, with diagnostic classification serving as the dependent variable of interest. In set 4 of the analyses (to test hypothesis 7) DEPI and CDI total scores will comprise the independent variables, along with MMPI-A Scale 2 T-scores and RADS total scores, with predicted
group membership comprising the dichotomous dependent variable of interest (i.e.,
depressed vs. non-depressed). As recommended by Exner (1993), CS total scores will be
considered in continuous terms ranging from 0 to 7 for the DEPI, and 0 to 5 for the CDI.
While not evaluated independently, numerous coded variables comprise the CS DEPI and
CDI indices, and bear brief explanation.

**DEPI Components**

1. $V$ (Vista, or responses in which the subject interprets inkblot shading features as depth
   or dimensionality).
2. $C.Sh$ (Color-Shading blends, or responses involving both chromatic color and
   shading).
3. $3r(2)/R$ (Egocentricity Index, or proportion of responses involving reflections,
   relative to overall responses).
4. $C'$ (Responses involving achromatic color).
5. $MOR$ (Morbid content, or responses in which percepts are identified as dead,
   destroyed, ruined, spoiled, damaged, injured, or broken).
6. $SumSh$ (Sum of all responses involving specified elements of shading).
7. $FM$ (responses involving the movement of animals).
8. $m$ (responses involving the movement of inanimate objects).
9. $COP$ (Cooperative, or responses involving cooperative activity among percepts).
10. Sum of $AB$, $Art$, and $Ay$ (Sum of responses involving abstract content, and percepts
    involving art and anthropology).
11. $Isolate/R$ (Isolation Index, or ratio of overall responses to those containing the
    following types of percepts: Botany, Clouds, Geography, Landscape, and Nature).

**CDI Components**

1. $EA$ (Experience Actual, or the sum of Human Movement responses and a weighted
   calculation of the sum of all chromatic color responses (FC, CF, and C).
2. Adj. D (adjusted calculation involving the sum of D location responses).
3. $COP$ (Cooperative Movement, or movement responses in which two objects are
   reported to be engaged in positive or cooperative interaction).
4. *AG* (Aggressive Movement, or movement responses in which an aggressive action is reported).
5. *Afr* (Affective Ratio, or proportion of chromatic color responses given in response to chromatic color Rorschach inkblots, relative to overall responses on all inkblots).
6. *p* (passive, or movement response in which percepts are engaged in a passive activity).
7. *a* (active, or movement response in which percepts are engaged in an active activity).
8. *H* (Human, or response in which a human percept is identified).
9. *T* (Texture, or responses in which inkblot shading is specified as a tactual phenomenon).
10. *Isolate/R* (see above)
11. *Fd* (Food, or responses containing food as a percept).

**MMPI-A Scale 2 (Depression scale)**

This study will utilize T-scores obtained by subjects on the MMPI-A scales as continuous independent variables. Scale 2 (i.e., the Depression, or D scale) of the MMPI-A Basic Scales will be utilized to address research hypothesis six for the purpose of evaluating the incremental validity of the Rorschach and MMPI-A.

**The Reynolds Adolescent Depression Scale (RADS)**

For research question number four, total scores obtained on the Reynolds Adolescent Depression Scale (RADS) will be inserted into the hierarchical discriminant function analysis equation as continuous independent variables in order to evaluate the incremental validity of this instrument, as well as that of the Rorschach DEPI and CDI.

**Instrumentation**

**Rorschach Inkblot Test**

Results of the Rorschach were obtained from administering and scoring the test according to nomothetic Comprehensive System (CS; Exner, 1993) guidelines and procedures. With respect to the Rorschach, the focus of the study was constrained to the Depression Index (DEPI) and the Coping Deficit Index (COP), both of which were designed to measure depression, and evaluated the specific variable derivations
comprising these indices. For a detailed review of DEPI and CDI psychometric properties, please refer to Chapter 2 of this study.

Of note, precautions were undertaken to ensure that only valid Rorschach protocols were utilized in the study. As several studies have advocated for caution when interpreting Rorschach “brief” protocols containing fewer than 14 responses (Exner, 1993; Exner & Weiner, 1995; Finch, Imm, & Belter, 1990; Leichtman, 1996b), no such protocols were used in the study. Further, as the presence of low proportions of Lambda responses (i.e., <1) for individuals of average intelligence has been hypothesized to be associated with defensiveness on the Rorschach (Weiner, 1996), protocols with Lambda ratios of 1 or lower were discarded.

The Minnesota Multiphasic Personality Inventory – Adolescent (MMPI-A)

The MMPI-A (Butcher, Williams, Graham, Archer, Tellegen, Ben-Porath, and Kaemmer, 1992) is a 478-item self-report checklist designed to measure broad psychopathology and personality characteristics. The MMPI-A has been extensively validated in studies with other self-report inventories, and with categorical diagnoses of psychopathology (including depression)(Archer & Krishnamurthy, 1993). It is currently one of the most commonly administered instruments used by clinicians to assess adolescent personality and psychopathology, and demonstrates excellent convergent and discriminant validity across instruments, methods, and observers (please see Archer, 1997 and Archer & Krishnamurthy, 1993 for a review).

Precautions were undertaken to ensure MMPI-A validity for all subjects, based on the numerous validity indicators included within this instrument. These indicators include T-score elevations of 65 or higher on the VRIN (i.e., Variable Response Inconsistency), TRIN (True Response Inconsistency), L and K, scales, F1, and F2 (i.e., “back F,” or latter portion of items) T-scores of >110, and raw ‘?’ (i.e., “Cannot Say” scale, or items to which subjects did not respond) scores of >30. Pertaining to the “cannot say” scale, no subjects exhibited raw scores in excess of 30, as examination protocol consisted of re-administering the MMPI-A to these subjects and/or determining the rationale for missing responses (i.e., limited reading ability, oppositional behavior,
etc.). Individuals exhibiting clinically significant T-scores in reaching or exceeding 65 on these validity indicators were excluded from the study.

The Reynolds Adolescent Depression Scale (RADS)

The RADS is a self-report checklist designed for adolescents between the ages of 12 and 18 that is comprised of 30 items on which responses are given using a four-point likert-type scale (Reynolds, 1987). The RADS is a popular and widely used instrument that has been found to demonstrate acceptable psychometric validity and reliability (for a detailed description of its psychometric properties, please refer to page 56).

Procedures

Data collection for this study took place in Fall, 2002 at the Behavioral Health Center affiliated with Tallahassee Regional Memorial Hospital. Prior to data collection, Institutional Review Board approval was obtained for all procedures from Tallahassee Regional Memorial Hospital’s Medical Staff Office, as well as from the Florida State University’s Human Subjects Committee. Available archival data located within the medical records of child and adolescent inpatients admitted to the Behavioral Health Center within the past seven years was collected. This data consisted of demographic information pertinent to each subject (i.e., age, race, gender, socioeconomic status, marital status of parents, and previous/current attempts to commit suicide), categorical disorders diagnosed at the time of each subject’s inpatient admission to the Behavioral Health Center, and psychological test results germane to the interests of this study (i.e., data from the Rorschach Comprehensive System pertinent to the DEPI and the CDI, MMPI-A Scale 2 T-scores, and RADS total scores). Exemplary standards of confidentiality were maintained by the study, and it was ensured that no personal, direct contact or professional interaction occurred between the examiner and/or individuals assisting in data collection and selected subjects associated with the archival data.

The process of admission, diagnosis, assessment, and discharge of patients referred to the Behavioral Health Center occurred according to a consistent, predictable schedule. Categorical Axis I diagnoses are administered during triage by a psychiatrist via a structured clinical interview, and in accordance with DSM-IV criteria. This occurs immediately prior to hospital admission, and before the patient is administered any
standardized psychological assessments. Additionally, mental health professionals obtain comprehensive background and demographic information upon intake, based on structured interviews conducted with the patient and presenting family members.

Psychology interns employed by the Behavioral Health Center administered the Rorschach, typically within the first two days following a patient’s admission to the hospital. Psychology interns consisted of upper-level graduate psychology students who have completed at least two years in either a doctoral (Ph.D. or Psy.D.) program of study. Although the extent of intern training in the administration and scoring of the Rorschach Comprehensive System varied, interns typically possessed some extent of formal instruction within an academic setting, and some clinical experience with the instrument. Psychology interns transcribed verbatim all Rorschach responses provided by children and adolescents.

Scoring was accomplished exclusively using Exner’s (1993) nomothetic Comprehensive System guidelines, and the structural summary was calculated using the RIAP scoring system. Scoring accuracy was verified by a clinical supervisor possessing greater than 25 hours of approved training and more than seven years of clinical experience, according to the standards advocated by Guarnaccia, Dill, Sabatino, and Southwick (2001). Prior to discharge from the hospital, inpatient adolescents completed the MMPI-A independently via paper-and-pencil format under the supervision of a direct care unit staff member, psychologist, or psychology intern. Responses were subsequently scored by psychology interns using a computerized scoring method, and were plotted using T-score conversions for the Basic scales.

Data Analysis

All statistical analyses will be computed using the Statistical Package for the Social Sciences (SPSS). The archival data collected will represent a sample of convenience likely consisting of non-equivalent sample sizes that may delimit the analyses to be performed; however, statistical procedures will be conducted based on parameters suggested by Tabachnick and Fidell (1996). In order to ensure adequate power to investigate medium-sized relationships, the authors advocate for alpha level of .05 and effect size equal to .20. As such, a sufficient sample size is determined by the
following estimates: \( N > 50 + 8m \) (\( m \) = number of independent variables) for tests of overall correlation, and \( N > 104 + m \) for testing individual predictors. Based on these guidelines, it does not appear that sufficient \( N \) will be problematic. To test the 6 research hypotheses, four series of analyses will be performed.

**Analyses Set 1**

The first set of analyses will address hypotheses 1 through 3, and will consist of a series of t-tests to compare both DEPI and CDI total scores provided by the three subject groups of interest: non-psychiatric/“normal” control adolescents, depressed-only psychiatric adolescents, and non-depressed psychiatric adolescents. Analysis of Variance (ANOVA) or other more optimal statistical methods of evaluating for reliable differences between two or more means (Tabachnick and Fidell, 1996) will not be undertaken in analysis 1 due to limitations in the non-psychiatric treatment group data provided by Exner and Weiner’s (1995) Comprehensive System normative study.

Specifically, it is problematic that no relationship was specified pertaining to the normal subjects’ DEPI and CDI scores, which were presented as frequency data only limited to total \( N \), and number of individuals obtaining scores equal to or greater than Exner’s (1993) recommended clinical score cutoff points (i.e., DEPI \( \geq 5, 6, \) and 7, and CDI \( \geq 4 \) and 5). As such, analysis of the data for all groups will be performed using independent samples, one-tailed t-tests for ages 13-16 (corresponding to the “normal” age groupings available). Scores less than 5 will be coded as scores of 4 to avoid over-inflation of mean group differences and to provide the strictest criteria possible with which to measure these differences. T-scores, means, standard deviations, and level of significance for each test will be reported.

One-tailed t-tests will be used, as recommended by Witte (1989), due to the unidirectional nature of the research hypothesis (i.e., it is predicted that DEPI and CDI Total scores will be higher among the depressed subject group than in the non-depressed and “normal” control groups). Frequency data pertaining to the “normal” group of non-psychiatric adolescents will be coded and entered into an SPSS database, and no post hoc analysis of significant differences in psychiatric vs. non-psychiatric variables will be undertaken due to the aforementioned limitations in the data. Further, a Bonferroni
correction will be utilized to control for family-wise error, as recommended by Tabachnick and Fidell (1996). Due to the anticipated uneven sample size (N) when comparing clinical and “normal” groups, Levene’s Test for Equality of Variances will also be performed for each t-test.

**Analyses Set 2**

The second set of analyses will address research hypotheses 4 and 5 to evaluate the overall relationship between the DEPI and CDI, and to investigate the main effect of chronological age (i.e., 13, 14...17) on mean DEPI and CDI total scores. To test hypothesis 4, which states that there will be significant relationships between DEPI and CDI total scores for all psychiatric groups of adolescents, an omnibus Pearson product moment correlation will be performed using the combined group of both depressed and non-depressed psychiatric adolescents of all ages (i.e., 13-17 years old). Based on Tabachnick and Fidell’s (1996) specifications for adequate power (i.e., N > 50 + 8m for tests of overall correlation), it is not anticipated that the estimated sample size of roughly 340 individuals will be problematic.

Hypothesis 5 posits that no significant differences in either DEPI or CDI total scores will be attributable to chronological age. To test this hypothesis, two Analysis of Variance (ANOVA) techniques will be performed using the dependent variables of DEPI and CDI total scores, respectively, and chronological age as the independent variable in both analyses. The main effect of chronological age on both DEPI and CDI total scores, significance, and means and standard deviations will be reported. In the event that the main effect of age on either DEPI or CDI total scores is found to be significant, post-hoc pair-wise analyses will be conducted between each chronological age group (i.e., 13, 14...17) and CDI and/or DEPI total score using Pearson correlations.

**Analyses Set 3**

The third set of analyses will test research hypothesis 6 by evaluating the diagnostic efficiency of the DEPI and CDI, both independently and jointly. In determining a test’s diagnostic efficiency, Ganellen (1996) posits that test scores are most useful if they maximize the examiner’s ability to accurately identify test subjects into diagnostic categories, while minimizing the probability of misclassification. Ganellen,
and also Streiner (2003), have advocated for the use of the following descriptive and statistical computations in evaluating a test’s diagnostic efficiency: sensitivity, specificity, false positive rate, false negative rate, positive predictive power, negative predictive power, overall correct classification, and kappa. For general definitions of the aforementioned terms, please refer to the Definitions of Terms section (p. 16) of this study.

Operational definitions of the aforementioned terms within this study are as follows: (a) sensitivity refers to the proportion of adolescents diagnosed with categorical depression (per DSM-IV criteria) who were detected as depressed by the DEPI and CDI regardless of the overall prevalence of depressive disorders within the population of the analysis; (b) specificity refers to the proportion of adolescents who were not diagnosed with depression (per DSM-IV criteria) and who, correctly, were not identified as depressed by the DEPI and CDI, regardless of the overall prevalence of non-depressive disorders within the population of the analysis; (c) false positive rate refers to the proportion of adolescents not diagnosed with categorical depression who were incorrectly identified by the DEPI and CDI as depressed; (d) false negative rate refers to the proportion of adolescents diagnosed with categorical depression who were incorrectly not identified by the DEPI and CDI as depressed; (e) positive predictive power refers to the percentage of adolescents identified by the DEPI and CDI as being depressed, relative to the overall number of adolescents actually diagnosed with depression in the sample; (f) negative predictive power refers to the percentage of adolescents identified by the DEPI and CDI as being non-depressed, relative to the overall number of adolescents who were actually not diagnosed with depression in the sample; (g) overall classification rate refers to the rate at which the DEPI and CDI were able to correctly identify all of the adolescents as being both depressed and non-depressed; and (h) Kappa refers to the statistically represented level of agreement between DEPI and CDI total scores and adolescents’ administered diagnoses beyond that accounted for by chance.

The aforementioned diagnostic efficiency statistics will be computed based on various sets of total score cutoff points for both the DEPI, the CDI, and the DEPI and CDI in combination. Exner (1993) has recommended that total score cutoff points of 5 and between 6 and 7 be considered for the DEPI with respect to improving diagnostic
classification of depression. Similarly, Exner recommends that CDI scores of 4 and 5 be considered as being increasingly indicative of depressive symptoms. As such, diagnostic efficiency statistics will be computed for the DEPI and CDI independently, and also both indices in combination, to evaluate the accuracy/efficiency of the indices at various total score cutoff points towards prediction/classification of depressed adolescents by diagnostic category.

**Analyses Set 4**

In the fourth set of analyses, a hierarchical logistic regression analysis will be performed to test the seventh and final research hypothesis. Analyses set 4 will assess the incremental validity of the DEPI and the CDI by evaluating the soundness of a series of logistic regression models through inspection of the following: (a) overall evaluation of various model combinations of the variables DEPI, CDI, MMPI-A, and RADS; (b) statistical tests of the individual predictors; and (c) evaluation of the goodness of fit of each model, as contributes to the outcome variable of diagnosis. The regression equation will consist of the dichotomous dependent variable of diagnostic classification (i.e., “depressed only” and “non-depressed other only”) pertaining to the criterion group of inpatient psychiatric adolescent participants, and will consist of approximately 250 individuals. The independent predictor covariates to be included in the study will consist of DEPI and CDI total scores, RADS total scores, and MMPI-A Scale 2 T-scores provided by both depressed and non-depressed adolescents.
CHAPTER 4

RESULTS

Introduction

A series of 5 sets of analyses was undertaken to test the research hypotheses. The first set of analyses (research hypotheses 1 through 3) was conducted to determine whether significant differences in both DEPI and CDI total scores (separately) would be identified at different age ranges (13-16) using t-score comparisons involving three separate adolescent groups: non-depressed “normal” controls, depressed psychiatric adolescents, and non-depressed psychiatric adolescents. The second research set was performed to test research hypotheses 4 and 5. The test of hypothesis 4 investigated the overall relationship between DEPI and CDI total scores irrespective of age, using a correlational analysis. The test of hypothesis 5 used two separate ANOVA analyses to investigate the influence of chronological age on both DEPI and CDI total scores.

The third set of analyses was conducted to test the diagnostic efficiency of the DEPI and CDI scales as indicators of depression using the following statistics: sensitivity, specificity, positive and negative predictive power, false positive and negative rate, overall correct classification, and Kappa values. DEPI and CDI total scores were evaluated using these statistics, both independently and in combination, based on Exner’s (1993) recommended total score cutoff points. The fourth and final set of analyses was conducted to evaluate the incremental validity of the DEPI, CDI, MMPI-A Scale 2 scores, and RADS total scores using a hierarchical linear regression analysis. The following chapter presents the results of this study including: (1) demographic characteristics of the sample, (2) means, standard deviations, and correlations of all variables of interest, and (3) data analysis results for each set of research hypotheses.

Sample Demographic Data

The clinical sample used in this study consisted of 335 adolescent subjects ranging in age from 13-17 years. All subjects were admitted on an inpatient basis to
Tallahassee Regional Memorial Hospital’s Behavioral Health Unit in Tallahassee, Florida. Demographic data for the sample are presented in Table 1. The total sample consisted of 179 females (53.4%) and 156 males (46.6%). The ethnic composition of the sample was comprised primarily of Caucasians (78.8%), followed by African Americans (16.7%), Hispanic/Latino Americans (3.3%), and a small proportion of individuals of other ethnic backgrounds (i.e., Asian American/Pacific Islander, Native American) (1.2%). The participants represented 53 13-year-olds (15.8%), 68 14-year-olds (20.3%), 83 15-year-olds (24.8%), 86 16-year-olds (25.6%), and 45 17-year-olds (13.4%).

Hospital records indicated that 29.3% of all clinical participants came from households with intact parental relationships, 40.6% from those with divorced and/or separated parents, 11.3% from single-parent households, and 18.8% from other parental arrangements (i.e., foster care/temporary custody arrangements, or raised by a relative).

The sample was made up of 158 individuals diagnosed exclusively with depression (66.5% female, 33.5% male), and 177 individuals diagnosed with one or more internalizing disorders (15.25% of non-depressed psychiatric sample), one or more externalizing disorders (58.76% of non-depressed psychiatric sample), or a combination of one or more internalizing and externalizing disorders (25.99% of non-depressed psychiatric sample). Distribution of subjects within these two groups by age is provided in Table 2. Within the non-depressed psychiatric sample, 77.8% of those diagnosed only with non-depressive internalizing disorders were female, and 22.2% were male. Within the externalizing-only group, 26.9% were female, 73.1% were male. Finally, within the comorbid internalizing and externalizing group, 54.3% were female, and 45.7% were male. Distribution of subjects according to age and administered diagnostic category is provided in Table 3.

Exner and Weiner’s (1995) Rorschach Comprehensive System standardization sample was used as a non-clinical sample in analysis set 1. Exner and Weiner report that this sample was comprised of adolescents who were never admitted to an inpatient psychiatric facility, and who denied the presence of any significant psychopathology. The authors did not specify information pertaining to the sex, ethnic background, and geographic area of origin of the subjects within their non-clinical sample. Further, no relationship was specified between the subjects’ DEPI and CDI total scores. The
“normal” group of adolescents was included in the first research set only in order to provide non-clinical representation, and represented a sample of convenience. The group consisted of 110 (23.66%) 13-year-olds, 105 (22.58%) 14-year-olds, 110 (23.66%) 15-year-olds, and 140 (30.11%) 16-year-olds.

<table>
<thead>
<tr>
<th>Table 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographic Characteristics of Clinical Sample</td>
</tr>
<tr>
<td>(N=335)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th># of Subjects</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Female</td>
<td>179</td>
<td>53.4%</td>
</tr>
<tr>
<td>Total Male</td>
<td>156</td>
<td>46.6%</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Combined, age 13</td>
<td>53</td>
<td>15.8%</td>
</tr>
<tr>
<td>Combined, age 14</td>
<td>68</td>
<td>20.3%</td>
</tr>
<tr>
<td>Combined, age 15</td>
<td>83</td>
<td>24.8%</td>
</tr>
<tr>
<td>Combined, age 16</td>
<td>86</td>
<td>25.7%</td>
</tr>
<tr>
<td>Combined, age 17</td>
<td>45</td>
<td>13.4%</td>
</tr>
<tr>
<td><strong>Ethnicity</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>264</td>
<td>78.8%</td>
</tr>
<tr>
<td>African American</td>
<td>56</td>
<td>16.7%</td>
</tr>
<tr>
<td>Hispanic, Latin American</td>
<td>11</td>
<td>3.3%</td>
</tr>
<tr>
<td>Other</td>
<td>4</td>
<td>1.2%</td>
</tr>
<tr>
<td><strong>Parental Structure</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intact Parental Relationship</td>
<td>98</td>
<td>29.3%</td>
</tr>
<tr>
<td>Divorced Parents</td>
<td>136</td>
<td>40.6%</td>
</tr>
<tr>
<td>Single Parent</td>
<td>38</td>
<td>11.3%</td>
</tr>
<tr>
<td>Other</td>
<td>63</td>
<td>18.8%</td>
</tr>
<tr>
<td><strong>DSM-IV Diagnoses</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depressed Only</td>
<td>158</td>
<td>47.2%</td>
</tr>
<tr>
<td>Non-Depressed</td>
<td>27</td>
<td>8.1%</td>
</tr>
<tr>
<td>Internalizing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Depressed</td>
<td>104</td>
<td>31%</td>
</tr>
<tr>
<td>Externalizing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Depressed Combined</td>
<td>46</td>
<td>13.7%</td>
</tr>
</tbody>
</table>
### Table 2
**Total Sample By Diagnosis and Age**
(N=800)

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Depressed-Only Psychiatric Group</th>
<th>Non-Depressed Psychiatric Group</th>
<th>Non-psychiatric “Normal” Control Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>13 year-olds</td>
<td>13</td>
<td>40</td>
<td>110</td>
</tr>
<tr>
<td>14-year-old</td>
<td>27</td>
<td>41</td>
<td>105</td>
</tr>
<tr>
<td>15-year-olds</td>
<td>40</td>
<td>43</td>
<td>110</td>
</tr>
<tr>
<td>16-year-olds</td>
<td>49</td>
<td>37</td>
<td>140</td>
</tr>
<tr>
<td>17-year-olds</td>
<td>29</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td><strong>Total:</strong></td>
<td><strong>158</strong></td>
<td><strong>177</strong></td>
<td><strong>465</strong></td>
</tr>
</tbody>
</table>

### Table 3
**Non-Depressed Psychiatric Sample by Age and Diagnosis**
(N=167)

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Internalizing/Non-Depressed-Only Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>13 year-olds</td>
<td>2</td>
</tr>
<tr>
<td>14-year-olds</td>
<td>6</td>
</tr>
<tr>
<td>15-year-olds</td>
<td>10</td>
</tr>
<tr>
<td>16-year-olds</td>
<td>4</td>
</tr>
<tr>
<td>17-year-olds</td>
<td>5</td>
</tr>
<tr>
<td><strong>Total:</strong></td>
<td><strong>27</strong></td>
</tr>
</tbody>
</table>

91
Table 3 (Continued)

<table>
<thead>
<tr>
<th>Externalizing/Non-Depressed-Only Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>13 year-olds</td>
</tr>
<tr>
<td>14-year-olds</td>
</tr>
<tr>
<td>15-year-olds</td>
</tr>
<tr>
<td>16-year-olds</td>
</tr>
<tr>
<td>17-year-olds</td>
</tr>
<tr>
<td><strong>Total:</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Internalizing and Externalizing/Non-Depressed-Only Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>13 year-olds</td>
</tr>
<tr>
<td>14-year-olds</td>
</tr>
<tr>
<td>15-year-olds</td>
</tr>
<tr>
<td>16-year-olds</td>
</tr>
<tr>
<td>17-year-olds</td>
</tr>
<tr>
<td><strong>Total:</strong></td>
</tr>
</tbody>
</table>

Results of Data Analyses Set 1

To address hypotheses 1 through 3, the relationship between DEPI and CDI total scores was evaluated using a series of independent-samples, one-tailed t-tests. Separate pair-wise comparisons were undertaken, by age (i.e., 13, 14, 15, 16), for both DEPI and CDI total scores between the following groups: non-depressed/“normal” control adolescents (N=465), depressed psychiatric adolescents (N=129), and non-depressed psychiatric adolescents (N=161). The 17-year-olds were excluded from the clinical group of adolescents in research question 1 because no same-age non-clinical comparison group was available. A Bonferroni correction was utilized to control for family-wise error, and $p$ was set at .004 (Tabachnick & Fidell, 1996). Because of greater sample sizes in the non-clinical/“normal” group and disproportionate between-group variance, Levene’s Test for Equality of Variances was performed for each t-test. For all subject groupings evaluated, significant inequality of variance was identified and correspondingly adjusted.

Hypotheses 1 Through 3
H1a: For the depressed psychiatric population of adolescents, there will be significant differences in DEPI total scores, compared with the population of non-depressed “normal” controls.

H2a: For the depressed psychiatric population of adolescents, there will be significant differences in DEPI total scores, compared with the population of non-depressed psychiatric adolescents.

H3: For the non-depressed psychiatric population of adolescents, there will be no significant differences in DEPI total scores compared with the population of non-depressed “normal” controls.

T-test results and descriptive data corresponding to DEPI comparisons are provided in Table 4. The data were consistent in supporting a fail to reject decision for H3, given findings of no significant differences between depressed-only clinical and non-depressed clinical group DEPI scores. Considered independently, the 15-year-old group of depressed-only psychiatric adolescents (M = 4.28, SD = .62) demonstrated higher DEPI total scores than the “normal” adolescent control group at the $p < .01$ level (M = 4, SD = 0), $t(35) = 2.712$, $p = .01$ (one-tailed), $d = .1024$; however, given the Bonferroni correction necessitating significance at the $p < .004$ value, results supported the decision to fail to reject the null hypothesis for all ages (i.e., including 15-year-olds) for both H1a and H2a. Namely, no significant differences in DEPI total scores were demonstrated by depressed-only psychiatric adolescents, compared with “normal” non-psychiatric or non-depressed psychiatric adolescents.

The non-depressed psychiatric 15-year-olds (M = 4.20, SD = .51) were found to demonstrate higher DEPI scores than the “normal” control group at the $p < .05$ level (M = 4, SD = 0), $t(40) = 2.446$, $p = .019$ (one-tailed), $d = .08$, which was also found to be non-significant in light of correction for family-wise error inflation (i.e., p<.004). The homogenous distribution of the DEPI “normals” group for all ages except 13-year-olds resulted due to the need to code all sub-clinical scores (i.e., DEPI ≤ 5) as DEPI = 4, given that Exner (1993) did not specify exact DEPI total scores for subjects not exceeding DEPI >4. For the entire DEPI non-clinical group, only one subject (one 13-year-old) obtained a DEPI total score of greater than 4.
Table 4

T-Scores, Means, and Standard Deviations For DEPI Pair-Wise Comparisons

<table>
<thead>
<tr>
<th>Age</th>
<th>NC M</th>
<th>SD</th>
<th>DP M</th>
<th>SD</th>
<th>NP M</th>
<th>SD</th>
<th>NC by DP</th>
<th>NC by NP</th>
<th>DP by NP</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>13</td>
<td>4.01</td>
<td>0.095</td>
<td>4.15</td>
<td>0.83</td>
<td>4.08</td>
<td>0.28</td>
<td>1.385</td>
<td>1.56</td>
</tr>
<tr>
<td>2.</td>
<td>14</td>
<td>4.00</td>
<td>0.49</td>
<td>4.19</td>
<td>0.49</td>
<td>4.09</td>
<td>0.29</td>
<td>1.995</td>
<td>1.787</td>
</tr>
<tr>
<td>3.</td>
<td>15</td>
<td>4.00</td>
<td>0.62</td>
<td>4.14</td>
<td>0.47</td>
<td>4.03</td>
<td>0.18</td>
<td>2.712</td>
<td>2.446</td>
</tr>
<tr>
<td>4.</td>
<td>16</td>
<td>4.00</td>
<td>0.47</td>
<td>4.14</td>
<td>0.47</td>
<td>4.03</td>
<td>0.18</td>
<td>1.961</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Note: NC = Normal, Non-Psychiatric Adolescent Control; DP = Depressed Psychiatric Adolescents; NP = Non-Depressed Psychiatric Adolescents; *p < .004

H1b: For the depressed psychiatric population of adolescents, there will be significant differences in CDI total scores, compared with the population of non-depressed “normal” controls.

H2b: For the depressed psychiatric population of adolescents, there will be significant differences in CDI total scores, compared with the population of non-depressed psychiatric adolescents.

T-test results and descriptive data corresponding to CDI comparisons are provided in Table 5. Pertaining to H1b, initial differences in CDI scores were noted at p < .05 for all ages when comparing the “normal” non-psychiatric control group with the depressed psychiatric group of adolescents; however, only the 14- and 15-year-old groups demonstrated significant differences based on the Bonferroni-corrected p level of <.004. Pertaining to H2b, the 16-year-old group of non-depressed psychiatric group was found to provide higher CDI total scores (M = 3.73, SD = .78), t(47.14) = 2.429, p = .019 (one-tailed), d = .16 than the depressed group (M = 3.33, SD = .53); however, the decision to support the failure to reject the null hypothesis for H2b was supported for all ages given corrected p<.004. Although not hypothesized to occur, the non-depressed psychiatric sample demonstrated significantly higher CDI total scores than the “normal” control group (p ≤.004), regardless of age.
Table 5

T-Scores, Means, and Standard Deviations For CDI Pair-Wise Comparisons

<table>
<thead>
<tr>
<th>Age</th>
<th>NC M SD</th>
<th>DP M SD</th>
<th>NP M SD</th>
<th>NC by DP</th>
<th>NC by NP</th>
<th>DP by NP</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. 13</td>
<td>3.13 0.33</td>
<td>3.85 0.90</td>
<td>3.64 0.64</td>
<td>2.861</td>
<td>4.60**</td>
<td>0.765</td>
</tr>
<tr>
<td>2. 14</td>
<td>3.12 0.33</td>
<td>3.65 0.80</td>
<td>3.52 0.71</td>
<td>3.32**</td>
<td>3.054*</td>
<td>0.695</td>
</tr>
<tr>
<td>3. 15</td>
<td>3.12 0.35</td>
<td>3.49 0.66</td>
<td>3.71 0.75</td>
<td>3.162**</td>
<td>4.837**</td>
<td>1.372</td>
</tr>
<tr>
<td>4. 16</td>
<td>3.10 0.32</td>
<td>3.33 0.53</td>
<td>3.73 0.78</td>
<td>2.725</td>
<td>4.341**</td>
<td>2.429</td>
</tr>
</tbody>
</table>

Note: NC = Normal, Non-Psychiatric Controls; DP = Depressed Psychiatric Adolescents; NP = Non-Depressed Psychiatric Adolescent; *\( p < .004 \); **\( p < .004 \)

Results of Data Analyses Set 2

To evaluate research hypothesis 4, the relationship between DEPI and CDI total scores was evaluated using an omnibus, bivariate Pearson product moment correlation to compare DEPI and CDI total scores from both the combined group of depressed-only and non-depressed psychiatric adolescents of all ages (i.e., 13, 14...17). Post-hoc correlational analyses were also conducted using two additional Pearson product moment correlations to compare DEPI and CDI total scores for each respective group (i.e., depressed-only adolescents of all ages, and non-depressed psychiatric adolescents of all ages) to observe preliminary differences in total scores potentially attributable to group membership. To test hypothesis 5, the main effect of chronological age (i.e., 13, 14,...17) on DEPI and CDI total scores was evaluated using two separate one-way analyses of variance (ANOVA’s). Two separate analyses were conducted with both the DEPI and the CDI as dependent variables, and subject age as the independent variable.

Hypothesis 4

H4: There will be significant correlations between DEPI and CDI total scores, regardless of group membership (i.e., depressed-only, non-depressed clinical only.)
**Pearson product moment results**

Pearson coefficients, significance, means, and standard deviations derived from the variables of interest are provided in Table 6. In evaluating correlational relationships, Cohen (1977) recommends the following conventional guideline for defining size of a correlational relationship: small=.10, medium=.30, and large=.50. For both the overall group of combined adolescents ($r = .144, p = .013$) and also the depressed-only psychiatric adolescents ($r = .229, p = .006$), significant, small-to-medium relationships were identified between DEPI and CDI total scores. While not specifically hypothesized, a non-significant, extremely small relationship was found between DEPI and CDI scores for the non-depressed psychiatric group of adolescents ($r = .055, p = .492$).

<table>
<thead>
<tr>
<th>Group</th>
<th>DEPI M</th>
<th>DEPI SD</th>
<th>CDI M</th>
<th>CDI SD</th>
<th>Pearson Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. ALL</td>
<td>3.322</td>
<td>1.104</td>
<td>3.243</td>
<td>1.124</td>
<td>.144 *</td>
</tr>
<tr>
<td>2. DP</td>
<td>3.366</td>
<td>1.201</td>
<td>3.207</td>
<td>1.111</td>
<td>.229 **</td>
</tr>
</tbody>
</table>

Note: ALL = Combined depressed-only and non-depressed psychiatric; DP = Depressed-Only Psychiatric Adolescents; NP = Non-Depressed Psychiatric Adolescent; * p < .05; **p < .01

**Hypothesis 5**

H5: For both depressed and non-depressed adolescent psychiatric populations, there will be no significant mean differences in DEPI and CDI total scores attributable to age. The prediction is that, for both psychiatric adolescent populations ages 13-17, there will be no significant main effect of age on DEPI or CDI total scores, respectively.
Results of analyses of variance (ANOVA’s)

Results of the analyses consistently call for acceptance of H5, and rejection of the null hypothesis. As predicted, for DEPI total scores the omnibus test of the main effect of age on total score was not significant $F(4, 300) = .569, p = .686$. In a consistent manner, the main effect of age on CDI total scores, $F(4, 300) = .698, p = .594$, was not significant. As a result, no post-hoc comparison analyses were undertaken to evaluate the effects of specific age groups on total scores.

<table>
<thead>
<tr>
<th>Age</th>
<th>DEPI M</th>
<th>DEPI SD</th>
<th>CDI M</th>
<th>CDI SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. 13</td>
<td>3.62</td>
<td>.77</td>
<td>3.77</td>
<td>1.01</td>
</tr>
<tr>
<td>2. 14</td>
<td>3.42</td>
<td>1.10</td>
<td>3.46</td>
<td>1.07</td>
</tr>
<tr>
<td>3. 15</td>
<td>3.22</td>
<td>1.35</td>
<td>3.17</td>
<td>1.08</td>
</tr>
<tr>
<td>4. 16</td>
<td>3.21</td>
<td>1.18</td>
<td>2.90</td>
<td>.96</td>
</tr>
<tr>
<td>5. 17</td>
<td>3.61</td>
<td>1.29</td>
<td>3.21</td>
<td>1.11</td>
</tr>
</tbody>
</table>

Results of Data Analyses Set 3

The analyses in set 3 were conducted to test research hypothesis 6 by evaluating the diagnostic efficiency of the DEPI and CDI. Overall, the analyses were performed in order to provide a basis for comparison with Ganellen’s (1996) findings pertaining to relatively high diagnostic efficiency for the CS depression indices. Ganellen’s (1996) study, similar to the present study, was conducted using depressed psychiatric adults (N=315), non-depressed psychiatric adults diagnosed with thoughts disorders (N=320), and non-patient adults (N=700), based on Exner’s (1991) published adult findings. Further, it remains of interest to determine whether findings from the current study are consistent with Exner and Weiner’s (1995) conclusions that the Rorschach CS is able to correctly classify categorically diagnosed depressed adolescents with 50% accuracy on
the basis of DEPI total score cutoffs, with 18% accuracy on the basis of CDI total score
cutoffs, and with 25% accuracy on the basis of co-joint DEPI and CDI total score cutoffs.

Based on statistical methods advocated by Streiner (2003) and Ganellen (1996),
several coefficients were used to evaluate the CS depression indices’ diagnostic
efficiency. For the current study, sensitivity refers to the proportion of categorically
depressed (i.e., DSM-IV-based) adolescents correctly identified as depressed by the DEPI
and CDI, based on pre-established score cutoff points. Specificity, conversely, refers to
the proportion of categorically (i.e., DSM-IV-based) depressed adolescents correctly
identified as non-depressed by the DEPI and CDI, based on pre-established score cutoff
points. False positive rate refers to the proportion of non-categorically depressed
adolescents incorrectly identified as being depressed by the DEPI and CDI, while false
negative rate refers to the proportion of categorically depressed adolescents incorrectly
identified as being non-depressed by the DEPI and CDI.

Ganellen (1996) reports that the diagnostic efficiency of a test depends not only
on being able to identify a particular disorder, but also on the prevalence of the disorder
being assessed among the sample of interest. Positive and negative predictive power
coefficients contribute valuable information about a test above and beyond measures of
sensitivity and specificity. The positive predictive power of the DEPI and CDI refers to
the proportion of adolescents correctly identified as depressed, relative to the overall
number of adolescents actually diagnosed with depression in the sample. Negative
predictive power of the DEPI and CDI refers to the proportion of adolescents correctly
identified as non-depressed, relative to the overall number of adolescents actually having
received non-depressive diagnoses in the sample. Overall correct classification rate
refers to the indices’ proportion of correct classification of subjects into both depressed
and non-depressed categories. The Kappa coefficient refers to the statistically
represented level of agreement between DEPI and CDI total scores and adolescents’
administered diagnoses beyond that accounted for by chance.

Based on recommendations suggested by Exner (1993), the diagnostic efficiency
of the CS depression indices was evaluated according to various total score cutoff points
(i.e., 4, 5, 6, and 7 for the DEPI, and 3, 4, and 5 for the CDI). Further, numerous
permutations of these cutoff points were evaluated using combinations of DEPI and CDI cutoff points to evaluate their contribution to diagnostic CS diagnostic efficiency.

**Hypothesis 6**

H6a: It is hypothesized that for the combined group of depressed and non-depressed psychiatric adolescents, the diagnostic efficiency of the DEPI will approximate the results derived by Ganellen (1996), based on Exner’s (1993) DEPI standardization sample of both depressed and non-depressed adults.

H6b: It is further hypothesized that the contribution of the CDI, according to Exner’s (1993) recommended clinical score cutoff points, will improve DEPI diagnostic efficiency.

**Diagnostic efficiency results**

Tables 7, 8, and 9, respectively, provide diagnostic efficiency statistics for the DEPI, the CDI, and both scales combined, pertaining to their ability to differentiate depressed psychiatric adolescents from non-depressed psychiatric adolescents. An examination of the diagnostic efficiency results reveals several important findings germane to each variable, which will be delineated separately for the DEPI, CDI, and for both indices combined. It is noted, however, that several trends remained consistent when considering the DEPI and CDI both independently and together.

For one, applying increasingly stringent cutoff criteria consistently resulted in notable decreases in false positive rate, and notable increases in false negative rate. Negative predictive power and overall correct classification rate also remained highly stable despite combining the variables or altering the cutoff score permutations, falling roughly around .50. Further, Kappa coefficients also remained roughly stable overall, and quite low (ranging from .004 to .078), indicating considerable lack of agreement between total scores and administered diagnoses.

**DEPI**

Overall, the DEPI demonstrated consistently high false negative rates (i.e., ranging from at least .58 to .99) and consistently low sensitivity (.42 at the highest for the most liberal cutoff of DEPI $> 4$), regardless of cutoff criteria used. It is noted that DEPI
cutoff scores based on Exner’s (1993) recommended clinical level of greater than or equal to 5 resulted in a false negative rate of .83. For the DEPI, an inverse relationship was observed between sensitivity and both positive predictive power and specificity when higher cutoff criteria were applied. That is to say, more stringent cutoff criteria resulted in increases in positive predictive power and specificity, but also a decrease in sensitivity.

The independent cutoff of ≥ 5 for the DEPI resulted in the highest overall classification rate (.55), highest Kappa coefficient (.078), and substantially lower false positive rate (.10, compared with .42 for the cutoff of DEPI ≥ 4), suggesting that it is likely the optimal cutoff in this sample; however, a considerably high false negative rate (.83) and extremely low sensitivity (.17) were also observed. Pertaining to the frequency of DEPI responses, it bears mentioning that only one adolescent from the depressed-only group produced a DEPI total score of seven, five produced total scores of six, and 19 produced total scores of five. No non-depressed psychiatric adolescents obtained scores of seven on the DEPI; however, 13 produced scores of five and two subjects produced scores of six.

**CDI**

The CDI, independently, demonstrated higher sensitivity, overall (i.e., .74 at highest) compared to the DEPI. Similar to the DEPI findings, more stringent cutoff criteria resulted in dramatic decreases in CDI sensitivity (i.e., .74 for CDI ≥3, .42 for CDI ≥4, .12 for CDI ≥5), and increases in specificity. No notable changes in positive or negative predictive power were observed as a product of altering cutoff criteria, both of which remained relatively stable at around .50. The CDI also demonstrated high false negative rates (FNR) similar to the DEPI when Exner’s recommended cutoffs were applied (i.e., CDI ≥4 yielded FNR of .58, compared with FNR of .26 for CDI ≥3).

**Combined DEPI and CDI**

Overall, considering DEPI and CDI total scores in conjunction served to decrease false positive rate (i.e., .31 at the lowest), but also served to increase false negative rate (.67 at the lowest). False Negative rate of classification was consistently worsened by combining the DEPI and CDI, and remained high for all cutoff criteria (i.e., ranging from
As observed for the DEPI, sensitivity remained quite low for all combinations of DEPI and CDI total score cutoffs (.33 at the highest, based on the most liberal sub-clinical criteria of DEPI ≥4, CDI ≥3).

Also similar to the independent DEPI total scores, more stringent cutoff criteria applied to combined DEPI and CDI scores resulted in increases in positive predictive power and specificity, but also a decrease in sensitivity. For DEPI scores of ≥6, higher CDI total score criteria made no appreciable difference in CS diagnostic efficiency. For DEPI scores of ≥5, increases in the stringency of CDI resulted in slight increases in positive predictive power (i.e., increased from .60 to .78), and a simultaneous increase in false negative rate (i.e., increased from .88 to .95).

Table 8
Diagnostic Efficiency Statistics For Different DEPI Total Score Cutoffs

<table>
<thead>
<tr>
<th>Cutoff Scores</th>
<th>SEN</th>
<th>SPEC</th>
<th>PPP</th>
<th>NPP</th>
<th>FPR</th>
<th>FNR</th>
<th>OCC</th>
<th>Cohen’s Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEPI ≥ 4</td>
<td>.42</td>
<td>.58</td>
<td>.48</td>
<td>.43</td>
<td>.42</td>
<td>.58</td>
<td>.50</td>
<td>.004</td>
</tr>
<tr>
<td>DEPI ≥ 5</td>
<td>.17</td>
<td>.90</td>
<td>.63</td>
<td>.54</td>
<td>.10</td>
<td>.83</td>
<td>.55</td>
<td>.078</td>
</tr>
<tr>
<td>DEPI ≥ 6</td>
<td>.04</td>
<td>.99</td>
<td>.75</td>
<td>.53</td>
<td>.01</td>
<td>.96</td>
<td>.53</td>
<td>.030</td>
</tr>
<tr>
<td>DEPI = 7</td>
<td>.01</td>
<td>1.0</td>
<td>1.0</td>
<td>.52</td>
<td>0.0</td>
<td>.99</td>
<td>.52</td>
<td>.007</td>
</tr>
</tbody>
</table>

Note: SEN = Sensitivity; SPEC = Specificity; PPP = Positive Predictive Power; NPP = Negative Predictive Power; FPR = False Positive Rate; FNR = False Negative Rate; OCC = Overall Correct Classification

Table 9
Diagnostic Efficiency Statistics For Different CDI Total Score Cutoffs

<table>
<thead>
<tr>
<th>Cutoff Scores</th>
<th>SEN</th>
<th>SPEC</th>
<th>PPP</th>
<th>NPP</th>
<th>FPR</th>
<th>FNR</th>
<th>OCC</th>
<th>Cohen’s Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDI ≥ 3</td>
<td>.74</td>
<td>.26</td>
<td>.48</td>
<td>.51</td>
<td>.74</td>
<td>.26</td>
<td>.49</td>
<td>-.006</td>
</tr>
<tr>
<td>CDI ≥ 4</td>
<td>.42</td>
<td>.52</td>
<td>.45</td>
<td>.49</td>
<td>.48</td>
<td>.58</td>
<td>.47</td>
<td>-.060</td>
</tr>
<tr>
<td>CDI = 5</td>
<td>.12</td>
<td>.87</td>
<td>.46</td>
<td>.52</td>
<td>.13</td>
<td>.88</td>
<td>.51</td>
<td>-.011</td>
</tr>
</tbody>
</table>

Note: SEN = Sensitivity; SPEC = Specificity; PPP = Positive Predictive Power; NPP = Negative Predictive Power; FPR = False Positive Rate; FNR = False Negative Rate; OCC = Overall Correct Classification
Table 10
Diagnostic Efficiency Statistics For Different DEPI and CDI Total Score Combined Cutoffs

<table>
<thead>
<tr>
<th>Cutoff Scores</th>
<th>SEN</th>
<th>SPEC</th>
<th>PPP</th>
<th>NPP</th>
<th>FPR</th>
<th>FNR</th>
<th>OCC</th>
<th>Cohen’s Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEPI ≥4, CDI ≥3</td>
<td>.33</td>
<td>.69</td>
<td>.50</td>
<td>.53</td>
<td>.31</td>
<td>.67</td>
<td>.52</td>
<td>.024</td>
</tr>
<tr>
<td>DEPI ≥4, CDI ≥4</td>
<td>.20</td>
<td>.81</td>
<td>.49</td>
<td>.52</td>
<td>.19</td>
<td>.80</td>
<td>.51</td>
<td>.008</td>
</tr>
<tr>
<td>DEPI ≥4, CDI ≥5</td>
<td>.07</td>
<td>.94</td>
<td>.53</td>
<td>.52</td>
<td>.06</td>
<td>.93</td>
<td>.52</td>
<td>.012</td>
</tr>
<tr>
<td>DEPI ≥5, CDI ≥3</td>
<td>.12</td>
<td>.92</td>
<td>.60</td>
<td>.53</td>
<td>.08</td>
<td>.88</td>
<td>.54</td>
<td>.049</td>
</tr>
<tr>
<td>DEPI ≥5, CDI ≥4</td>
<td>.09</td>
<td>.94</td>
<td>.57</td>
<td>.53</td>
<td>.06</td>
<td>.91</td>
<td>.53</td>
<td>.026</td>
</tr>
<tr>
<td>DEPI ≥5, CDI ≥5</td>
<td>.05</td>
<td>.99</td>
<td>.78</td>
<td>.53</td>
<td>.01</td>
<td>.95</td>
<td>.53</td>
<td>.037</td>
</tr>
<tr>
<td>DEPI ≥6, CDI ≥3</td>
<td>.03</td>
<td>1.0</td>
<td>1.0</td>
<td>.53</td>
<td>0.0</td>
<td>.97</td>
<td>.53</td>
<td>.029</td>
</tr>
<tr>
<td>DEPI ≥6, CDI ≥4</td>
<td>.02</td>
<td>1.0</td>
<td>1.0</td>
<td>.52</td>
<td>0.0</td>
<td>.98</td>
<td>.53</td>
<td>.021</td>
</tr>
<tr>
<td>DEPI ≥6, CDI ≥5</td>
<td>.007</td>
<td>1.0</td>
<td>1.0</td>
<td>.52</td>
<td>0.0</td>
<td>.993</td>
<td>.52</td>
<td>.007</td>
</tr>
<tr>
<td>DEPI=7, CDI ≥4</td>
<td>.007</td>
<td>1.0</td>
<td>1.0</td>
<td>.52</td>
<td>0.0</td>
<td>.993</td>
<td>.52</td>
<td>.007</td>
</tr>
</tbody>
</table>

Note: SEN = Sensitivity; SPEC = Specificity; PPP = Positive Predictive Power; NPP = Negative Predictive Power; FPR = False Positive Rate; FNR = False Negative Rate; OCC = Overall Correct Classification

Results of Data Analyses Set 4

The analyses in set 4 were conducted to test the final research hypothesis by evaluating the incremental validity of the DEPI and CDI, when compared with the RADS and the MMPI-A Depression scale. Tabachnick and Fidell (1996) and Davis and Offord (1997) recommend the use of hierarchical logistic regression to determine incremental validity when the dependent variable of interest (i.e., classification as depressed/non-depressed based on DSM-IV criteria) is dichotomous in nature. Peng, Lee, and Ingersol (2002) report that, when using hierarchical logistic regression, results should present sufficient information to address (a) an overall evaluation of the proposed model, (b) statistical tests of each individual predictor, (c) statistics pertaining to the goodness-of-fit of the proposed model, and (d) results pertaining to the assessment of the predicted probabilities of the model.
Tabachnick and Fidell (1996) note that in a hierarchical general linear model (GLM) the $F$ ratio serves as the omnibus test of the null hypothesis of a normally-distributed multivariate data set. In other words, the $F$ test provides an estimate of the significance of the overall relationship of the independent variables to the outcome variable, and assumes a normal distribution. In contrast, the likelihood ratio test provides a least squares comparison via chi-square statistics to demonstrate the extent of the difference between null and alternate models of the contribution of the predictor to outcome variables (Davis and Offord, 1997; Peng et al., 2002). That is to say, the overall test of the model provides an estimate of the improvement of the proposed model over the null, or intercept-only, model in which it is predicted that all observations belong to the largest outcome of the dichotomous outcome category without consideration of the contribution of individual predictors. The test of the overall logistic model is most often accomplished by means of the likelihood ratio test.

Differences between GLM and logistic models also extend to selected statistical tests of each individual predictor. Whereas $R^2$ is used in GLM to estimate the proportion of variance of the dependent variable explained by the model predictor variables (Tabachnick and Fidell, 1996), several researchers (Davis and Offord, 1997; Peng et al., 2002; Tabachnick and Fidell, 1996) concur that logistic regression models have failed to derive a satisfactory equivalent to $R^2$. Rather, to test the statistical significance of individual predictors in a logistic regression equation, logistic models commonly rely on beta coefficients ($\beta$) and odds ratio estimate ($e^{\beta}$), representations of the Wald chi-square statistic, to determine the degree and significance of each independent variable towards predicting membership within the dichotomous dependent variable of classification (i.e., depressed-only vs. non-depressed-only)(Peng et al., 2002).

Peng et al. (2002) further note that logistic models derive a Pearson chi-square statistic (i.e., Hosmer-Lemeshow, or H-L, test) to determine the goodness-of-fit of the proposed logistic model of independent variables compared to the actual observed outcomes of the dichotomous dependent variable. In other words, the goodness-of-fit statistic assesses the fit of a given logistic model to the actual outcomes described by the dependent variable (i.e., depressed vs. non-depressed). Davis and Offord (1997) and Peng et al. (2002) also caution against the use of “pseudo-$R^2$ measures” (Davis and
Offord, 1997; p. 501) often utilized as subsidiary measures of goodness-of-fit in logistic equations, the most common of which include Cox and Snell’s $R^2$ and Nagelkerke’s $R^2$. Davis and Offord (1997) and Peng et al. (2002) report that neither measure accounts for the meaning of explained variance in a logistic equation, neither corresponds to predictive efficiency, and “neither can be tested in an inferential framework,” (Peng et al., 2002; p. 6). For these reasons, Peng et al. (2002) advocate that Cox and Snell, and Nagelkerke measures be considered as “supplementary to other, more useful evaluative indices, such as the overall evaluation of the model, tests of individual regression coefficients, and the goodness-of-fit statistic,” (p.6) and call for the use of diagnostic efficiency statistics such as those performed in analysis set 3 of this study.

**Preliminary Analysis**

A matrix containing correlations, means, and standard deviations pertaining to all variables in Analyses set 4 is presented in Table 11. Pearson coefficients were derived to describe the relationship between the continuous variables (i.e., MMPI-A, RADS, DEPI, and CDI). Point biserial correlations were derived to describe the relationship between the dichotomous dependent variable, diagnosis, based on recommendations provided by Tabachnick and Fidell (1996) and Witte (1989). An inspection of the matrix indicates no cause for concern due to excessive covariate effects, given Tabachnick and Fidell’s (1996) recommendation that correlations above .90 are sufficiently high to create problematic multicollinearity among variables of interest in a regression equation. Tabachnick and Fidell (1996) also caution against using small sample sizes relative to the number of predictor variables used in the model, and recommend that a logistic analysis contain at least 5 times the number of cases as predictor variables in the research design. Given the initial N of 222 cases, this was not a cause for concern in the current study.

Tabachnick and Fidell (1996), Peng et al. (2002), and Davis and Offord (1997) note that logistic, non-linear regression equations are robust to violations of multivariate normality due to their binomial distribution; however, as Tabachnick and Fidell (1996) note, power may be enhanced if the predictor variables demonstrate normality and linearity because the distribution of the dichotomous dependent variable is presumed to describe the distribution of standard error equal to observed minus actual dichotomous
outcomes. Thus, the presence of excessive outliers among the data distribution has the potential to adversely affect the actual dichotomous outcome by decreasing the overall model evaluation and goodness of fit if a sufficient number of cases result in skewed representation of the data.

Screening of the initial data set (n=295) indicated that there were 73 missing cases (i.e., adolescents who had not completed either the RADS or the MMPI-A), which were deleted from the sample. A visual inspection of these cases indicated that they encompassed a roughly equivalent proportion of both depressed and non-depressed subjects, and was not felt to be problematic. Further inspection of the remaining cases (n=222) using logistic regression diagnostics suggested that there were six outliers with excessive studentized residual values (i.e., cases with outlying deviance scores of greater than +/- 2 or 3)(Tabachnick and Fidell, 1996). The highest of these values reached 2.58, while the lowest was –2.23. Further inspection of the data indicated problematically high values pertaining to RADS and MMPI-A scores, given the DSM-IV diagnosis assigned to each case (i.e., one non-depressed case had a RADS score of 117 and an MMPI-A scale 2 score of 68, one depressed subject had a RADS score of 38 and an MMPI-A scale 2 score of 46, etc.). The presence of this data was not easily explained, and ultimately it was felt likely that these outlying scores were the product of examiner or investigator error. In two of the cases it was felt that the excessive influence was the direct result of over-inflated RADS scores, which appeared likely to be attributable to a “fake bad” response set on the part of the examinee given the instrument’s lack of validity indices. In the majority of cases, however, high residual values were attributable both to over-inflated RADS and MMPI-A Scale 2 scores, relative to their administered DSM-IV diagnosis.

On the basis of the excessive influence toward the improvement of the proposed model over null, a decision was made to remove the six outliers exerting excessive influence. This resulted in slight improvement in the overall evaluation of the model (i.e., increase in $\chi^2$ value by 27.234), a slight improvement in goodness of fit (i.e., an increase in $\chi^2$ value by .176), and slight-to-moderate improvements for all individual predictor variables (i.e., most significant increases consisted of increase in $\beta$ coefficients of 1.91 pertaining to the Constant, and increase of .03 pertaining to the RADS).
Based on these improvements, it was determined that the modified model represented a better fit of the predictor variables to the dichotomous outcome of depression, and the initial cases demonstrating excessive residual influence were excluded from subsequent analyses.

It is noted that seven additional outliers were also identified in the revised model with studentized residual values slightly in excess of 2 (i.e., 2.2087 at highest value, and –2.2479 at lowest value); however, the decision was made to retain the cases with identified residual values because a subsequent analysis of the effect of removing those outlying values did not contribute significantly to the logistic regression outcomes (i.e., overall evaluation of the model, tests of the individual predictors, and goodness-of-fit), particularly with regard to the CS predictor variables of interest. Further, theoretical justification for removing the outlying cases was dubious due to lacking empirical findings concerning the false positive/false negative rate of the RADS and MMPI-A Scale 2. In other words, it was felt that there was a likely probability that removing numerous cases with smaller residual values from the equation would detract from the representativeness of the clinical sample.

### Table 11

Pearson and Point Biserial Correlation Coefficients, Means, & Standard Deviations Among MMPI-A, RADS, DEPI, CDI, and Diagnosis (N=222)

<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5(^a)</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. MMPI</td>
<td>1.0</td>
<td>.67**</td>
<td>.04</td>
<td>-.04</td>
<td>.55**(^a)</td>
<td>60.90</td>
<td>13.05</td>
</tr>
<tr>
<td>2. RADS</td>
<td>1.0</td>
<td>-.057</td>
<td>-.06</td>
<td>.64**(^a)</td>
<td>70.78</td>
<td>19.97</td>
<td></td>
</tr>
<tr>
<td>3. DEPI</td>
<td>1.0</td>
<td>.13*</td>
<td>.04(^a)</td>
<td>3.32</td>
<td>1.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. CDI</td>
<td>1.0</td>
<td>.04(^a)</td>
<td>3.21</td>
<td>1.12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. DX</td>
<td>1.0</td>
<td>1.46</td>
<td>.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: MMPI = MMPI-A Basic Scale 2; RADS = Reynolds Adolescent Depression Scale; DEPI = CS Depression Index; CDI = CS Coping Deficit Index; DX = Diagnostic Group; \(^a\) = coefficient derived from point biserial coefficient; \(* p = .05; ** p < .01\)
Results of Hierarchical Logistic Regression Analysis 1

Significant correlations were observed between the MMPI-A 2 scale and the RADS ($r = .67$), the DEPI and CDI ($r = .13$), and the MMPI-A 2 scale and diagnosis, ($r = .55$) and RADS and diagnosis ($r = .64$). A non-significant relationship was observed between DEPI and diagnosis and between CDI and diagnosis.

Licht (1995) notes that the order of entry of predictor variables into a hierarchical regression equation is an important factor to consider, as the effect of variables entered earlier in the equation is controlled by their relationships with subsequently-entered variables. As a result, many authors (Blais, Hilsenroth, Castlebury, Fowler, & Baity, 2001; Cohen and Cohen, 1983; David and Offord, 1997; Licht, 1995; Tabachnick and Fidell, 1996) advocate that the decision regarding the order of inclusion of variables from a given hierarchical regression equation ultimately be made by the investigator on the basis of prior research (i.e., a priori information regarding the relationship between variables), or on the basis of the observed, empirical relationship among the variables of interest (i.e., observation of the contributions of each variable to the maximum likelihood estimate and goodness-of-fit of the prediction equation).

The decision was made to perform two separate sets of analyses with the predictor variables entered into the logistic equation in a step-wise, contingent manner. Based on the exploratory nature of this study and general lack of consensus regarding the incremental validity of the CS depression indices for adolescent populations, initially it was intended that two sets of analyses would be performed, with the order of entry of the predictor variables reversed. In so doing, the degree of effect of each of the subsequently-entered variables (i.e., block containing both DEPI and CDI in the first analysis, and two separate blocks containing RADS and MMPI-A Scale 2, respectively, in the second set of analysis) would be determined, while controlling for their relationships with the previously entered variables.

Table 12 contains the results of the first logistic regression analysis in which the RADS and MMPI-A were entered first into the equation using a forward stepwise conditional procedure to evaluate the contribution of each predictor variable to the overall maximum likelihood of prediction of diagnostic status. Table 12 contains statistics
pertinent to the evaluation of the overall model, individual contributions and odds ratio of
the individual predictor variables, goodness-of-fit of the model, standard error, degrees of
freedom, and significance. The variables were entered in separate blocks with RADS
scores entered in the first block, MMPI-A Scale 2 scores entered in the second block, and
DEPI and CDI entered in a step-wise manner in the third block.

Based on the overall contribution of all predictor variables, the likelihood ratio
test of the overall model \( \chi^2(2) \) of 113.780 was significant \((p<.001)\), indicating that the
maximum likelihood model represented by the variables selected for inclusion
demonstrates improvement over the intercept-only, or null, model in predicting group
membership (depressed vs. non-depressed adolescents). The Hosmer-Lemeshow (H-L)
test of the goodness-of-fit of the model further yielded a \( \chi^2(8) \) of 7.1369 and was non-
significant \((p>.05)\), suggesting that the model fit the data well. In other words, the H-L
provided justification for acceptance of the null hypothesis that the model represented a
good fit to the data. Based on actual vs. predicted values of depressed vs. non-depressed
adolescents, the model represented by Table 12 was able to predict group membership
within the depressed group (i.e., sensitivity) 83.90% of the time, membership within the
non-depressed group (i.e., specificity) 79.59% of the time. Overall correct classification
of the model was 81.94%.

An inspection of the statistical tests of the individual predictors indicated that,
based on the forward stepwise conditional order of entry procedure, the predictor
variables of DEPI and CDI were not selected into the model. In other words, the
contribution of the DEPI and CDI towards predicting the maximum likelihood of
determining diagnostic classification was not sufficient (i.e., they resulted in change in
Log Likelihood of less than .01 percent) to warrant inclusion. Wald’s chi-square tests of
both RADS scores, \( \chi^2(1) \) of 27.3513, \( p<.001 \), and MMPI-A Scale 2 scores, \( \chi^2(1) \) of
9.0308, \( p<.01 \), indicate that both variables were significant predictors of depression.

More specifically, the odds ratio estimate \( e^\hat{\theta} \) for RADS indicates that for each point
increase in RADS total score, the odds of being classified as depressed increase by .3617
(and, conversely, odd of being classified as non-depressed decrease by .3617). Similarly,
the odds ratio estimate for the MMPI-A Scale 2 indicates that for each point increase the
odds of being classified as depressed increase by .1905.
Table 12
Logistic Hierarchical Regression Analysis 1, with Conditional Variable Entry of Variables in the Following Order: RADS, MMPI-A, DEPI, CDI (N=216)

<table>
<thead>
<tr>
<th>Predictor</th>
<th>β</th>
<th>SE β</th>
<th>Wald's $\chi^2$</th>
<th>df</th>
<th>p</th>
<th>$e^\beta$ (odds ratio)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>8.8140</td>
<td>1.2185</td>
<td>52.3259</td>
<td>1</td>
<td>.0000</td>
<td>NA</td>
</tr>
<tr>
<td>RADS</td>
<td>.0800</td>
<td>.0153</td>
<td>27.3513</td>
<td>1</td>
<td>.0000</td>
<td>.3617</td>
</tr>
<tr>
<td>MMPI-A</td>
<td>.0580</td>
<td>.0193</td>
<td>9.0308</td>
<td>1</td>
<td>.0027</td>
<td>.1905</td>
</tr>
</tbody>
</table>

Test

$\chi^2$  df  p
Overall model evaluation
Likelihood ratio test  113.780  2  .0000
Goodness-of-fit test
Hosmer & Lemeshow     7.1369  8  .5219
Cox and Snell $R^2$   .409   NA  NA
Nagelkerke $R^2$ (Max rescaled $R^2$) .548   NA  NA

Note. $-2$ Log Likelihood=183.806; RADS=Reynolds Adolescent Depression Scale total score; MMPI-A=MMPI-A Scale 2 t-scores; NA=not applicable.

Results of Hierarchical Logistic Regression Analysis 2

The results of the first model ostensibly predicted that entering the DEPI and CDI prior to the RADS and MMPI-A using a step-wise, conditional method would not improve on the contribution of the RADS and MMPI-A towards prediction of depression. When an attempt was made to enter the DEPI and CDI prior to the RADS and MMPI-A in block one of the second model, the step-wise conditional entry criteria (i.e., change in Log Likelihood of less than .01 percent) rejected the CS variables even before considering the contribution of the RADS and MMPI-A. The equation ultimately resulted in a model identical to that expressed in Table 12.

In the interest of demonstrating the extent of the contribution of DEPI and CDI towards the model prediction of the dichotomous outcome variable, however, the
decision was made to enter all variables using the forced-entry method. Table 13 contains the results of the final, follow-up logistic regression analysis containing all four predictor variables entered in a step-wise, forced order of entry not contingent upon degree of contribution towards maximum predictive likelihood. As was done in the first logistic analysis, RADS scores were entered first in block 1, MMPI-A Scale 2 scores were entered in block 2, and DEPI and CDI total scores were entered into the equation in Block 3.

Compared to the first logistic model containing only RADS and MMPI-A, the second model, $\chi^2(2)$ of 114.364, $p<.001$, continued to retain its effectiveness in predicting dichotomous group membership above that of the null. That is to say, despite the negligible contribution of the DEPI and CDI, the model continued to retain its predictive utility based on the contributions of the RADS and MMPI-A. The H-L test of goodness of fit similarly yielded a $\chi^2(8)$ of 6.3014, $p<.6135$, indicating a good fit of the model to the data based on actual vs. predicted case values. The second model demonstrated sensitivity of 82.20%, specificity of 78.57%, and overall correct classification rate of 80.56%. As previously indicated, Wald's chi-square tests of both DEPI, $\chi^2(1)$ of .0535, $p<.8170$, and CDI scores, $\chi^2(1)$ of .5612, $p<.4538$, indicate that neither variable had a significant multivariate effect on the outcome of diagnosis.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>$\beta$</th>
<th>SE $\beta$</th>
<th>Wald's $\chi^2$</th>
<th>$df$</th>
<th>$p$</th>
<th>$e^\beta$ (odds ratio)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>8.5475</td>
<td>1.3665</td>
<td>39.1241</td>
<td>1</td>
<td>.0000</td>
<td>NA</td>
</tr>
<tr>
<td>RADS</td>
<td>.0811</td>
<td>.0155</td>
<td>27.3357</td>
<td>1</td>
<td>.0000</td>
<td>.3713</td>
</tr>
<tr>
<td>MMPI-A</td>
<td>.0569</td>
<td>.0195</td>
<td>8.5342</td>
<td>1</td>
<td>.0035</td>
<td>.1885</td>
</tr>
<tr>
<td>DEPI</td>
<td>.0380</td>
<td>.1642</td>
<td>.0535</td>
<td>1</td>
<td>.8170</td>
<td>.0000</td>
</tr>
<tr>
<td>CDI</td>
<td>-.1250</td>
<td>.1668</td>
<td>.5612</td>
<td>1</td>
<td>.4538</td>
<td>.0000</td>
</tr>
</tbody>
</table>
Table 13 (Continued)

<table>
<thead>
<tr>
<th>Test</th>
<th>$\chi^2$</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall model evaluation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Likelihood ratio test</td>
<td>114.364</td>
<td>4</td>
<td>.0000</td>
</tr>
<tr>
<td>Goodness-of-fit test</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hosmer &amp; Lemeshow</td>
<td>6.3014</td>
<td>8</td>
<td>.6135</td>
</tr>
<tr>
<td>Cox and Snell $R^2$</td>
<td>.411</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Nagelkerke $R^2$ (Max rescaled $R^2$)</td>
<td>.550</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

Note. -2 Log Likelihood=183.221. RADS=Reynolds Adolescent Depression Scale total score. MMPI-A=MMPI-A Scale 2 t-scores. DEPI=DEPI total scores. CDI=CDI total scores. NA=not applicable.

Hypothesis 7

H7: Within the clinical adolescent group, there will be a significant incremental relationship between Rorschach DEPI and CDI total scores, MMPI-A Basic Scale 2 Scores, Reynolds Adolescent Depression Scale Total Scores, and diagnoses of depression administered using the DSM-IV.

Davis and Offord (1997) assert that evaluation of incremental validity using logistic regression equations can be accomplished by comparing the two estimates of log likelihood derived by two models of interest: one including the predictor of variable and one without. As anticipated by the aforementioned lack of significant effect by both DEPI and CDI, the log likelihood estimates of the two models does not indicate significant change. Subtracting the log likelihood for model 1 (183.806) from that of model 2 (183.221) results in a difference of .585 (Chi-square value of 1.17, p>.20). In other words, the model efficiency in predicting depressed vs. non-depressed was not sufficiently improved by the addition of the DEPI and CDI.
CHAPTER 5

DISCUSSION

The current study sought to better understand the utility of the Rorschach Comprehensive System’s (CS; Exner, 1993) Depression Index (DEPI) and Coping Deficit Index (CDI), which serve as the CS’s dimensional predictors of depressive symptoms and disorders. To date, the majority of research investigating the DEPI and CDI has been conducted with adults (Meyer & Archer, 2001). The current study was conducted using depressed and non-depressed adolescent inpatients to address questions regarding the depression indices’ psychometric validity with regard to that population. The study was also conducted to contribute to the limited and somewhat discrepant body of Rorschach empirical research involved in evaluating the psychometric efficacy and validity of the CS depression indices (Wood, Lilienfeld, Garb, and Nezworski, 2000; Wood, Nezworski, and Stejskal; 1997) and the need to replicate Exner’s (1993; Exner and Weiner, 1996) findings.

The current study compared DEPI and CDI total scores by age and DSM-IV diagnosis, examined their degree of co-relationship, and investigated their diagnostic efficiency, predictive validity, and incremental validity pertaining to Major Depressive Disorder. Primary variables included in the study were DEPI and CDI total scores. MMPI-A Scale 2 scores and RADS total scores were used in the evaluation of CS incremental validity. Additionally, Exner’s (Exner and Weiner, 1995) normative sample of non-clinical adolescents was used as a comparison group as part of the analysis of predictive validity.

Summary of Results

Four separate sets of analyses were conducted to evaluate the utility of using Exner’s (1993) Rorschach Comprehensive System’s (CS) depression indices as a dimensional assessment method. The DEPI and CDI were tested by examining the
following: differences in total scores by age and diagnostic status; the relationship between the two variables (DEPI and CDI) considering diagnostic group (i.e., depressed-only, non-depressed clinical, and combined clinical) and age (i.e., 13-17 years of age); the diagnostic efficiency of the DEPI and CDI, considered as both individual and combined indices; and the incremental validity of the DEPI and CDI, when examined with MMPI-A and RADS.

In the first set of analyses age-based, pair-wise t-tests for significant differences between the means of DEPI and CDI total scores were undertaken for depressed-only clinical adolescents, non-depressed/“normal” adolescents, and non-depressed clinical adolescents. After statistical correction to account for over-inflation of family-wise test error rate, results revealed that neither the depressed-only nor non-depressed-only clinical groups performed significantly higher on the DEPI than their non-clinical “normal” control group counterparts at any age. Further, comparison of the two clinical adolescent groups revealed no significant differences in depressed vs. non-depressed DEPI scores at any age.

Pertaining to the CDI, depressed-only clinical 14- and 15-year-olds, as well as non-depressed clinical adolescents of all ages, obtained significantly higher total scores than their non-depressed/“normal” counterparts, given correction for group-wise error inflation. Further, there were no significant differences in CDI scores for any age, based on a comparison of the depressed and non-depressed clinical groups.

Results of the second set of analyses revealed that there were significant relationships between the DEPI and CDI, but not for all comparison groups. A small-to-medium-sized relationship (.229) was observed between the two indices for the depressed-only clinical group, and a small relationship (.144) was observed for the combined clinical group of both depressed and non-depressed adolescents. The DEPI and CDI failed to demonstrate a significant degree of co-relationship for the non-depressed group of clinical adolescents. The omnibus test of the main effect of age on both DEPI or CDI total scores (i.e., two separate tests performed) indicated that age was not a significant contributor to the variance in test scores for either test, based on the combined group of depressed-only and non-depressed clinical adolescents.
The third set of analyses provided information pertaining to the diagnostic efficiency of the DEPI and CDI. Overall, for both DEPI, CDI, and combinations of the two indices, many coefficients remained stable. Use of increasingly stringent cutoff criteria consistently resulted in substantial decreases in false positive rate and increases in false negative rate. Kappa coefficients remained extremely low (ranging from .004 to .078), and negative predictive power and overall correct classification rate remained consistent at around .50 for all combinations of variables.

The DEPI, independent of the CDI, demonstrated extremely low overall sensitivity (.42 at the highest for the most liberal cutoff of DEPI ≥ 4) and consistently high false negative rates (i.e., ranging from a low of .58 to a high of .99), regardless of total score cutoff criteria used. The DEPI’s positive predictive power and specificity improved with increasingly strict cutoff criteria, while sensitivity decreased. A comparison of the various cutoff score points tested as part of the sensitivity study revealed that DEPI ≥ 5 yielded the highest overall classification rate (.55), highest Kappa coefficient (.078), and substantially lowered false positive rate (.10, compared with .42 for the cutoff of DEPI ≥ 4). As a result, while not ideal, that cutoff point was felt to represent the most optimal cutoff. Overall, a low number of depressed adolescents actually produced markedly highly elevated scores (i.e., 6 or 7) on the DEPI (i.e., one score of 7, and five scores of 6).

CDI sensitivity was somewhat improved over that of the DEPI at its highest value (.74), but decreased dramatically when more stringent cutoff score guidelines (CDI ≥4 = .42, CDI ≥5 = .12) were applied. Conversely, specificity and false negative rates increased substantially with more stringent cutoff scores. No notable changes in positive or negative predictive power were observed as a result of altering cutoff criteria, with both remaining stable at around .50.

Combining DEPI and CDI resulted in decreased false positive rate (i.e., .31 at the lowest), but also substantially increased false negative rate (.67 at the lowest). False negative rate was also consistently worsened by combining the DEPI and CDI, remaining extremely high for all cutoff criteria (i.e., ranging from .67 to .993). For all combinations of DEPI and CDI total scores, sensitivity remained quite low (.33 at the highest, based on the most liberal sub-clinical criteria of DEPI ≥4, CDI ≥3). Similar to DEPI alone,
application of more stringent cutoff criteria resulted in increases in positive predictive power and specificity, but also a decrease in sensitivity. For DEPI scores of $\geq 6$, CDI, total scores made no appreciable difference in CS diagnostic efficiency. For DEPI scores of $\geq 5$, increases in the stringency of CDI resulted in slight increases in positive predictive power (i.e., increased from .60 to .78), and a slight increase in false negative rate (i.e., increased from .88 to .95).

Analysis of the incremental validity of the CS depression indices using a logistic hierarchical regression analysis indicated that neither the DEPI nor the CDI contributed significantly towards predicting depressive diagnostic status relative to or independent of the contributions of the MMPI-A Scale 2 (depression scale) or the RADS. Wald’s chi-square tests of both DEPI, $\chi^2(1)$ of .0535, $p<.8170$, and CDI, $\chi^2(1)$ of .5612, $p<.4538$, indicated that neither variable had a significant multivariate effect on the outcome of diagnosis. Order of entry of the DEPI and CDI proved to be negligible due to the insufficient degree of contribution of each towards overall maximum likelihood of prediction of diagnostic status (change in Log Likelihood of less than .01 percent accounted for by the variables).

Interpretation of Results

Differences in Scores by Diagnosis

The first set of analyses tested the relationship between mean DEPI and CDI scores by comparing them separately for the three diagnostic groups of interest (i.e., non-clinical “normal,” depressed-only clinical, and non-depressed clinical). Although overall findings were quite discouraging compared with hypothesized outcomes, the results remain noteworthy and have practical importance given the dearth of research investigating the construct and predictive validity of the CS depression indices. Interpretation of results is presented subsequently by diagnostic group-wise comparison, according to the pre-stated hypotheses. Also, overall interpretation of the combined findings follows the group-based discussions.
Clinical vs. Non-Clinical

For the DEPI, group-wise comparisons for all ages failed to support hypothesis 1 when multiple test-wise error inflation was accounted for using a Bonferroni correction (Tabachnick & Fidell, 1996). Both depressed-only and non-depressed-only clinical 15-year-olds performed higher than non-depressed “normals” of the same age, but not significantly so given this test-wise correction. It is discouraging that the findings in analysis set 1 are consistent with previous research reporting disappointing DEPI predictive validity (Meyer & Archer, 2001; Wood, Lilienfeld, Garb, and Nezworski, 2000; Wood, Nezworski, and Stejskal; 1997), and also that the current data are inconsistent with Ganellen’s (1996) psychometric efficiency study involving Exner’s (1993) adult DEPI data. While limited findings by Exner and Weiner (1995) suggest that the DEPI’s predictive validity may be somewhat diminished for clinical adolescent populations (i.e., positive predictive power of .50), the consistent findings in analysis set 1 suggest that this is highly problematic for adolescent populations.

For the CDI, significant differences between depressed-only clinical and non-clinical/“normal” populations were observed for only the 14- and 15-year-old groups, which was somewhat difficult to understand. While the 13- and 16-year-old groups yielded higher CDI scores, these were not significant when accounting for test-wise inflation of error. All age groups within the non-clinical group also demonstrated significant differences in CDI scores compared with those of the non-depressed clinical group, which is not supported by hypothesis 2. While largely preliminary due to the lack of confirmatory studies replicating Exner and Weiner’s (1995) prior CDI research, the current results appear consistent with their findings that the CDI can be expected to demonstrate low predictive accuracy (18%) for adolescent populations.
Clinical depressed vs. clinical non-depressed

Results of analysis set 1 were unilaterally unsupportive of hypothesis 2 for the DEPI in that no significant differences were observed between depressed-only and non-depressed psychiatric adolescents within the clinical groups. Alternately, no significant differences in CDI scores were observed between depressed and non-depressed clinical groups.

Overall Interpretation

Smith and Glass (1987) assert that if a test is said to demonstrate adequate construct validity, it is purported to be able to accurately measure/identify the construct it was designed to measure. Thus, if results of a test that purports to measure one, specific construct are derived from two populations presumed to represent discrepant representations of that construct, significant differences should be identified between those two groups as a result of the test’s construct validity. Smith and Glass also add that the validity of a psychometric test must always depend on the test’s context and the purpose for which it is used (e.g., measurement of depression for adolescents, children, or adults; measurement of depression for non-clinical, inpatient, etc.). As a result, it is felt that the preliminary findings in analysis set 1 call for one of the following conclusions, to be tested and clarified further in the remainder of analyses: (1) the DEPI is not a valid dimensional measure of the overall construct of depression for adolescents, or (2) there was sufficient error in the data set used in analysis 1 to account for a high proportion of variance over and above that accounted for by the test scores.

To accept the first of these explanations is to conclude that unknown factors related to the DEPI and CDI, but at the same time unrelated to depression, are contributing to considerable variance in the indices’ total scores. Findings pertaining to the CDI in analysis set 1 are not entirely surprising, given the empirical research to date suggests that the scale is not a good independent predictor of depressive disorders (Exner, 1993; Exner and Weiner, 1995). In other words, given that the CDI in and of itself has been proposed to predict only one typology of depressive symptomatology (i.e., hopelessness ideation) and contribute minimally on an independent basis to prediction of depressive disorders (18%), the current data appear to be consistent with prior research.
The current results pertaining to the DEPI are consistent with several past studies that have reported low predictive validity coefficients for both adolescents and adults (Archer & Krishnamurthy, 1997; Exner and Weiner, 1995; Meyer & Archer, 2001; Wood, Lilienfeld, Garb, and Nezworski, 2000; Wood, Nezworski, and Stejskal, 1997), and are not consistent with those reporting adequate predictive validity pertaining to adults (Exner, 1993; Ganellen, 1996). Although this will be explored further in the current study, these preliminary findings appear to have negative implications for the Rorschach CS. To date Exner (1993), Weiner (Exner and Weiner, 1995), and Ganellen (1996) have remained the strongest proponents of the DEPI’s ability to assess the construct of depression. As previously mentioned, the DEPI’s construct and predictive validity remain largely unconfirmed for adolescents. Further, the limited amount of research that has addressed the predictive validity of the DEPI for adolescents (Archer & Krishnamurthy, 1997; Exner and Weiner, 1995) has not demonstrated promising results. The current study indicates that the DEPI does not accurately measure the construct(s) for which it was designed for adolescents. Further, it is felt that this explanation likely accounts for lack of consistently significant differences in DEPI total scores for depressed-only and non-depressed clinical adolescents.

**Relationship Between DEPI And CDI**

The second set of analyses resulted in mixed findings pertaining to the relationship between DEPI and CDI. As predicted by hypothesis 4, the two scales demonstrated significant relationships between the two variables for members of both the depressed-only, and combined depressed-only and non-depressed clinical groups. The data did not completely support hypothesis 4, however, in that DEPI and CDI were not significantly related for the non-depressed clinical group.

Overall, these results are not easily understood, and may best be explained by limitations in the psychometric validity of the CS depression indices and/or error variance, as previously discussed in reference to the current study’s first set of analyses. Proponents of the Rorschach CS, however, might posit that it is reasonable to suspect that lack of a relationship pertaining to the non-depressed clinical group is supportive of Exner’s (1993) claim that the DEPI and CDI measure different typologies of depression-
related phenomena. In other words, perhaps the qualitative difference in constructs being measured by both DEPI and CDI is sufficiently unique so as to be significantly correlated only when depressive symptoms and/or disorders are present.

For example, the Tripartite model of depression and anxiety posits that the constructs of physiological hyperarousal (PH) and positive affect (PA) are present primarily in cases of anxiety, and not in depression (Clark, Steer, & Beck, 1994). That is to say, perhaps the inability of the DEPI and/or the CDI to assess variance attributable to these factors/constructs, which are unique to other non-depressive disorders, is detracting from the degree of co-relation of the DEPI and CDI for the clinical non-depressed group.

In a similar, converse manner perhaps the relationship between DEPI and CDI is best explained by factors common only to depressive disorders (i.e., lack of positive affect). That is to say, if the DEPI and CDI are actually valid measures of the factor/construct “lack of positive affect,” this could perhaps account for the observed total score increases and corresponding increase in shared variance as a result of the contribution of the depressed-only group to the combined clinical group. This explanation appears particularly applicable to the CDI, which is purported to be primarily a measure of helplessness ideation (Exner, 1993).

Main Effect of Age On DEPI and CDI Scores

Interest in the relationship between age and DEPI and CDI performance stemmed largely from Exner’s (1993) assertion that test-retest reliability of the Egocentricity Index [(EI); 3r+(2)/R], a key component of the DEPI, becomes increasingly stable during adolescence. Further, numerous epidemiological studies have suggested that older individuals may demonstrate a higher prevalence of depressive symptoms (Angold & Costello, 1995; Harvard Medical School, 2002; Maier, Gansicke, Gater, Rezaki, Tiemens & Urzua, 1999; Robbins, Alessi, Cook, Pooznanski, & Yanchysyn; Strober, Green, & Carlson, cited in Poznanski & Mokros, 1994).

While largely unexplored in relation to actual performance on the DEPI, age was felt to be a potentially important moderating factor to include in the present study. Results of analysis set 2, however, were supportive of hypothesis 5 in that no significant differences in either DEPI or CDI were attributable to subject age for the combined
depressed-only and non-depressed clinical groups. These results, in and of themselves, do not clearly account for findings of significant difference observed in the first set of analyses (i.e., between “normals” and both clinical and non-clinical 15-year-olds for the DEPI, and between depressed vs. non-depressed clinical 16-year-olds for the CDI). It remains evident, based on these findings, that age is not a clear moderator of the ability of the DEPI and CDI to accurately classify individuals as depressed.

**Diagnostic Efficiency of the DEPI and CDI**

**DEPI**

The third set of analyses investigating the psychometric efficiency of the DEPI resulted in a fail to accept decision for hypothesis 6. In other words, results clearly indicated that the DEPI demonstrated inadequate diagnostic efficiency, compared with coefficients reported by Ganellen’s (1996) adult study. Although conducted with adult subjects, Gannellen’s (1996) study contains the most complete overview of DEPI diagnostic efficiency data to date, and was therefore selected to serve as a basis of comparison with results of the current study. Ganellen’s study, based on Exner’s (1993) standardization sample, reports the following coefficients pertaining to DEPI diagnostic efficiency: sensitivity = .75, specificity = .81, positive predictive power = .79, negative predictive power = .76, false positive rate = .23, false negative rate = .25, and overall correct classification = .79. Kappa coefficients were not reported for Ganellen’s (1996) study.

Exner’s (1993) data, summarized in Ganellen’s (1996) study, present optimistic results pertaining to the DEPI’s diagnostic efficiency. Although based on adult populations, these findings suggest that by using Exner’s (1993) recommended cutoff scores the DEPI is able to correctly identify (i.e., sensitivity) 75% of all subjects diagnosed with depression. Taking account the proportion of all subjects actually diagnosed with depression within the utilized sample, the DEPI was purportedly able to identify a slightly increased rate of 79%, based on recommended cutoff scores (i.e., positive predictive power). Conversely, Ganellen’s (1996) study reports that the DEPI was able to screen out 81% of actual non-depressed adult subjects based on sub-clinical cutoff scores (i.e., specificity), and that this proportion dropped minimally (i.e., 76%)
when the actual proportion of non-depressed adults within the sample was taken into account (i.e., negative predictive power).

Ganellen’s (1996) findings also report that the DEPI was able to discriminate between depressed vs. non-depressed adults with a reasonable degree of accuracy. That is to say, based on the results of the study, there was a 23% probability of misclassifying non-categorically depressed subjects as depressed (i.e., false positive rate), a 25% probability of misclassifying categorically depressed subjects as non-depressed (false negative rate), and a 21% probability of misclassifying all subjects (i.e., both depressed and non-depressed), given their categorical diagnoses (i.e., overall correct classification).

It bears mentioning that, in comparing Ganellen’s study with the current one, some methodological differences were found to exist. For one, the current study examined several different cutoff-point criteria in evaluating the DEPI’s diagnostic efficiency, whereas Ganellen relied exclusively on Exner’s (1993) recommended criteria of DEPI >5 to determine “correct” vs. “incorrect” classification of subjects as depressed without consideration of more or less stringent cutoffs (i.e., DEPI >4, >6, or >7). Further, Ganellen did not include the CDI in his analyses, and reported that diagnoses administered in Exner’s (1993) adult data set were based on DSM-III (APA, 1980), not DSM-IV (APA, 1994) depression criteria.

A comparison of Ganellen’s (1996) adult study with the current adolescent study, considering only the cutoff point of DEPI >5, indicates that the results of the current study differed considerably in almost all areas. Most notably, the DEPI demonstrated much lower sensitivity for the adolescent group (.17) than the adult study (.75). Although specificity (.90) and false positive rate (.10) in the adolescent study were slightly improved over Ganellen’s (1996) adult-based findings, adolescent false negative rate (.83) and overall correct classification (.55) were substantially worsened. Further, positive predictive power (.63) and negative predictive power (.54) were also found to be notably lower than that observed in the adult study. For the DEPI alone, the cutoff of >5 resulted in the highest overall classification rate (.55), highest Kappa coefficient (.078), and substantially lower false positive rate (.10, compared with .42 for the cutoff of DEPI ≥ 4).
CDI

The influence of the CDI on CS diagnostic efficiency was not evaluated in Ganellen’s study and, as such, the current study constitutes a confirmatory analysis of the limited results reported by Exner and Weiner (1995) in their normative CS adolescent data. Most notably, Exner and Weiner (1995) have proposed that the CS can predict adolescent depressive diagnoses with 18% accuracy on the basis of CDI total score cutoffs, although no further diagnostic efficiency coefficients were published. It is reasonable to assume that the authors utilized the recommended clinical cutoff of CDI $>4$ as the basis for their prediction. Results of the current study, consistent with that cutoff, yielded sensitivity (.42) and positive predictive power (PPP; .46) slightly higher than that anticipated by the adolescent normative sample; however, as was the case for the DEPI, this occurred at a cost of substantially lowered overall correct classification (OCC; .47), false positive and negative rate (FPR, FNR; .48, .58), and extremely low Kappa (-.060).

Combined DEPI and CDI

It is also discouraging that the current results indicated that the CDI did not contribute meaningfully towards enhancing the diagnostic efficacy of the DEPI, as was predicted in hypothesis 6 of the current study. The current findings were consistent in that, regardless of the cutoff criteria used, addition of the CDI to DEPI total scores resulted in no appreciable improvement in combined sensitivity and positive predictive power. These findings do not support Exner’s (1993) assertion that the CDI increases CS predictive validity of 79% of clinically depressed subjects exhibiting false negative results based only the DEPI.

Similar to results observed pertaining solely to the DEPI, no combination of DEPI and CDI total score cutoff points was able to approximate the higher levels of diagnostic efficiency reported by Ganellen (1996), pertaining to the DEPI alone. It is noted that numerous permutations of DEPI and CDI total score cutoffs resulted in significant increases in diagnostic efficiency in some areas, but consistently resulted in sacrifices in other areas. For example, use of increasingly stringent cutoff criteria (i.e., DEPI $>6$, CDI $>3$) resulted in significant improvement in specificity, positive predictive power, and false positive rates above those reported by Ganellen; however, application of these
criteria did not improve on modest negative predictive power rate and overall correct classification rate (i.e., both around .50), and resulted in extremely low sensitivity (i.e., between .03 to .007).

**Comparison with other research**

The current study’s findings differed greatly from those reported by numerous other research studies. Archer and Krishnamurthy (1997), for example, concluded that the CS yielded acceptable Kappa coefficients (between .66 to .98) for all CS variables in an adolescent-based research study. One of the more salient results derived from the current analysis set three is that Kappa coefficients describing the degree of agreement between DEPI, CDI, and combined indices for the current clinical adolescent populations were consistently low, regardless of cutoff criteria. Although no studies addressing the DEPI have provided a full range of diagnostic efficiency coefficients for children and adolescents, two studies have given partial information.

Exner (1990) has reported that the first version of the DEPI demonstrated 10%-30% false positive rates and greater than 60% false negative rates among adult samples, based on comparison with DSM-III (APA, 1980) criteria. Exner and Weiner (1995) also report that 69 out of 100 clinically depressed adolescents were correctly identified by clinically suggested DEPI cutoff scores, based on preliminary studies conducted in 1990. The poor psychometric efficiency of the first version of the DEPI led to construction of the current, revised DEPI, which Exner (1990; 1993) asserts demonstrates higher discriminant validity; however, Exner and Weiner (1995) also claim that the revised DEPI is able to correctly classify categorically diagnosed depressed adolescents with 50% accuracy on the basis of DEPI total score cutoffs, which does not appear to be an improvement over the previous DEPI.

Findings from the current study are not consistent with the research involving the older DEPI, but are quite consistent with specific reports pertaining to the revised version. Exner and Weiner (1995) have also stated that the Rorschach CS is able to correctly classify categorically diagnosed depressed adolescents with 25% accuracy on the basis of co-joint DEPI and CDI total score cutoffs. As previously discussed, results
of the current study revealed that the CDI did not contribute meaningfully towards increasing any of the psychometric efficiency coefficients for the DEPI.

Exner (1993) has also reported that the DEPI has demonstrated a false positive rate (FPR) of less than 2%, pertaining to a large sample of non-clinical children (Exner, 1993). Unfortunately, this information may or may not be clinically useful in the absence of further diagnostic efficiency information that was not included in Exner’s study. In the current study, for example, FPR reached as high as 0.1 or 0.0 for the DEPI based on the cutoff point of \( \geq 6 \) or \( 7 \); however, corresponding sacrifices in false negative rate (.96, .99) and sensitivity (.04, .01) occurred. Further, only one adolescent in the present study produced a DEPI score of 7, and only five produced scores of 6, revealing a highly skewed distribution limiting the practical significance of the results. It would be useful to inspect the distribution of Exner’s (1993) child sample to verify the practical significance of those findings, as well.

In light of the multiple, discrepant findings reported by the aforementioned data sources, it is noteworthy that the current results are roughly consistent with the diagnostic efficiency data reported as part of Exner and Weiner’s (1995) adolescent normative study for the revised DEPI. Although information provided by Exner and Weiner was highly limited in scope, it is interesting to find that results pertaining to adolescents in both their study and the current research produce consistently lower efficiency coefficients than those reported in Ganellen’s (1996) adult study. While numerous possible explanations exist for these differences it appears likely, based on the aforementioned results and review of the literature, that differences in diagnostic efficiency may be attributable to the influence of development (i.e., adolescent vs. adult), or differences in categorical classification associated with the change from DSM-III to DSM-IV diagnostic criteria.

**Incremental Validity of the DEPI and CDI**

The seventh hypothesis anticipating a significant incremental contribution of the DEPI and CDI towards prediction of categorical depression was not supported by the data. This finding calls for a fail to reject decision for the null hypothesis, namely that no significant incremental relationship was demonstrated by either DEPI or CDI total scores, in conjunction with MMPI-A Scale 2 scores and RADS total scores, towards the
maximum likelihood of prediction of depression for the adolescent population. While this finding is disappointing given the hypothesized outcome and negative implications for the CS depression indices, the results nevertheless are consistent with the aforementioned findings of the current study pertaining to weak DEPI and CDI diagnostic efficiency, and the limited empirical research to date investigating the psychometric properties of the DEPI and CDI for adolescents. Given such dearth of research pertaining to the CS depression indices, the current incremental validity results are considered clinically and theoretically useful towards understanding the function and nature of the DEPI and CDI with this population.

**Clinical Implications**

Despite discouraging findings, the aforementioned data have far-reaching practical implications for clinical practice. The purpose of this section is to delineate the applicability of these findings to both the assessment of depressive symptoms and disorders, and to ongoing attempts to understand what constitutes appropriate (and inappropriate) use of the Rorschach CS. For one, it is felt necessary and responsible to respond to the ongoing controversy about the use of the CS as a dimensional depression instrument for adolescents on the basis of the current findings.

On the basis of the current study, results remain consistent and quite clear in advocating for considerable caution by clinicians in interpreting the DEPI and CDI as dimensional predictors of adolescent depression. Results of the current study consistently demonstrate that the CS depression indices, considered by total scores, demonstrate poor predictive and incremental validity, poor diagnostic efficiency, and as a result poor construct validity for clinical adolescents.

Of greatest practical significance remains the finding that, based on the most optimal cutoff point identified for the DEPI in the current study (DEPI >5), both the DEPI and CDI together resulted in an overall correct classification rate (i.e., classification of “truly” depressed adolescents as depressed, and “truly” non-depressed adolescents as not depressed) of only slightly better than chance (.55). The same cutoff point resulted in an excessively high false negative rate (83%), and sensitivity of only 17%. Given that the CS is often used in court/forensic settings (McCann, 1998) and psychological/
psychiatric assessment and treatment facilities (Weiner, 1997), the current findings call for a high degree of caution when interpreting adolescent DEPI and CDI results in the context of important life-altering decisions (i.e., forensic or academic evaluation and/or placement, parental custody, or medical/pharmaceutical treatment).

Kazdin’s (1987) assertion remains theoretically valid that no one dimensional assessment measure, administered by itself, is capable of assessing categorical disorders with 100% accuracy. To this end, numerous replicated studies have confirmed adequate psychometric stability and diagnostic efficiency of numerous other dimensional depression inventories such as the CDI (Kovacs, 1992), the Reynolds Adolescent Depression Scale (Reynolds, 1987), and the MMPI-A (Archer, 1997; Butcher et al., 1992). In the case of the CS depression indices, the current study has failed to provide evidence that the DEPI and CDI demonstrate sufficient incremental validity to warrant their use in this regard for adolescent populations.

Consideration should also be given to the manner and purpose for which the CS is administered to adolescent populations. Because the CS is able to simultaneously measure numerous different constructs related to personality as well as discrete psychopathology (i.e., thought disorders/perceptual organization, anxiety/trauma, coping skills deficits, etc.) (Blatt, 1990; Exner, 1988; Weiner, 1997), and because administering and scoring the CS is a time-consuming and labor-intensive process, it is likely that few clinicians rely on the CS solely as a dimensional form of assessment. Ample research (Blatt, 1990; Exner, 1988; Weiner, 1997, among others) has demonstrated that, independent of the construct of depression, the Rorschach remains a highly useful assessment tool due to its ability to measure multiple constructs.

Caution is also warranted in making an outright rejection of the DEPI and CDI as dimensional measures of depression for all individuals, given encouraging empirical research findings pertaining to these indices for non-adolescents (Archer and Krishnamurthy, 1997; Ganellen, 1996, 2001; Meyer & Archer, 2001; Spigelman and Spigelman, 1991; Viglione & Hilsenroth, 2001; Weiner, 2001). In addition, given the few studies that have attempted to replicate Exner’s (1993) and Exner and Weiner’s (1995) findings related to the diagnostic efficiency and predictive validity of the DEPI
and CDI for adolescents, the results of the current study should be considered preliminary at best.

A number of possible explanations exist to explain why the DEPI and CDI are demonstrating such poor predictive validity. It remains possible that the disappointing findings observed in the current study are largely attributable to differences in the manner in which adolescents as opposed to adults (and perhaps children) perceive the Rorschach inkblot cards cognitively, based on unique developmental variance. Further, as previously mentioned in the study, adolescents may be prone to experience depression and/or depressive symptoms in a manner more consistent with differing depressive typologies (i.e., cognitive, affective, hopeless, etc.) that may affect different individuals.

Correspondingly, the possibility remains that adolescents, as opposed to children or adults, tend to provide more defensive response sets pertaining to the Rorschach CS (i.e., whether deliberately or unintentionally), although this was not examined in the current study. As adolescent protocols with insufficient responses or high/low Lambda rate were automatically excluded from the current study on an a-priori basis, it remains uncertain what proportion of the clinical population sampled would have demonstrated such defensive or underreporting response patterns, overall.

If such were the case, the sample of convenience may not be truly representative of depressed and/or clinical populations of adolescents, assuming that adolescents are simply more self-conscious and/or guarded. This is not felt to be likely, however, given that adolescents used in the current study were also excluded based on excessively elevated MMPI-A scores (i.e., T-scores of 65 or higher on the VRIN, TRIN, L, and K scores, F scores of >110, and raw “cannot say” scores of >30), which would have been more likely to screen out atypical response patterns.

Recommendations For Future Research

Due to the pervasive lack of research addressing the use and utility of the CS depression indices with populations of children and adolescents, the current results should be considered largely preliminary, and should serve as a guide for future research. Overall, it is hoped that the current results, although somewhat disappointing, will serve to inform future research. For one, replication of the study to determine whether the CS
DEPI and CDI consistently demonstrate low predictive and incremental validity, diagnostic efficiency, and co-relation for both clinical and non-clinical groups of adolescents would provide further evidence useful in generalizing findings to similar populations.

It was noted that a very few number of studies have been conducted to derive DEPI and CDI data for by non-patient populations of adolescents, thus necessitating the use of Exner and Weiner’s (1996) normative demographic data as a comparison group in the current study. In light of the surprisingly and somewhat disappointingly high variability in DEPI and CDI scores evidenced by the clinical participants of the current study (i.e., numerous cases of excessively high scores for depressed adolescents and sub-clinical scores for non-depressed adolescents), it would be helpful to know whether or not results derived by non-patients would, similarly, yield excessively high scores.

Lack of confirmation of the proposed hypotheses in the current study also provided evidence that there remains a pervasively limited comprehension of the precise constructs measured/not measured by the DEPI and CDI. Very few studies have been performed to replicate Exner’s (1993) research exploring the factor structure of the DEPI and CDI in order to better understand the function and underlying constructs associated with these indices. In particular, given the CDI’s lack of contribution towards prediction of depression in the current study, further research is warranted to elucidate the role, function, and best clinical use of this scale, including clarification of its role in assessing depression and other symptoms and disorders. Such research would serve to bolster theoretical and empirical support for the continued use of the CDI, and could serve to increase its clinical utility.

As undertaken by Khouri and Greenway (1996) and Archer and Krishnamurthy (1997), further research is called for to investigate the convergent and discriminate validity and degree of co-relation of the DEPI and CDI and various different assessments, such as the MMPI-A Harris-Lingoes or Content scales. Such studies could possibly provide greater insight into underlying differences in both divergent assessment methodologies and various proposed typologies depression.

Given the lack of predictive validity demonstrated by the DEPI and CDI in the current study for adolescent populations, research investigating the respective
contributions of depression-related variables (i.e., isolation index, morbid responses, etc.) as has been undertaken in other studies (Archer & Krishnamurthy, 1997; Khouri & Greenway, 1996), would also be beneficial for a number of reasons. For one, contribution towards a better understanding of the function of these variables could lead towards construction of a more effective and precise depression scale to increase predictive power and efficiency for adolescent populations. Also, such an understanding could possibly contribute towards a better understanding of the differences in predictive, concurrent, and discriminate validity of the DEPI and CDI between adolescents, adults, and perhaps children.

As called for by past research (Exner, 1993; Exner and Weiner, 1995; Wenar and Curtis, 1991), continued research investigating the effect of age on the ability of the DEPI and CDI to classify depressed individuals using the DEPI and CDI warrants future research, particularly based on some of the current study’s findings. The sporadic distribution of DEPI and CDI scores of individuals of different ages in analysis set 1, for example, and the lack of effect of age on total scores in set 2 provide evidence that this effect remains largely unexplained. Also, current findings providing partial corroboration of weaker diagnostic efficiency for the DEPI associated with adolescent, as opposed to adult, populations calls for further clarification of these differences.

The current study was necessarily restricted to adolescent populations, due to limitations in the clinical population from which the current study’s variables were derived. In addition to adolescents, there remains a need to better understand the CS’s psychometric properties, particularly its diagnostic efficiency/accuracy, with respect to depressed children. Particularly in light of the dynamic developmental factors contributing to the unique manifestation of depressed symptoms for children (Cicchetti & Toth, 1998; Hammen & Rudolph, 1996; Weiss, Weisz, Politano, Carey, Nelson, & Finch, 1991; Worchel, Hughes, Hall, Stanton, Stanton, & Little, 1990), further research is clearly warranted. For one, the ability to validly predict categorical depressive disorders otherwise overlooked by assessment instruments with more overt face validity would likely entail greater clinical applications and perhaps hold important significance towards incremental validity.
As mentioned in the previous section, further research pertaining to the contribution of the DEPI and CDI towards predicting depression, as well as other categorical disorders, in conjunction with different types of test-taking response sets would also perhaps serve to elucidate the utility of these scales. For both adult and child/adolescent populations, for example, it would be clinically useful to investigate the relationship between fake-good and fake-bad responding, such as that identified by the MMPI-2/MMPI-A L, F, and F scales and performance on the CS depression indices. Correspondingly, it would be useful to determine the degree of relationship between overly-desirable and overly-undesirable response sets, as indicated by elevated validity indices on other tests, and proportion of CS Lambda (L) to further evaluate the utility of this scale. Overall, it is felt that these findings could provide considerable information and support to the CS, given the poor face validity of the Rorschach, and test the commonly-held belief that the Rorschach CS is useful with resistant or uncooperative subjects.

Finally, in light of the predictive validity differences that appear to exist between adult and adolescent populations, based on the current and other research studies, further study is warranted to investigate the relationship between the DEPI, the CDI, and divergent theoretical typologies of depression that may manifest themselves differently for individuals of different ages, cultures, or other demographic factors (i.e., race, sex, etc.). In particular, future research is called for involving assessment instruments that have been proven to validly and reliably measure discrete typologies of depression (i.e., cognitive, affective, hopelessness, etc.), particularly with clinically depressed populations of adolescents (as well as adults).

**Limitations of the Study**

Attempts were made to control for threats to the validity of this study; however, as in all studies, limitations were inherent that bear further mention. Most notable are threats to the generalizability of the study, given the criterion samples from which the current study data were derived. With the exception of the non-clinical sample, only adolescents between the ages of 13-17 who were admitted on an inpatient basis to the same psychiatric inpatient facility were considered for inclusion in the study. Further,
only adolescents meeting DSM-IV diagnostic criteria for the various disorders of interest to the current study were included as research subjects. From the onset of the current study, the decision was made to exclude other research subjects not fulfilling these criteria acknowledging that use of such a sample would limit the studies generalizability to like populations.

Factors related to the demographic composition of the sample subjects also serve to restrict generalizability. Most notably, these included the geographical area, racial/ethnic composition (largely Caucasian), and socioeconomic status of the subjects. It was noted that the sample appeared roughly equal with respect to gender composition, but that there was a slight overrepresentation of adolescents from divorced/separated parental households (40.6%, compared with 29.3% intact and 18.8% “other”). While this may be representative of clinical populations, this remains a factor of consideration in replicating the results of the current study.

The necessity of relying on Exner’s (Exner and Weiner, 1995) CS normative sample of adolescents constituted a considerable limitation in the current study. Pertaining to generalizability and reproducibility, the conditions under which said data was obtained are highly questionable, given that Exner and Weiner omitted many important considerations from their research methodology. While information used in the study could be replicated in a similar manner as that of the current study by consulting Exner and Weiner’s (1995) published data, it remains a reality that without much of the omitted information (i.e., geographic location, ethnicity, demographic information pertaining to family history and mental health status) their results cannot be replicated. It is also regrettable that Exner and Weiner chose not to report the full range of sub-clinical DEPI and CDI total scores from their research publication, which could have contributed in a more meaningful way to the current study’s analyses.

Pertaining to confirmatory findings, particularly those related to the CDI, the current research study was at times limited to comparisons with unrepresentative data sets. For example, it would have been extremely helpful to have been able to compare the full range of diagnostic efficiency coefficients for Exner and Weiner’s (1995) adolescent normative sample with that of the current study, given the similarity of the current data to the limited information reported by the authors. Due to the lack of that
data, the current study was limited to comparing the results of the current adolescents with Ganellen’s (1996) adult clinical sample, which was based on DSM-III categorical criteria. As a result, it remains uncertain what proportion of the variance of the difference in results was attributable to changes in categorical criteria associated with the use of the DSM-IV to diagnose depressive disorders.

Threats to the reproducibility inherent in the current study also bear mentioning. For one, the archival data was generated from a sample of convenience based on one inpatient psychiatric facility in which multiple examiners, diagnosticians, and mental health personnel contributed to data collection. While reproduction of the collection of archival data would be possible, reproduction of the archival data, itself, is not due to a number of factors. For one, it was not possible to oversee the presentation, scoring, and recording of results of every test administered to the adolescent population of interest. As a result, it would not be possible to ensure exact reproduction of the results used in the current study.

The involvement of multiple individuals, and the fact that many of those individuals were students, also limits the reproducibility of the current study due to the increased probability of error; however, it was planned from the onset of the study that having one, consistent clinical supervisor who would oversee all individuals would control for the effect of this error to the highest extent possible. Further, it was not possible to verify first-hand the clinical interview process by which diagnoses were administered by psychiatrists to the population of interest; however, it is known that diagnoses were administered based on DSM-IV categorical criteria on the basis of subject and/or parent report.

Although the current study sought to corroborate Exner’s (1993) published findings relating to the Comprehensive System’s normative and standardization samples, in a sense it also constituted an exploratory effort, given the lack of research addressing the psychometric utility of the DEPI and CDI to date. In light of this, the results of the current investigation should be considered preliminary, and should serve to inform future research rather than be considered definitive results. Replication of the study resulting in findings that support those of this study would provide further evidence useful in generalizing findings to similar populations.
A further limitation consists of the manner by which diagnoses of depression were administered. As recommended by Kazdin (1987), the most valid manner in which depression can be diagnosed consists of a combination of both categorical and dimensional methodology. It is noted that diagnoses of depression were administered based entirely on categorical methodology in the current study. Although this method represents the most common method by which inpatients are administered DSM-IV diagnoses (Kazdin 1987, 1989), this is not ideal for research purposes and does constitute a limitation in research methodology.

**Conclusions**

Findings of the current study provide corroborating evidence supporting arguments that use of CS depression indices’ total scores does not provide the most useful and efficient means by which to assess the construct of depression for adolescent populations. Specifically, findings indicative of poor predictive validity and diagnostic efficiency call for caution, as well as further consideration and research, pertaining to the use of the DEPI and CDI in clinical practice. These findings remain applicable to clinical populations of adolescents, and pertain to the use of the Rorschach CS depression indices as a dimensional means of assessing the construct of depression.

The current study also elucidates that fact that despite considerable advancements in understanding and improving categorical taxonomy over the past several decades, the construct of depression remains a complex phenomenon that is difficult to adequately and holistically conceptualize and explain. Correspondingly, depression is also difficult to consistently and efficiently assess. This difficulty stems from and also is compounded by factors relating to the practice of categorical and dimensional assessment, differences in conceptualization of the construct of depression, and also by the need to explain complex developmental factors contributing to unique manifestation of depressive symptoms across the lifespan.

Given the significant implications and outcomes of assessment-based decisions, and the increasing need for accurate and expedient assessment mandated by managed care organizations, the need for increasingly cost-efficient, valid, and reliable assessment is becoming increasingly clear. As such, the burden of responsibility falls to the mental
health practitioner, informed by research, to select instruments that validly and reliably assess the construct they are designed to measure.

The CS depression indices represent one of the most commonly used, yet least-often researched dimensional assessment tool of depression, particularly with regard to its use with non-adult populations. In light of this, results of the current investigation underscore the need for better understanding of this instrument and its many components, particularly those related to depressive symptoms and disorders. In particular, the study provides promising preliminary, albeit limited, conceptual support for a relationship between age/development and the DEPI and CDI for clinical adolescents, compared with clinical populations of adults. Greater understanding of this relationship through future research would provide valuable insights needed to further mental health practitioners’ understanding of the relationship between the CS and depression, and the validity of this instrument for clinical and non-clinical adolescent populations. Further, such understanding could lend valuable support and insight towards understanding the construct of depression, and improving the practice of assessment.
APPENDIX A: Cognitive Information Processing Memory Constructs

Directed Attention

Short-term Immediate Memory

Auditory Verbal  Visual Nonverbal

Working Memory

Learning

Long-term Memory

Declarative

Episodic Events  Semantic Facts

Procedural

Skill Learning  Classical Conditioning

Retrieval

Free Recall  Recognition Recall

Source: Cohen, 1997
APPENDIX B: Steps and Variables Involved in Scoring the Rorschach CS

I. Location
   A. \textit{W=Whole} (subject uses the entire inkblot in percept)
   B. \textit{D=Common Detail} (frequently identified inkblot area)
   C. \textit{Dd=Unusual Detail} (infrequently identified inkblot area)
   D. \textit{S=Space} (white space area is use in response)

II. Developmental Quality
   A. \textit{+ : Synthesized Response} (Two or more objects, with form demand, listed as separate but related)
   B. \textit{o : Ordinary Response} (inkblot area identified as a single object with form demand)
   C. \textit{v/+ : Synthesized Response} (two or more objects, without form demand, listed as separate but related)
   D. \textit{v : Vague Response} (one object listed without specific form demand)

III. Determinants
   A. Form (F): responses based exclusively on the form features of the inkblot
   B. Movement (active vs. passive, or \textit{a} vs. \textit{p})
      1. \textit{M: Human Movement} (responses involve human movement, or animal or fictional character in human-like activity)
      2. \textit{FM: Animal Movement} (responses involve movement of animals that is congruent to their species)
      3. \textit{m: Inanimate Movement} (responses involving movement of inanimate, inorganic, or insensate objects)
   C. Chromatic Color
      1. \textit{C: Pure color} (responses based exclusively on the color of the inkblot without reference to form)
      2. \textit{CF: Color-form} (responses based primarily on inkblot color, yet also on form)
3. **FC: Form-color** (responses based primarily on inkblot form, yet also on color)

4. **Cn: Color naming** (colors on the inkblot are intentionally identified exclusively by color name)

D. **Achromatic Color**
   1. **C’: Pure achromatic color** (responses based exclusively on gray, black, or white features of the inkblot, clearly used as color)
   2. **C’F: Achromatic color-form** (responses based primarily on black, gray, or white features, yet also include form)
   3. **F’C: Form-achromatic color** (responses based primarily on form, yet also include black, gray, or white features)

E. **Shading-Texture**
   1. **T: Pure texture** (responses interpreted exclusively to represent tactual phenomena, and do not include form)
   2. **TF: Texture-form** (responses primarily interpreted as tactual phenomena, but also includes form)
   3. **FT: Form-texture** (responses made primarily on the basis of form, yet also interpreted as tactual phenomena)

F. **Shading-Dimension**
   1. **V: Pure vista** (responses interpreted as depth, based exclusively based on shading features)
   2. **VF: Vista-form** (responses interpreted as depth, based primarily on shading features, yet form is involved)
   3. **FV: Form-vista** (responses based primarily on form, yet also interpreted as depth)

G. **Shading-Diffuse**
   1. **Y: Pure shading** (responses based exclusively on light-dark features of inkblot)
   2. **YF: Shading form** (responses based primarily on light-dark features of inkblot, yet form is involved)
3. **FY:** Form-shading (responses based primarily on form, yet light-dark features of inkblot are involved)

H. **FD:** Form Dimension (responses in which depth, distance, or dimensionality is expressed using size and/or shape of inkblot contours, without the use of shading.

I. **Pairs & Reflections**
   1. **(2):** pair (response in which two identical objects are expressed, based on the inkblot symmetry)
   2. **rF: Reflection-form** (response in which
   3. **Fr: Form-reflection** (response in which the percept is reported as a reflection, yet has a specific form demand)

IV. **Form Quality**
   A. **+: Superior-Overelaborated** (articulated response is unusually detailed and appropriate to inkblot contours)
   B. **o: Ordinary** (response is common without unusual enrichment)
   C. **u: Unusual** (response is not often reported, yet inkblot contours used in the percept are not unreasonably identified)
   D. **- : Minus** (response represents distorted, arbitrary, and unrealistic use of form)

V. **Contents and Populars**
   A. **Contents:** Discrete categories include Whole or Fictional/Mythological Human, Human or Fictional/Mythological Human Detail, Human Experience, Whole or Fictional/Mythological Animal, Animal or Fictional/Mythological Detail, Anatomy, Art, Anthropology, Blood, Botany, Clothing, Clouds, Explosion, Fire, Food, Geography, Household, Landscape, Nature, Science, Sex, X-ray, and Idiographic.
   B. **Popular (p):** responses that occur with unusually high frequency (one in three protocols)

VI. **Organizational Activity**
   A. **ZW:** Whole response with a DQ coding of +, v/+, or o.
B. **ZA**: response in which two or more separate, yet touching, objects are reported in meaningful relationship.

C. **ZD**: response in which two or more separate, non-adjacent, objects are reported in meaningful relationship.

D. **ZS**: response in which white space is integrated with other inkblot areas.

### VII. Special Scores

A. Deviant Verbalizations (DV, severity level 1 or 2): verbalized content is particularly odd or inappropriate.
   1. **Neologism**: incorrect word is substituted in place of a correct word
   2. **Redundancy**: subject identifies twice the nature of the object(s) described

B. Deviant Responses (DR, severity level 1 or 2): responses reflect a tendency to detach from or distort the task at hand.
   1. **Inappropriate Phrases**: phrases are irrelevant or completely inappropriate to the task at hand
   2. **Circumstantial Responses**: fluid, rambling/circuitous responses in which the subject becomes elaborates extensively and inappropriately

C. Inappropriate Combinations: responses violate realistic expectations of percepts
   1. **Incongruous Combinations (INCOM)**: inkblot details are inappropriately merged into a single object
   2. **Fabulized Combination (FABCOM)**: response posits an implausible relationship between two or more objects
   3. **Contamination (CONTAM)**: two or more impressions violate reality by becoming fused into a single response

D. Inappropriate Logic (ALOG): responses reflect strained, unconventional, or illogical reasoning

E. Perseveration (PSV): marked preoccupation with particular inkblot aspects
   1. **Within Card Perseveration**: identical scoring is applied to consecutive responses to the same card
2. **Content Perseveration**: subject intentionally and overtly identifies a percept as being the same object viewed previously.

3. **Mechanical Perseveration**: subject mechanically reports the same object repeatedly, despite changing stimuli.

F. **Confabulation (CONFAB)**: subject attends only to one detail of the inkblot, yet generalizes the response to a larger area.

G. **Special Content Characteristics**: responses that reflect cognitive features or projected characteristics of the self.
   1. **Abstract Content (AB)**: responses clearly incorporating symbolic representation.
   2. **Aggressive Movement (AG)**: responses in which action described is clearly aggressive.
   3. **Cooperative Movement (COP)**: response in which two or more objects are interacting in positive, or cooperative ways.
   4. **Morbid Content (MOR)**: response contains depictions of objects as dead, destroyed, ruined, spoiled, damaged, injured, or broken, or description of dysphoric feelings or characteristics.

H. **Personalized Responses (PER)**: responses clearly suggestive of self-reference (i.e., personal knowledge or experience used to justify the response).

I. **Color Projection (CP)**: an achromatic inkblot area is identified as being chromatically colored.

Sources: Exner et al., 1995; Groth-Marnat, 1997
APPENDIX C: Comparison of Rorschach Depression Variables with Cognitive Symptoms

<table>
<thead>
<tr>
<th>Rorschach Variable</th>
<th>Specific Cognitive Ramifications</th>
<th>Salient Implications for Cognitive Theory</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEVELOPMENTAL QUALITY</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low # of W Responses</td>
<td>Distracting/irrelevant negative thoughts</td>
<td>B, RA</td>
</tr>
<tr>
<td>High # of D, Low # of D+ responses</td>
<td>Negative thoughts, poor overall adjustment</td>
<td>TM, LH, B, RA</td>
</tr>
<tr>
<td>High # of Dd responses</td>
<td>Rigid, dichotomous thinking</td>
<td>B, LH, RA</td>
</tr>
<tr>
<td>High # of S (3 or more)</td>
<td>Cognitive Triad</td>
<td>B, LH, RA</td>
</tr>
<tr>
<td>DETERMINANTS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High # of M&lt;sup&gt;p&lt;/sup&gt; responses</td>
<td>Extrinsic Locus of control</td>
<td>B, LH</td>
</tr>
<tr>
<td>Low # of M responses</td>
<td>Cog. Rigidity, extrinsic locus of control, cog. Triad (higher in adolescents)</td>
<td>B, RA, LH, TM</td>
</tr>
<tr>
<td>Low # of FM responses</td>
<td>Cog. triad, extrinsic locus of control, learned helplessness</td>
<td>B, LH</td>
</tr>
<tr>
<td>High # of m, OR High # of Y responses</td>
<td>Low stress tolerance, disorganization</td>
<td>RA, B, TM</td>
</tr>
<tr>
<td>Low # of C and CF</td>
<td>Overcontrolled emotions, negative thoughts</td>
<td>B, RA, RM</td>
</tr>
<tr>
<td>High # of C’</td>
<td>Cog. Triad, pervasive negative thoughts, ext. locus of control, generalized dep.</td>
<td>B, LH, RA, TM</td>
</tr>
<tr>
<td>High # of T or TF responses</td>
<td>Interpersonal dependence, neg. view of self, ext. LOC</td>
<td>LH, B</td>
</tr>
<tr>
<td>Absence of or very low T</td>
<td>Emotional impoverishment</td>
<td>B, LH</td>
</tr>
<tr>
<td>High # of pure V responses</td>
<td>Self-critical thoughts, neg. view of self</td>
<td>B, LH</td>
</tr>
<tr>
<td>FORM QUALITY</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High # of u and/or - responses</td>
<td>High proportion of irrelevant/ distracting thoughts</td>
<td>B, RA</td>
</tr>
<tr>
<td>ORGANIZATIONAL ACTIVITY (Z)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High # of Zf (&gt;13)</td>
<td>High level of cog. Energy expended to organize perceptions</td>
<td>RA</td>
</tr>
<tr>
<td>Low # of Zf (&gt;13)</td>
<td>Low level of cog. Energy expended to organize perceptions</td>
<td>RA</td>
</tr>
<tr>
<td>SPECIFIC CONTENT SCORES</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------</td>
<td>-------</td>
</tr>
<tr>
<td>Low # of H responses</td>
<td>Social withdrawal, possibly negative view of others</td>
<td>B, LH</td>
</tr>
<tr>
<td>High # of A and Ad responses</td>
<td>Low creativity, predictable style of approaching life</td>
<td>LH, RA</td>
</tr>
<tr>
<td>High # of popular (P) or Food responses</td>
<td>Approval-seeking, overconforming, ext. LOC</td>
<td>B, LH, RA</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SPECIAL SCORES</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Presence of a perseveration score</td>
<td>Cognitive rigidity</td>
<td>B, LH, RA</td>
</tr>
<tr>
<td>Absence of or few Cooperative and Aggressive responses</td>
<td>Negative view of others/interaction, low interpersonal skill</td>
<td>B, LH,</td>
</tr>
<tr>
<td>Presence of Morbid responses</td>
<td>Cognitive triad, view of self as “damaged goods”</td>
<td>B, LH, RA, TM</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>RATIOS, PERCENTAGES AND DERIVATIONS</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Low # of responses (R)</td>
<td>Generalized depression, defensiveness</td>
<td>LH, RA</td>
</tr>
<tr>
<td>Low Affective Ratio (Afr)</td>
<td>Withdrawal from emotions/affect</td>
<td>B, TM</td>
</tr>
<tr>
<td>Low Egocentricity Index (EI) Ratio: (3r + 2)/R</td>
<td>Low Self-Esteem</td>
<td>B, LH, TM</td>
</tr>
<tr>
<td>High Lambda (pure F: non-pure F)</td>
<td>Generalized depression, self-guilt, cognitive triad</td>
<td>B, LH, TM</td>
</tr>
<tr>
<td>Uneven Experience Base (EB), and high D</td>
<td>Poor stress management, cognitive triad</td>
<td>B, LH, RA</td>
</tr>
<tr>
<td>Low Experience Actual (EA) or Adj. D</td>
<td>Low ability to use cognition and affect to organize experiences/stimuli</td>
<td>LH, RA</td>
</tr>
<tr>
<td>Elevated Experienced Stimulation (ES)</td>
<td>Low frustration tolerance, disorganization, sense of helplessness</td>
<td>B, LH, RA</td>
</tr>
<tr>
<td>Low Active, High Passive responses</td>
<td>Passive cognitive and behavioral proclivities</td>
<td>LH</td>
</tr>
<tr>
<td>High ratio of Isolation Index:R</td>
<td>Social withdrawal/alienation, limited interest in people</td>
<td>B, LH</td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>Key:</th>
<th>Symbol</th>
<th>Theory</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td></td>
<td>Beck’s Theory</td>
</tr>
<tr>
<td>LH</td>
<td></td>
<td>Learned Helplessness Model</td>
</tr>
<tr>
<td>RA</td>
<td></td>
<td>Resource Allocation Model</td>
</tr>
<tr>
<td>TM</td>
<td></td>
<td>Tripartite Model</td>
</tr>
</tbody>
</table>
REFERENCES


BIOGRAPHICAL SKETCH

Robert Van Noord was born and raised in Michigan, where he completed his undergraduate course of study at Calvin College in Grand Rapids. After earning a Bachelor of Arts degree with dual majors in Psychology and Spanish, he worked for several years in the field of Social Work providing case management services to adults with severe and persistent mental illnesses. Robert became enrolled in Florida State University’s College of Education in 1997, aspiring to earn the degree of Education Specialist and to obtain qualification as a School Psychologist. After his first year, he successfully applied to and was accepted into the combined Counseling Psychology and School Psychology doctoral program, the requirements of which he successfully completed in December, 2005. Robert completed his pre-doctoral internship through the Medical College of Ohio in Toledo, Ohio in the fall of 2003, since which time he married and moved with his wife to the Twin Cities area of Minnesota. Robert is proficient in Spanish, and currently works providing mental health services (i.e., individual counseling, psychological testing, consultation, etc.) predominantly to members of the Twin Cities’ Latino population.