CONFIRMING THE CONSTRUCTS OF THE CHILD INTERPERSONAL RELATIONSHIP AND ATTITUDES ASSESSMENT

Chiao-Feng Chung, B.S., M.Ed.

Dissertation Prepared for the Degree of

DOCTOR OF PHILOSOPHY

UNIVERSITY OF NORTH TEXAS

August 2013

APPROVED:

Dee C. Ray, Major Professor
Sue C. Bratton, Committee Member
Natalya Lindo, Committee Member
Leslie De Jones, Committee Member
Jan Holden, Chair of the Department of Counseling and Higher Education
Jerry R. Thomas, Dean of the College of Education
Mark Wardell, Dean of the Toulouse Graduate School
Chung, Chiao-Feng. Confirming the Constructs of the Child Interpersonal Relationship and Attitudes Assessment. Doctor of Philosophy (Counseling), August 2013, 122 pp., 14 tables, 2 figures, references, 65 titles.

The purpose of this study was to confirm the four-factor model of the Child Interpersonal Relationship and Attitudes Assessment (CIRAA) in order to establish the instrument’s factor/structure validity using a sample different than that used in instrument development. The CIRAA was the first parent-report instrument based on child-centered theory and designed to measure play therapy outcomes. Its four factors are Self-Regulation (formerly Self-Control), Interpersonal Relationships, Coping Skills, and Internal Locus of Evaluation.

For this study, the CIRAA was administered to 206 parents; their children were 75 females and 131 males aged 3 to 10 years old. The distributions of children’s genders and ages in this study were similar to the distributions of the sample used to develop the CIRAA.

Based on confirmatory factor analysis results of overall goodness-of-fit indices; localized areas of strain; and interpretability, size, and statistical significance ($p < .001$) of the model's parameter estimate, the four-factor model of the CIRAA was confirmed with both theoretical and empirical support. Internal consistency reliability for the subscales and total score were acceptable, with an overall reliability coefficient of .928. A medium negative correlation ($r = -.417, n = 47, p < .01$) was found between the CIRAA total scores and Child Behavior Checklist Total Problems scores, a result in the expected direction. All subscales and total scores, except the Internal Locus of Evaluation score, distinguished differences between non-clinical and clinical samples. Altogether, results indicated that the CIRAA is appropriate for making clinical decisions about individuals. Limitations, additional implications, and future research are discussed.
Copyright 2013

by

Chiao-Feng Chung
ACKNOWLEDGEMENTS

I would not become who I am and be where I am without help from my co-travelers who have offered me comfort, support, encouragement, understanding, and sometimes challenged me in different ways throughout this process.

Mom and Dad, I would never ever be able to accomplish what I have accomplished without you. I know you were worried to let me come to the U.S.A seven years ago. Yet, you still decided to support and encourage me to pursue my dream. I trust myself and have courage to conquer any obstacles because of your trust in me. Thank you for loving me unconditionally! My little brother, I have always cherished the relationship we have. We are always there for each other. Even though you don’t say this often, I know you are proud of me just like I am so very proud of you. Grandmom,

奶奶，我知道您總是為我禱告。儘管不喜歡我們的分離，您始終為我的成就感到開心和驕傲。

My major professor and mentor, Dr. Dee Ray, I feel fortunate to have known you, been guided by you, and learned from you. You have challenged me in various ways that have helped me grow. I appreciate your openness to my feedback and willingness to listen. My committee members, Dr. Sue Bratton, Dr. Natalya Lindo, and Dr. Leslie Jones, I appreciate your feedback, support, encouragement, and care for me.

Last but not the least, I want to thanks my dear friends. David and Kasie, I will miss our time in your office. I have always felt comforted and reenergized to keep going after talking to you. Yules and Terrbear, you have always believed in me and never doubted me even when I doubted myself. Lizer and Dina, you cannot imagine how staying up late at the clinic with you made finishing this dissertation possible. Thank you, everyone!!

iii
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>ACKNOWLEDGEMENTS</th>
<th>iii</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF TABLES</td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>viii</td>
</tr>
<tr>
<td><strong>CHAPTER 1 INTRODUCTION</strong></td>
<td>1</td>
</tr>
<tr>
<td>Statement of the Problem</td>
<td>3</td>
</tr>
<tr>
<td>Purpose of the Study</td>
<td>4</td>
</tr>
<tr>
<td><strong>CHAPTER 2 REVIEW OF LITERATURE</strong></td>
<td>5</td>
</tr>
<tr>
<td>Child-Centered Play Therapy</td>
<td>5</td>
</tr>
<tr>
<td>Assessment and Evidence-Based Research</td>
<td>8</td>
</tr>
<tr>
<td>Assessment in Play Therapy Research</td>
<td>9</td>
</tr>
<tr>
<td>Instruments Used in Play Therapy Research</td>
<td>11</td>
</tr>
<tr>
<td>Child Behavior Checklist</td>
<td>14</td>
</tr>
<tr>
<td>Parenting Stress Index</td>
<td>16</td>
</tr>
<tr>
<td>Other Instruments Used in Measurement of CCPT</td>
<td>19</td>
</tr>
<tr>
<td>Limitations of Assessments Used for CCPT</td>
<td>20</td>
</tr>
<tr>
<td>Child Interpersonal Relationship and Attitudes Assessment</td>
<td>21</td>
</tr>
<tr>
<td>Process of Developing the CIRAA</td>
<td>22</td>
</tr>
<tr>
<td>Outcome of the CIRAA Development</td>
<td>27</td>
</tr>
<tr>
<td>Strengths of CIRAA</td>
<td>30</td>
</tr>
<tr>
<td>Identified and Defined Factors on the CIRAA</td>
<td>32</td>
</tr>
<tr>
<td>Self-Control</td>
<td>32</td>
</tr>
<tr>
<td>Interpersonal Relationship</td>
<td>35</td>
</tr>
<tr>
<td>Coping Skills</td>
<td>39</td>
</tr>
<tr>
<td>Internal Locus of Evaluation</td>
<td>41</td>
</tr>
<tr>
<td>Confirmatory Factor Analysis and the Assessment of Model Fit</td>
<td>44</td>
</tr>
<tr>
<td>Conclusion</td>
<td>48</td>
</tr>
<tr>
<td><strong>CHAPTER 3 METHODS AND PROCEDURES</strong></td>
<td>50</td>
</tr>
<tr>
<td>Research Questions</td>
<td>50</td>
</tr>
</tbody>
</table>
PARTICIPANTS ........................................................................................................................ 50

INSTRUMENT ......................................................................................................................... 53

CIRAA .................................................................................................................. 53

Child Behavior Checklist (CBCL) ........................................................................ 54

DATA COLLECTION PROCEDURES .............................................................................. 55

Clinical Sample Procedures .................................................................................. 55

Non-Clinical Sample Procedures .......................................................................... 55

DATA ANALYSES ................................................................................................................... 56

Confirmatory Factor Analysis ............................................................................... 56

Multiple Regressions ............................................................................................ 61

Independent-Sample t-Tests ................................................................................ 62

CHAPTER 4 RESULTS ............................................................................................................... 64

Research Question 1: Will CFA Support the Factor Structure of CIRAA (Self-Control, Interpersonal Relationship, Coping Skills, and Internal Locus of Evaluation) Previously Defined by Original EFA? .............................................................. 64

Model 1: Uncorrelated Four Factor Solution ........................................................ 66

Model 2: Correlated Four Factor Solution ............................................................ 71

Research Question 2: How are Children's Characteristics, if any, such as Age and Gender, related to Factors and Total Score on the CIRAA? ......................................................... 71

Research Question 3: Will Children from Non-Clinical and Clinical Samples Differ on the CIRAA Mean Subscale Score (Self-Control, Interpersonal Relationship, Coping Skills, and Internal Locus of Evaluation and Total Score? ......................................................... 76

Internal Consistency Reliability .............................................................................. 78

Concurrent Validity with Child Behavioral Checklist ............................................. 81

CHAPTER 5 DISCUSSION ......................................................................................................... 82

Validity ............................................................................................................................. 83

Reliability .......................................................................................................................... 84

Generalizability of Sample ...................................................................................... 84

Utility of Instrument ....................................................................................................... 85

Recommended Modifications to the Instrument ....................................................... 86

Limitations ......................................................................................................................... 87

Implications ....................................................................................................................... 89

Recommendations for Further Studies ........................................................................ 90
Conclusion ........................................................................................................................ 91

APPENDIX A INSTITUTIONAL REVIEW BOARD INFORMED CONSENT DOCUMENTS ....................................................................................................................................................... 93

APPENDIX B  LETTER TO PARENTS ..................................................................................... 97

APPENDIX C CIRAA .................................................................................................................. 99

REFERENCES ........................................................................................................................... 102
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 3.1</td>
<td>Demographic Characteristics of Participants (n = 202)</td>
<td>52</td>
</tr>
<tr>
<td>Table 3.2</td>
<td>CIRAA Items and Factors</td>
<td>57</td>
</tr>
<tr>
<td>Table 4.1</td>
<td>Parameter Estimate from the CFA Model 1</td>
<td>70</td>
</tr>
<tr>
<td>Table 4.2</td>
<td>Comparison of Goodness-of-Fit Indices of Two Models</td>
<td>71</td>
</tr>
<tr>
<td>Table 4.3</td>
<td>Summary of the Regression of Two Predictors and Self-Control</td>
<td>73</td>
</tr>
<tr>
<td>Table 4.4</td>
<td>Summary of the Regression of Two Predictors and Interpersonal Relationship</td>
<td>74</td>
</tr>
<tr>
<td>Table 4.5</td>
<td>Summary of the Regression of Two Predictors and Coping Skills</td>
<td>74</td>
</tr>
<tr>
<td>Table 4.6</td>
<td>Summary of the Regression of Two Predictors and Internal Locus of Evaluation</td>
<td>75</td>
</tr>
<tr>
<td>Table 4.7</td>
<td>Summary of the Regression of Two Predictors and Total Score</td>
<td>75</td>
</tr>
<tr>
<td>Table 4.8</td>
<td>Independent Sample t-Test on Four Subscales and Total Score on Two Groups</td>
<td>77</td>
</tr>
<tr>
<td>Table 4.9</td>
<td>Inter-item Correlations for CIRAA Factor 1: Self-Control</td>
<td>79</td>
</tr>
<tr>
<td>Table 4.10</td>
<td>Inter-item Correlations for CIRAA Factor 2: Interpersonal Relationship</td>
<td>79</td>
</tr>
<tr>
<td>Table 4.11</td>
<td>Inter-item Correlations for CIRAA Factor 3: Coping Skills</td>
<td>80</td>
</tr>
<tr>
<td>Table 4.12</td>
<td>Inter-item Correlations for CIRAA Factor 4: Internal Locus of Evaluation</td>
<td>80</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 3.1</td>
<td>CFA Model 1: Uncorrelated four factor solution</td>
<td>59</td>
</tr>
<tr>
<td>Figure 3.2</td>
<td>CFA Model 2: Correlated four factor solution</td>
<td>60</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

Ray (2011) stated that child-centered play therapy (CCPT) “...is the preferred and developmentally appropriate intervention for children who are experiencing emotional, behavioral, and developmental problems” (p. xi). In adapting Roger’s person-centered philosophy, Axline (1947) presented the non-directive play therapy approach for working with children. She described the concepts, philosophy, and therapist actions related to the work with children in her book, *Play Therapy* (1947). Subsequently, in 1991, Landreth integrated Axline’s philosophy of working with children and Ginott’s (1975) limit setting strategies and systematically defined and described child-centered play therapy in his book, *Play Therapy: The Art of the Relationship*. Landreth (2012) stated that CCPT “is an encompassing philosophy for living one’s life in relationships with children” (p. 53). The fundamental concept of child-centered play therapy is a deep and constant belief “in the capacity and resiliency of children to be constructively self-directing” (Landreth, 2012, p. 53). Ray (2009) developed the manual for standardizing CCPT use in order to encourage the development of CCPT research and exploration of its effectiveness. Currently, CCPT is the most widely-used treatment modality by professional play therapists (Ray, 2011). However, based on the criteria for empirically validated treatments (APA, 2006), play therapy is not yet an evidence-based intervention.

A crucial process in the development of any mental health intervention is to establish a body of evidence for its effectiveness. Therefore, one of the major tasks for researchers is the appraisal of their selected assessments’ abilities to measure the named constructs (APA, 2008; Heppner, Wampold, & Kivlighan, 2008). The validity of empirically-supported treatments (EST) is strengthened using measurement assessments with reasonable reliabilities and validities,
and those that truly measure the constructs of most interest (APA, 2006). Because CCPT is widely used, it is imperative to develop an instrument that can truly assess the objects of CCPT in order to move this discipline to an evidence-based intervention. Among the assessments used in CCPT research presented by Lin (2011), it is evident that most instruments were not developed for the specific use of assessing the effectiveness of play therapy. Moreover, most of these assessments that have been used in research attend to behavioral changes only, which do not align with CCPT's philosophy of emphasizing holistic and internal changes occurring within children (Holliman, 2010; Landreth, 2012). The assessments developed specifically for play therapy usually require trained observers who observe in-session behaviors of children. Hence, Holliman and Ray (2013) believed that developing a rigorous theoretically-based instrument for measuring the impact of intended outcomes in CCPT would move this discipline toward EST.

Holliman and Ray (2013) created and developed the Child Interpersonal Relationship and Attitudes Assessment (CIRAA) through five phases of assessment development, including literature review, parent focus group, expert review, expert focus group, and pilot administration (Springer, Abell, & Hudson, 2002). The CIRAA, the Child Behavior Checklist (CBCL), and the Parenting Stress Index (PSI) were administered to 136 parents of children aged 3 to 10 years. After conducting exploratory factor analysis, four factors were retained, and the factor solution explained 53.86% of the variance. Reliability coefficients for total and subscale scores were acceptable, with an overall Cronbach’s alpha reliability coefficient of .93. Concurrent validity among the CIRAA and CBCL, and CIRAA and PSI were reported with Pearson's $r$ of .75 and .74, respectively. ROC analysis demonstrated reasonable discrimination of adaptive scores and clinical scores, resulting in a clinical cut-off score of 2.5. Holliman and Ray (2013) named the four factors retained from the EFA as Self-Control, Interpersonal Relationship, Coping Skills,
and Internal Locus of Evaluation. These four factors were named in order to reflect content of items and align with objectives of CCPT (Landreth, 2012).

Holliman and Ray's (2013) suggested that the next step in validation of the CIRAA is to conduct confirmatory factor analysis (CFA) in order to strengthen the factor structure found in EFA. CFA is typically used when researchers already have theories in mind or want to evaluate instrument structures based on theoretical expectations (Brown, 2006). Following exploratory identification of factors through EFA with the CIRAA, completion of CFA will possibly provide verification of factor structure and credibility for the factor constructs.

Statement of the Problem

Instruments used to measure the effectiveness of CCPT do not typically align with the intended objectives of the intervention (Holliman, 2010; Holliman & Ray, 2013). Most assessment instruments focus more on children's external behaviors, yet CCPT emphasizes the importance of changes occurring in children.

One of the common issues related to measurement in evidence-based instruments is the lack of gold standard to evaluate the validation of the instrument (Kazdin, 2005). Additionally, because of the comorbidity and varying facets of children and adolescents’ presenting problems, no one instrument utilized in intervention research can truly capture the complexity of these issues. Due to a lack of agreed-upon standards, researchers may select assessments that do not truly measure the constructs of interest. In response to a lack of valid and reliable measures related to CCPT constructs of interest, Holliman and Ray (2013) developed the CIRAA, which aligns with CCPT objectives described by Landreth (2012). Although early exploration of
overall psychometrics and factor structure of the CIRAA are encouraging, Holliman and Ray (2013) indicated that further confirmation of CIRAA validity is needed.

Purpose of the Study

This study aimed to fulfill several goals. The first objective was to examine the factor structure and item-factor relationships found in Holliman's study (Brown, 2006; Holliman, 2010; Holliman & Ray, 2013) with a new sample using confirmatory factor analysis (CFA). Brown (2006) stated that exploratory factor analysis (EFA) is often used to develop a theory; and CFA is utilized to test the theory. The second objective was to explore the relationship between four factor and total score and children's characteristics, such as gender and age. This exploration helps to confirm the instrument's applicability across gender and age. The third objective was to assess if the ability of the CIRAA to differentiate between clinical and non-clinical samples.
CHAPTER 2

REVIEW OF LITERATURE

Child-Centered Play Therapy

Landreth (2012) stated child-centered play therapy (CCPT) “is an encompassing philosophy for living one’s life in relationships with children” (p. 53). The fundamental concept of CCPT is a deep and constant belief “in the capacity and resiliency of children to be constructively self-directing” (Landreth, 2012, p. 53). If children are granted freedom in the play therapy relationship to be themselves, children are capable of leading their own growth. In the playroom, play therapists not only believe deeply in the inner person of the child but also respect the direction decided by him or her. Because of this profound belief about children’s inner capabilities, “the play therapist’s objective is to relate to the child in ways that will release the child’s inner-directional, constructive, forward-moving, creative, self-healing power” (Landreth, 2012, pp. 53-54).

Before Landreth (2012) systematically defined and described CCPT in his book, Play Therapy: The Art of the Relationship, which was first published in 1991, several theorists had shared similar beliefs about children and play therapy. Axline (1947) described that “play is the child’s natural medium of self-expression” (p. 9). Therefore, play is a basic yet important activity in which children have engaged since they were born (Axline, 1947; Piaget; 1962). Children under 10 years of age gradually shift from manipulating materials to symbolic ways of playing to assist in their understanding the world (Piaget, 1962; Ray, 2011). Play offers children a vehicle for communication without requiring them to verbally describe their thoughts and feelings (Axline, 1947; Landreth, 2012; Ray, 2011).
Due to the intrinsic nature of play in children's lives, the use of play therapy as a type of treatment modality is suitable. Because children are not yet capable of abstract thinking and articulating their struggles clearly enough to gain support and assistance from therapists, play therapy offers children environments and opportunities to play out their worlds (Axline, 1974; Kottman, 2003; Landreth, 2012; Piaget, 1962). Scholars have developed different types of play therapy (Axline, 1974; Carroll & Oaklander, 1997; Freud, 1946; Klein, 1932/1975; Knell, 1997; Kottman, 2003; Landreth, 2012).

Axline (1947), one of Carl Rogers’ students and colleagues, adapted the person-centered approach when working with children, developing nondirective play therapy. Nondirective play therapy has several alternative names, including CCPT (Landreth, 2012). Many therapists, such as Moustakas (1959), Ginott (1961), and Landreth (2012) followed, modified, and popularized Axline’s nondirective play therapy (Lin, 2011). Axline (1947) believed that children use play to express their emotions and thoughts. Therefore, play therapy establishes a therapeutic environment allowing children to play as a natural medium for expressing themselves. Both Axline (1947) and Landreth (2012) highly regarded the importance of therapists’ relationships with children as part of communicating unconditional positive regard, empathy, and genuineness. Play therapists offer children a permissive environment in which children can grow and progress at their own pace (Axline, 1947; Landreth, 2012).

Axline (1947) proposed eight basic principles as guidance to play therapists when conducting play therapy with children. Axline’s principles were the following:

1. The therapist must develop a warm, friendly relationship with the child.
2. The therapist accepts the child exactly as he/she is.
3. The therapist develops a feeling of permissiveness so that child feels free to express feelings completely.
4. The therapist recognizes and reflects feelings (of a child) so that the child can gain insight into his/her behaviors.

5. The therapist respects that the child can solve his/her problems and believes that the responsibility to change rests on the child.

6. The therapist does not attempt to direct the child but lets the child lead the way as the therapist follows.

7. The therapist understands that therapy is a gradual process and does not rush the child.

8. The therapist establishes only those limitations that are necessary to anchor the therapy to the real world and to facilitate the child’s awareness for his/her responsibility in the relationship. (pp. 73-74)

Landreth (2012) adapted Axline’s (1947) nondirective strategies and Ginott’s (1975) limit setting techniques into a comprehensive and consistent approach. The approach requires trained play therapists to utilize developmentally appropriate materials in a nondirective fashion when working with children in the playroom (Holliman, 2010). Landreth incorporated some of the basic relationship-building skills, such as tracking, reflection of feelings, and reflection of content, proposed by Axline, and adapted the limit-setting philosophy advocated by Ginott to formulate his own child-centered approach.

Landreth (2012) emphasized that “permissiveness in the child-centered play therapy approach does not mean the acceptance of all behaviors” (p. 257). A relationship has little value without limits. Through the limit-setting process, children have opportunities to choose. Limits not only offer structure for the development of the therapeutic relationship but also assist with creating the therapeutic experience that is close to a real-life one (Landreth, 2012). Children learn to be responsible for themselves and their own well-being. Landreth (2012) believed that the limit setting process permits children to develop self-awareness, self-acceptance, self-control, self-responsibility, coping skills, and eventually self-growth (Landreth, 2012; Holliman, 2010). Landreth believed that the therapist’s main responsibility is to facilitate the therapeutic
relationship, providing an environment in which children are able to fully express themselves and experience the here and now.

Assessment and Evidence-Based Research

American Psychological Association Presidential Task Force on Evidence-Based Practices (APA, 2006) agreed that “Evidence-based practice in psychology (EBPP) is the integration of the best available research with clinical expertise in the context of patient characteristics, culture, and preferences” (p. 273). APA (2006) recommended implementing EBPP to “promote effective psychological practice and enhance public health by applying empirically supported principles of psychological assessment, case formulation, therapeutic relationship, and intervention” (p. 273). APA (2008) stated that the general principles for EBPP not only apply to the treatment but also the psychological assessment process that is fundamental to effective treatment.

It is crucial to illuminate the relationship between EBPP and empirically supported treatments (EST). ESTs require focus on the treatment to inquire if a treatment works for a specific problem under certain circumstances. EBPP ensures that research evidence drives the choices of clinicians in providing effective interventions. Moreover, EBPP includes a wider range of clinical activities, such as treatment, psychological assessment, and therapeutic relationships. ESTs are only those specific psychological treatments found to be effective in controlled clinical experiments. Therefore, “EBPP articulates a decision-making process for integrating multiple streams of research evidence—including but not limited to RCTs—into the intervention process (APA, 2006, p. 273). According to the criteria for empirically validated treatments as established by APA (2008), play therapy is not yet an evidence-based discipline.
In order to move a therapeutic technique toward EST, researchers and practitioners need to consider the following characteristics: “randomized controlled trials, well-described and replicable treatment, test with clinical samples, tests of clinical significance, broad-based outcome assessment including measure of real-world function, and others” (Kazdin & Weisz, 1998, p. 19). Specifically, assessment criteria for the establishment of EST require valid and reliable measurement of real-world change related to the constructs of the mental health intervention.

Assessment in Play Therapy Research

Researchers have continually used different assessments to assess play therapy outcomes in order to capture constructs that Landreth (2012) identified for CCPT. Landreth listed the following objectives for CCPT:

1. Develop a more positive self-concept.
5. Become more self-reliant.
7. Experience a feeling of control.
8. Become sensitive to the process of coping.
9. Develop an internal source of evaluation.
10. Become more trusting of himself. (p. 84)
Bratton and Ray (2000) conducted a comprehensive literature review of 82 play therapy studies, which were published between 1942 and 2000. They summarized studies demonstrating the effectiveness of play therapy with different presenting concerns and populations, such as children confronted with various emotional and behavioral challenges resulting from chronic illness, domestic violence, adjustment difficulties, and learning disabilities (Bratton, & Ray, 2000; DeGangi, Wietlisbach, Goodin, & Scheiner, 1993; Elliott, & Pumfrey, 1972; House, 1970; Jones & Landreth, 2002; Kot, Landreth, & Giordano, 1998; McGuire, 2000; Pelham, 1972; Rae, Worchel, Upchurch, Sanner, & Daniel, 1989; Ray, 2011; Thombs & Muro, 1973; Trostle, 1988; Wall, 1979). Other researchers studied the effectiveness of play therapy as a treatment method with different types of mental health diagnoses, such as anxiety, conduct disorder, depression, attention deficit, and other disorders (Brandt, 1999; Clatworthy, 1981; Milos & Reiss, 1982; Ray, Schottelkorb, & Tsai, 2007; Seeman, Barry, & Ellinwood, 1964). In addition to individual research studies, three meta-analyses on the effectiveness of play therapy demonstrated positive effects of play therapy on childhood behavioral problems, providing a more integrative summary of a combination of studies and suggesting play therapy as an evidence-based therapy (Bratton, Ray, Rhine, & Jones, 2005; LeBlanc & Ritchie, 2004; Lin, 2011).

Another method to assess the effectiveness of CCPT and move CCPT toward EST status involves developing a rigorous theory-based instrument that will truly measure the objectives of CCPT. In 2010, Holliman used CCPT philosophy to develop the Child Interpersonal Relationships and Attitudes Assessment (CIRAA). Holliman (2010) reviewed various child assessments, such as Piers-Harris Children's Self-Concept Scale, Behavior Assessment System of Children-2 (BASC-2), Eyberg Child Behavior Inventory, Behavior Dimension Rating Scale, Parent Child Relationship Inventory, Parenting Stress Index (PSI), and Child Behavior Checklist.
(CBCL; Abidin, 1995; Achenbach & Rescorla, 2001; Bullock & Wilson, 1989; Eyberg & Pincus, 1999; Gerard, 1994; Piers & Hersberg, 2002; Reynolds & Kamphaus, 2004) to understand current assessments available to clinicians in the field. However, none of the evaluated instruments had been specifically designed for assessing play therapy outcomes. Moreover, most of these assessments focused on children’s behavioral changes. Such focus is not consistent with the CCPT philosophy of emphasizing the changes occurring within children (Holliman, 2012; Landreth, 2012). Although CCPT research dates back to the 1940s and many experimental studies have been conducted, CCPT does not yet meet APA (2008) criteria as an evidence-based discipline (Bratton et al.; Ray, 2011). The purpose of generating the CIRAA was to develop a parent report instrument congruent with CCPT philosophy that could be used to research the effectiveness of CCPT (Holliman, 2010).

Instruments Used in Play Therapy Research

A crucial process of any psychotherapy development is to establish a body of evidence for its effectiveness. Therefore, one of the major tasks for researchers is to determine that selected assessments truly measure the named constructs of interest (APA, 2008; Heppner, Wampold, & Kivlighan, 2008). A review of research in play therapy indicates that various assessment instruments are used to measure multiple constructs. Packman and Bratton (2003) and Garza and Bratton (2005) chose the Behavior Assessment System for Children (BASC; Reynolds & Kamphaus, 1992) to evaluate the effectiveness of CCPT with learning disabled preadolescents exhibiting behavior problems and with Hispanic children. The BASC-2 (Reynolds & Kamphaus, 2004) incorporated multiple sources of information to assess the self-perceptions and adaptive and clinical behaviors of individuals 2 to 25 years of age. The Index of
Teaching Stress (ITS; Abidin, Greene, & Konold, 2004) was utilized in Ray’s (2007) and Ray et al.’s (2007) studies to assess the effectiveness of CCPT on reducing teacher-child relationship stress and symptoms of attention deficit hyperactivity disorder in children. The ITS assesses the stress a teacher experiences in the relationship with a particular student (Abidin et al., 2007). Abidin et al. (2007) developed the ITS assessment based on the belief that a student’s academic and personal success and the relationship between a teacher and student are highly related to each other.

Porter Parental Acceptance Scale (PPAS; Porter, 1954), Filial Problem Checklist (FPC; Horner, 1974), Revised Children’s Manifest Anxiety Scale (RCMAS; Reynolds & Richmond, 1985), Parenting Stress Index (PSI; Abidin, 1995), and Child Behavior Checklist (CBCL; Achenbach & Rescorla, 2001) are some examples of frequently used assessments in studies of the effectiveness of CCPT with different populations and varieties of concerns (Brandt, 2001; Bratton & Landreth, 1995; Ceballos & Bratton, 2010; Chau & Landreth, 1997; Costas & Landreth, 1999; Glover & Landreth, 2000; Grskovic & Goetze, 2008; Harris & Landreth, 1997; Holt, 2011; Jang, 2000; Jones & Landreth, 2002; Jones, Rhine, & Bratton, 2002; Packman & Bratton, 2003; Shen, 2002). Among these assessments, CBCL (Achenbach & Rescorla, 2001) and PSI (Abidin, 1995) are two of most used assessments when studying the effectiveness of CCPT (Brandt, 2001; Bratton & Landreth, 1995; Ceballos & Bratton, 2010; Chau & Landreth, 1997; Costas & Landreth, 1999; Grskovic & Goetze, 2008; Harris & Landreth, 1997; Holt, 2011; Jang, 2000; Jones et al., 2002; Packman & Bratton, 2003).

Although play therapy researchers attempted to provide evidence to empirically validate CCPT as a therapeutic treatment, researchers have typically utilized global assessment instruments that may or may not be appropriately linked with play therapy constructs (Holliman,
13

To date, no single instrument has been specifically validated for assessing the effectiveness of play therapy based on child-centered theory.

Holliman (2010) reviewed play therapy research across different disciplines and presenting concerns and concluded that commonly used child assessment instruments in these studies are: the PSI (Abidin, 1995), the CBCL (Achenbach & Rescorla, 2001), and the BASC (Reynolds & Kamphaus, 2004). Lin (2011) conducted a meta-analysis of controlled outcome studies on CCPT modalities. Lin aimed to estimate the overall effectiveness of child therapy intervention utilizing CCPT and discover the relationships between the studies’ characteristics and treatment outcomes. Lin (2011) reviewed 52 studies from 1995 to 2011, and one of the inclusion criteria for the analysis was the use of standardized psychometric assessments. Among the 52 studies Lin (2011) reviewed, the PSI (Abidin, 2002) and CBCL (Achenbach & Rescorla, 2001) seemed to be the most commonly used assessments for measuring the results of CCPT (Brandt, 2001; Bratton & Landreth, 1995; Ceballos & Bratton, 2010; Chau & Landreth, 1997; Costas & Landreth, 1999; Grskovic & Goetze, 2008; Harris & Landreth, 1997; Holt, 2011; Jang, 2000; Jones et al., 2002; Packman & Bratton, 2003). Lin’s (2011) findings further supported Holliman’s (2010) discovery of the frequent use of the PPAS (Porter, 1954), FPC (Horner, 1974), BASC-2 (Reynolds & Kamphaus, 2004), and ITS (Abidin et al., 2004) for CCPT measurement. These instruments were not developed with play therapy in mind, and they often do not correspond with theoretical orientation and outcome objectives of the therapy being examined (Holliman & Ray, 2013). Most of these assessments were designed to emphasize problem and deficit-based behaviors of children, instead of focusing on internal factors consistent with the objectives of CCPT, such as self-acceptance or self-direction. Even though CCPT acknowledges behavioral changes as probable outcomes of therapy (Post, Ceballos, &
Penn, 2012), the goals of CCPT are not to change children’s behaviors. Heppner et al. (2008) emphasized that a crucial activity for researchers is evaluating an instrument’s ability to accurately and effectively measure the planned constructs. Therefore, the historical mismatch between the CCPT intervention and assessments used for the measurement of CCPT outcomes is problematic.

Child Behavior Checklist

Achenbach and Rescorla (2001) developed the CBCL in order to gather information regarding demographic data, competencies at school and home, and adaptive and maladaptive behaviors for children 6 to 18 years of age. Achenbach and Rescorla contracted Temple University’s Institute for Survey Research to execute the national sampling frame in order to build a normative group for the CBCL. The CBCL had 1,753 respondents, and the ethnic composition was 60% Caucasian, 20% African American, 9% Latino, and 12% biracial or undefined. The CBCL encompasses four sets of scales: the Competence scales, the Syndrome scales, the Internalizing/Externalizing/Total Problem scales, and the DSM-Oriented scale (Achenbach & Rescorla, 2001; Holliman, 2010). The Competence scale suggests a child’s competence in the areas of extracurricular activities, socialization, academics, and overall competence. The Syndrome scale includes eight scales: Anxious/Depressed, Withdrawn/Depressed, Somatic Complaints, Social Problems, Thought Problems, Attention Problems, Rule-Breaking Behavior, and Aggressive Behavior. The Internalizing scale shows a child’s level of functioning in relation to internalizing problems, such as depressive, withdrawn, and anxious behaviors. The Externalizing scale suggests a child’s level of functioning in relation to externalizing problems, such as aggressive and rule-breaking behaviors. The Total Problem
scale is the combination of the Internalizing and Externalizing scale. Finally, the DSM-Oriented scale comprises Affective Problems, Anxiety Problems, Attention-Deficit/Hyperactivity Problems, Conduct Problems, Oppositional Defiant Problems, and Somatic Problems subscales. The scores on the CBCL belong to one of the three categories: adaptive/normal; borderline, meaning that a parent or main caregiver may be concerned about the child’s behaviors; and clinical, indicating that a child’s behavior may need immediate intervention.

The assessment developers utilized internal consistency, test-retest, and cross-informant agreement to establish reliabilities of the CBCL (Achenbach & Rescorla, 2001). The Cronbach’s alpha coefficient ranged from .63 to .79 for the Competence scale indicating good reliability. Moreover, the Cronbach’s alpha coefficient ranged from .78 to .97 for the Syndrome scale and .72 to .91 for the DSM-Oriented scale implying very good reliabilities (Achenbach & Rescorla, 2001). Achenbach and Rescorla (2001) also examined the test-retest reliability through asking the same parents to complete the assessment a second time at an interval of 8 to 16 days. The Person \( r \) correlation coefficient was used to estimate the test-retest reliability. The \( r \) ranged from .91 to .95 for the Total Problems scale, and the \( r \) means for the Syndrome and Total Problems scales were not lower than .90, suggesting a strong reliability (Achenbach & Rescorla, 2001; Holliman, 2010). Cross-informant agreement is an approach when investigating the differences between different respondents using the same instrument (Achenbach & Rescorla, 2001). The mean correlation coefficient \( r \) was .69 for Competence scale, .76 for Problem-based scale, and .73 for DSM-oriented scale for mother and fathers of the same children. These results showed a high reliability on cross-informant comparison. Even though mothers tended to rate children higher than fathers on the Syndrome and DSM-oriented scales, this difference only accounted for
less than 4% of the total variance and was considered to be small (Achenbach & Rescorla, 2001; Holliman, 2010).

Achenbach and Rescorla (2001) established the CBCL’s validity through examining content validity, criterion-related validity, and construct validity. Content validity is the most basic kind of validity. The developers established the content validity through careful selection of items and identified and omitted problem items. Achenbach and Rescorla stated that through nearly four decades of research, feedback, consultation, and modification, the findings strongly supported the content validity of the CBCL. They utilized different analyses, such as multiple regression analyses, clinical cut points, and discriminant analyses, to explore the criterion-related validity of the CBCL. Finally, Achenbach and Rescorla established the construct validity through the investigation of correlations between the CBC Problem scales with DSM diagnoses and scores from other instruments, such as the Conner’s Parent Rating Scale-Revised (CPRS-R; Conners, 1997) and the BASC (Reynolds & Kamphaus, 1992). The correlation coefficient between the CBCL and CPRS-R ranged from .71 to .85, indicating a high correlation (Achenbach & Rescorla, 2001).

Parenting Stress Index

Abidin (1995) designed the PSI in 1983 in response to the need to measure child characteristics, parent characteristics, life stress events, and family context of parent-child system. This instrument is an assessment used to measure stress in the parent-child relationship. The PSI measures stress mainly from three facets: parent’s perception of the influence of given temperamental characteristics of one’s children, parent’s attitude of being a parent, and situational/demographic life stress. The PSI is a self-administered questionnaire that contains
120 items. The PSI can be given to parents of children ranging from 1 month old to 12 years of age. The PSI encompasses six subscales for the Child Domain and seven subscales for the Parent Domain. The two domains’ scores are added up to obtain a total stress score. Other than the scores for child domain, parent domain, and total stress, the PSI offers a defensive responding measure and a life stress measure.

The PSI was administered to 2,633 mothers ranging in age from 16 to 61 years and a mean age of 30.9 years (Abidin, 1995). The mothers’ ethnic composition was 76 % Caucasian, 11% African American, 10% Hispanic, and 2% Asian. In addition to the sample of 2,633 mothers, normative data were gathered from 200 fathers. The age of the fathers ranged from 18 to 65 years, with the mean age being 32.1 years. The fathers’ ethnic makeup was predominately Caucasian at 95%, with 5% African American.

Abidin (1995) investigated the factorial validity through three factor analyses. The responses of the 47 items of the Child Domain scale formulated the data for the first analysis. The six-factor solution accounted for 41% of the variance explained. The responses of the 54 items of the Parent Domain scale formulated the data for the second analysis. The seven-factor solution accounted for 44% of the variance explained. The third factor analysis was conducted using the data from the 13 subscales. The two-factor solution accounted for 58% variance explained. Evidence for the construct and predictive validity of the PSI has also been found by a number of researchers (Cameron & Orr, 1989; Chaffee, Cunningham, Secord-Gilbert, Elbard, & Richards, 1991; Jarvis & Creasey, 1991; Zakreski, 1983).

Through factor analysis, Abidin (1995) found the Child Domain to include the following scales: (a) Distractibility/Hyperactivity, a measure of a parent’s perception of a child exhibiting behaviors related to attention-deficit disorder with hyperactivity; (b) Adaptability, a scale that
indicates a child is perceived to have struggles in adjusting to novel situations; (c) Reinforces Parent, a measure of parental experience of a child as origin of encouragement and reinforcement; (d) Demandingness, a scale of a parent’s perception of how much demand a child places on him or her; (e) Mood, a measure related to a child whose emotional functioning is an area of concern; and (f) Acceptability, an indication of a parent’s perception of a child as having intellectual, emotional, or physical characteristics that are not satisfied with a parent’s expectation (Abidin, 1995; Holliman, 2010).

Abidin (1995) reported the Parent Domain to encompass the following scales: (a) Competence, a scale that indicates the degree of a parent’s knowledge about child development or child management skills; (b) Isolation, a measure of a parent’s perception of his or her degree of isolation from family and social support system; (c) Attachment, a scale that suggests how strongly a parent feels emotionally connected to a child or how accurately a parent can understand a child’s feelings; (d) Health, a measure that shows possible health issues that may lead to parental stress; (e) Role Restriction, a measure that assesses the degree to which a parent feels the role of parent limits his or her freedom and inhibits his or her development or preservation of personal identify; (f) Depression, a scale that suggests the presence of symptoms of depression; and (g) Spouse, a measure of emotional and active support from the spouse in a parent role (Abidin, 1995; Holliman, 2010).

Abidin (1995) utilized internal consistency and test-retest methods to establish reliability for the PSI. Cronbach’s (1951) coefficient alpha reliability coefficients were estimated for each subscale, each domain, and the total stress score. For the subscales of the Child Domain, the coefficients ranged from .70 to .83. For the subscales of the Parent Domain, the coefficients ranged from .70 to .84. Also, the reliability coefficients for the total stress scale and two domains
were .90 or greater. These coefficients suggested a high degree of internal consistency for the measures. Burke (1978), Hamilton (1980), and Zakreski (1983) studied the test-retest reliability of the PSI to support the stability of the PSI scales and determined that the PSI has high test-retest reliability coefficients, supporting stability of scores across the time intervals.

Other Instruments Used in Measurement of CCPT

Other than the CBCL and PSI, additional assessments have been used in the measurement of CCPT. These assessments include the BASC, ITS, and RCMAS (Abidin et al., 2004; Flahive & Ray, 2007; Garza & Bratton, 2005; Jones & Landreth, 2002; Packman & Bratton, 2003; Ray, 2007; Ray et al., 2007; Reynolds & Kamphaus, 1992; Reynolds & Richmond, 1985). Garza and Bratton (2005), Packman and Bratton (2003), and Flahive and Ray (2007) employed the BASC (Reynolds & Kamphaus, 1992) for measurement. Reynolds and Kamphaus (2004) further developed the BASC and published the BASC-2 as the second edition of this instrument. The BASC is a multi-method approach to evaluate self-perception and the behavior of people ages 2 to 25 years. The BASC-2 (Reynolds & Kamphaus, 2004) encompasses the four rating components of Children-Teacher Rating scale (TRS), Children-Parent Rating scale (PRS), Children-Self-Report of Personality (SRP), and Structured Developmental History (SDH). The different forms of the BASC-2 may be administered either individually or in any combination (Reynolds & Kamphaus, 2004). The BASC-2 offers a wide range of information that can be utilized for evaluation for educational services, clinical diagnosis, assessment of sensory impairment, forensic evaluation, and evaluation of student progress (Reynolds & Kamphaus, 2004).
Both Ray (2007) and Ray et al. (2007) utilized the ITS (Abidin et al., 2004) as a measurement for their play therapy studies. The ITS is used to measure the stress a teacher may experience in the relationship with a particular student and enables evaluation of the independent factors that relate strongly with the quality of the teacher-child relationship. This relationship includes the student’s behavioral characteristics, the teacher’s perception of the teaching process, and the teacher’s perceptions of support from others who interact with the student (Abidin et al., 2004; Ray et al., 2007). The ITS is appropriate for use with teachers of students from preschool through the 12th grade and contains 90 Likert-scale items (Abidin et al., 2004).

The RCMAS (Reynolds & Richmond, 1985) was used by Shen (2002) and Jones and Landreth (2002) in studying the effectiveness of short-term group play therapy with Chinese earthquake victims and the efficacy of intensive individual play therapy for chronically ill children, respectively. Reynolds and Richmond (2008) further developed the RCMAS and published the second edition as RCMAS-2. The RCMAS-2 is a self-report instrument utilized to assess the level and nature of anxiety, including physiological anxiety, worry/oversensitivity, social concerns/concentration, and a lie factor. The RCMAS-2 is suitable for people from ages 6 to 19 years who have a third grade reading level. The RCMAS-2 offers children opportunities to report on their mental status.

Limitations of Assessments Used for CCPT

From the review of these commonly used assessments, several suggestions for developing assessment instruments are evident. First, large, diverse sample sizes are crucial to represent the population for whom the assessment was designed. Second, these assessment developers utilized rigorous methodology to identify the reliability and validity and multiple
measures to establish the psychometric properties of their instruments. Third, none of the instruments were specially designed for measuring the outcomes of play therapy. The few available play therapy-based instruments, such as the Children’s Play Therapy Instrument (CPTI; Kernberg, Chazan, & Normadin, 1998), Carmichael Therapist/Client Interaction Matrix (Carmichael, 1993), Play Therapy Observational Instrument (PTOI; Howe & Silvern, 1981), and Trauma Play Scale (TPS; Findling, Bratton, & Henson, 2006) require trained observers to fill out the assessment. Also, they only focus on in-session behaviors. Lastly, these assessments were developed to assess problem behaviors and assist clinicians identifying children with struggles (Holliman, 2010). None of these instruments specifically address the objectives of CCPT. Such objectives include assuming greater self-responsibility, experiencing a sense of self-control, and developing an internal source of evaluation (Landreth, 2012). Therefore, the need for developing an instrument that can be utilized to measure the play therapy outcomes and objectives of CCPT is evident and necessary to further the research of the CCPT field.

Child Interpersonal Relationship and Attitudes Assessment

Because of the limitations and restrictions of assessments utilized in the measurement of the effectiveness of CCPT, Holliman and Ray (2013) developed a theory-based assessment with clinical utility and reasonable reliability and validity, recommended the instrument be used to establish a global assessment of CCPT effectiveness, and suggested CIRAA be used to study play therapy. With CIRAA, CCPT practitioners may collect timely and accurate information regarding client progress.
Process of Developing the CIRAA

Instrument Development

Holliman (2010) adapted Springer, Abell, and Hudson's (2002) system for developing rapid assessment instruments (RAI). Clinicians are able to utilize RAIs to quickly and accurately monitor and assess change in clients (Springer et al., 2002). Two steps are needed in the process of developing a RAI. First, researchers need to design the instrument conceptually. Second, researchers need to conduct psychometric validation.

Holliman and Ray (2013) engaged in five phases of assessment development. These five phases encompass the literature review, parent focus group, expert review, expert focus group, and pilot administration. This five-step process corresponded to Springer et al.’s (2002) first step for the developing of RAI. Springer et al. (2002) stated that defining constructs are crucial and often difficult. Researchers ought to define constructs in such a manner that is apparent, differentiated from other similar constructs, and obviously connected to operational definitions. Researchers need to consider how to measure and for whom when developing RAI (Springer et al., 2002). The literature review confirmed the identified objectives, such as self-concept, self-direction, self-responsibility, self-acceptance, self-control, internal source of evaluation, and coping, as delineated by Landreth (2012).

In the second phase, Holliman and Ray (2013) explored parents’ views of objectives for children in play therapy. They identified 12 parents, who reported improvement of their children’s behaviors on the CBCL and PSI after participating in 10 sessions of CCPT. They sought to conduct interviews with these parents. Parents of seven children responded to the interview request, and the interviews were conducted by two advanced doctoral students, who had obtained sufficient training in play therapy and research interview procedures (Holliman,
2010; Holliman & Ray, 2013). The goal of this process was to incorporate the theoretical aspirations of CCPT with parents’ practical perceptions regarding the value of CCPT. The research team allocated parent responses to one of identified objectives of CCPT or formed new categories if parent responses were not stated within the listed CCPT objectives when reviewing transcripts of interviews. The research team identified five pertinent categories through this process. These five categories were self-concept, disruptive behaviors, social skills, self-direction/self-responsibility, and coping skills. Three categories, self-concept, self-direction/self-responsibility, and coping skills, resonated with an identified CCPT objective, and the other two categories, social skills and disruptive behaviors, were expressly significant to parents but did not explicitly address CCPT objectives. Holliman and Ray utilized these categories as a guide and created 63 items. They designed 17 items for social skills, 8 items for self-concept, 15 items for disruptive behaviors, 12 items for self-direction/self-responsibility, and 11 items for coping skills. Items were worded both positively and negatively. Holliman (2010) developed 63 items to be measured through a five-point Likert scale in which parents rated the child's behaviors. The item responses ranged from strongly agree to strongly disagree.

After developing the original 63 items, the third phase of instrument development began (Holliman & Ray, 2013). Holliman and Ray (2013) obtained feedback about the 63 items from eight identified CCPT experts. Each of the eight experts possessed an advanced degree in mental health, advanced education in CCPT, a mental health license, and a registered play therapist designation. The experts assessed the instrument for adherence to psychometric standards, readability, clinical usefulness, and relevance of each item to designated category according to recommendations of Springer et al. (2002). More than 85% of the experts attested that 35 items
accurately described their assigned construct category (Holliman & Ray, 2013). The remaining 29 items were recognized as correctly describing the construct by less than 85% of the experts.

In the fourth phase, Holliman and Ray (2013) conducted a focus group with four university faculty members, who specialized in CCPT, and with four advanced doctoral students, who had taken at least three didactic and four clinical courses in CCPT (Holliman & Ray, 2013). The focus group revisited, modified, and offered ideas about the items that had not been supported by more than 85% of the experts from the previous phase. The focus group also offered balance among positively- and negatively-worded items and addressed multicultural sensitivity. Fifty-two items (social skills, n = 11; self-concept, n = 11; disruptive behaviors, n = 11; self-direction/self-responsibility, n = 12; coping skills, n = 7) were finalized and used in the pilot study. In addition to the 52 items, a demographics section was included to allow respondents to report ethnicity, age, grade, and gender of children in a free-response format (Holliman, 2010; Holliman & Ray, 2013).

In the final phase of development, Holliman and Ray (2013) administered the procedures recommended by Pett et al. (2003) for the pilot study by using a sample of approximately 10% of the size of the anticipated sample for the main study. Therefore, Holliman (2010) conducted a pilot study with 20 parents of children who received counseling services at a local mental health clinic to examine the instrument. This sample included 11 girls and 9 boys from 3 to 9 years old with an average age of 6.85 years old. Nineteen of the pilot study parents identified themselves as Caucasian, and one parent identified as other (Holliman & Ray, 2013). The Cronbach’s coefficient alpha yielded a reliability estimate of .85, which met the prerequisite of reliability estimates for basic research on instrument development as suggested by Bernstein and Nunnaly (1994). Holliman (2010) then conducted the exploratory factor analysis (EFA) with the pilot
sample data and deleted four items due to low loading values (smaller than .40) as recommended by Henson and Roberts (2006). Forty-eight items (social skills, \( n = 9 \); self-concept, \( n = 11 \); disruptive behaviors, \( n = 10 \); self-direction/self-responsibility, \( n = 11 \); coping skills, \( n = 7 \)) were retained for use in the main study (Holliman, 2010; Holliman & Ray, 2013).

Instrument Validation

Holliman (2010) followed Worthington and Whitaker’s (2006) suggestion of using the subject to item ratio of 3:1. For the main study, Holliman successfully recruited 136 parents of children aged 3 to 10 years, with the median age being 6.5 years, who completed the CIRAA. Among the sample of 136 parents, 97 (71%) parents were viewed as a clinical sample because of their referral to counseling services through the university-based counseling clinic or local schools (Holliman, 2010; Holliman & Ray, 2013). The rest of the 39 parents were recruited through a flyer circulated in school and were considered to be a normal sample and reported no behavioral or emotional concerns for their children. In addition to the CIRAA items, parents filled out the PSI and CBCL for assessing concurrent validity. However, only 101 PSIs and 80 CBCLs were appropriately completed and returned (Holliman, 2010; Holliman & Ray, 2013).

RAI developers are required to examine reliability and validity in various forms in order to determine the psychometric qualities of the instruments (Springer et al., 2002). Through the process of literature review, parent focus group, expert review, expert focus group, and finally pilot administration, Holliman and Ray (2013) formulated, according to the underlying item construct, the five category conceptual theory of the CIRAA. The CIRAA then was ready to undergo the instrument validation process via EFA.
Kerlinger (1979) stated that factor analysis is “one of the most powerful methods yet for reducing variable complexity to greater simplicity” (p. 180). EFA can help reduce a large number of variables to a manageable set of factors (Henson & Robert, 2006). Thompson and Daniel (1996) stated that factors and constructs "can be seen as actually causing the observed scores on the measured variables" (p. 202). Holliman (2010) utilized EFA with the entire sample ($n = 136$) in order to determine which items to retain in CIRAA’s factor structure. The EFA represented the second step, conducting psychometric validation, of developing RAIs stressed by Springer et al. (2002).

Holliman (2010) utilized eigenvalue-greater-than-one ($EV<1$) rule (Kaiser, 1960), scree plot (Cattell, 1966), and parallel analysis (O’Conner, 2000) to assist with the decision of factor retention. Henson (2010) stressed the importance of extracting the correct number of factors because this process would directly influence the interpretation of the final outcomes. Different rules can be applied during the factor retention process, such as the $EV<1$ (Kaiser, 1960) and the scree plot (Cattell, 1966). Henson and Roberts (2006) stated that $EV<1$ and the scree plot are the most commonly used methods during the factor retention process. Henson (2010) clarified that "each factor comes with a statistic called an eigenvalue, which is a representation of how much variance the factor can reproduce out of all the variance present within and between the items in the matrix of associations" (p. 18). Byrd and Henson (2007) also suggested that $EV<1$ is often inaccurate and may lead to retaining too many factors, particularly when a large number of items are presented (Zwick & Velicer, 1982).

Because of this disadvantage, a more accurate method, parallel analysis, was applied using O'Conner's (2000) syntax. Parallel analysis is considered to be a more stringent factor retention strategy (Zwick & Velicer, 1986). Through the parallel analysis, "only factors from the
real data whose eigenvalues are larger than the eigenvalues from the respective factors in the random data are retained” (Henson, 2010, p. 20). Factor rotation is often a necessary step to make the meaning of factor more evident (Henson, 2010). Holliman (2010) used varimax rotation, an orthogonal rotation, based on the assumption of the uncorrelated nature between factors. Items were retained based on two standards (Holliman & Ray, 2013). One was the strength of structure coefficients. Based on Thurston's (1947) suggestion, variables with structure coefficients lower than .4 (16%) are usually considered too weak to interpret. The other was the absence of cross-loading on factors. Pett et al. (2003) stated that items should be deleted if their structure coefficients attain a value of .4 or higher on two or more factors.

For the purpose of establishing reliability for the subscales and the whole scale, Holliman (2010) conducted the Cronbach’s alpha coefficient to assess the internal consistency of the CIRAA. Dimitrov (2012) suggested researchers should gather both internal consistency and alternate forms of reliabilities whenever possible while developing instruments. Holliman and Ray (2013) chose not to employ methods for test-retest reliability due to the expectation that CIRAA scores change over time due to the effects of intervention or maturation. Holliman (2010) explored criterion validity through correlation analysis between the CIRAA, CBCL, and PSI. He also used the receiver operator characteristic (ROC) curve for determining a clinical threshold score for the CIRAA. The outcomes of the instrument validation process are presented below.

Outcome of the CIRAA Development

*Exploratory Factor Analysis*

Based on the EFA, Holliman (2010) retained 30 items, extracted four factors from the
data, named these factors the following: (a) Self-Control (12 items), (b) Interpersonal Relationships (7 items), (c) Coping Skills (6 items), and (d) Internal Locus of Evaluation (5 items). Holliman and Ray (2013) named these factors according to the conceptualization of correlated items within the context of CCPT’s structure. The 30 items had strong structure coefficients for designated factors and explained 53.83% of the variance in the original matrix of associations. Holliman (2010) constructed items in a way that high scores indicated the presence of less desirable behavior of a child from the perspective of a CCPT practitioner.

**Reliability**

Dimitrov (2012) stated that reliability scores of .90 or higher are usually desirable for achievement tests, and reliability scores of .75 or higher are respectable for assessments measuring more steady attributes, such as behavioral or psychological instruments. The internal consistency reliability for the CIRAA was .93, which met the criteria of acceptability of reliability for an instrument (Nunnaly & Bernstein, 1994). The Self-Control factor had a Cronbach's alpha coefficient of .91, with inter-item correlations varying from .286 to .706 and an average correlation of .47. The Interpersonal Relationships factor had a Cronbach's alpha coefficient of .86, with inter-item correlations varying from .308 to .686 and an average correlation of .46. The Coping Skills factor had a Cronbach's alpha coefficient of .77, with inter-item correlations varying from .174 to .614 and an average correlation of .349. The Internal Locus of Evaluation factor had a Cronbach's alpha coefficient of .72, with inter-item correlations varying from .197 to .569 and an average correlation of .369.

When considering the relationship between instrument length and strong Cronbach’s alpha coefficient values, Clark and Watson (1995) stated that it is usually acceptable for
subscales’ reliability coefficients to be lower than the entire instrument’s reliability. They also suggested that scale coefficient correlations ranging from .15 to .50 can be considered acceptable. Therefore, the internal consistency reliabilities were respectable for the whole scale and subscales; and the inter-item correlations were acceptable for each of the four factors.

Criterion Validity

Concurrent validities were obtained to assess the correlation between the CIRAA total score and scores for the PSI and the CBCL (6-18). Holliman and Ray (2013) conducted correlational analyses of the CIRAA and the CBCL with a clinical sample of 80 parents, whose children were referred by school counselors for play therapy at local elementary schools. The Pearson’s $r$ coefficient of .75 ($n = 80, p < .001$) for the total score on the CIRAA and Total Problems score on the CBCL was obtained indicating the positive correlation between reported undesirable behaviors on the CIRAA and clinical behaviors on the CBCL (Holliman & Ray, 2013). Holliman and Ray also used the Child Domain score on the PSI as a criterion reference because of its emphasis on children’s emotions and behaviors being comparable to items on the CIRAA. The Pearson’s $r$ coefficient of .74 ($n = 101, p < .001$) for the total score of the CIRAA and the Child Domain score on the PSI indicated significant positive relationships between these two instruments.

ROC Analysis

ROC curve analysis was conducted in order to determine the ability of the CIRAA to identify children with clinical behavior problems accurately. Holliman (2010) applied Fan, Upadhye, and Worster’s (2006) guideline of .75 area under the curve (AUC) for minimum
clinical utility. When ROC curve analysis was conducted with CIRAA and CBCL outcomes, the outcome yielded an AUC of .866, with a 95% confidence interval ranging from .813 to .958, locating under the respectable border for discriminatory power. When analyzed with the CBCL, the CIRAA yielded a clinical threshold score of 2.52 with an ability to identify clinical cases 85% of the time (Holliman & Ray, 2013).

Holliman and Ray conducted the ROC curve analysis with CIRAA and PSI Child Domain outcomes for data from 101 parents. Analysis showed an AUC of .863, with a 95% confidence interval ranging from .791 to .934, surpassing the minimum values required for clinical utility (Fan et al, 2006). When explored with the PSI, the CIRAA yielded a clinical threshold score of 2.55 with an ability to distinguish clinical cases 85% to 87% of the time. Based on the ROC curve analyses, Holliman (2010) concluded that the score of 2.5 was the clinical threshold score on the CIRAA.

Strengths of CIRAA

When developing a mental health instrument, rigorous analyses are required and demanded in order to resolve its precise psychometric properties (Dimitrov, 2012; Holliman & Ray, 2013). Aside from psychometric properties, a consistent theoretical basis and some degree of clinical and research utility are needed for a well-developed instrument. Holliman and Ray (2013) adapted Springer et al.’s (2002) approach for developing the RAI and used the five phases discussed above. They conducted a pilot administration in order to ensure the items and corresponding factors were appropriate for the CCPT constructs. The literature review, parent focus group, expert review, and expert focus group ensured the CIRAA would be grounded in theory (Holliman & Ray, 2013). The CIRAA went through rigorous analysis to identify its
psychometric properties (Holliman & Ray, 2013). Holliman and Ray utilized preliminary construct validity, EFA, criterion validity, content validity, reliability, and clinical exploration to demonstrate the CIRAA as a feasible instrument for assessing the effectiveness of process in play therapy.

The CIRAA showed an overall reasonable validity. Holliman and Ray (2013) not only consulted CCPT experts twice but also interviewed parents to offer a more thorough approach for ensuring content validity. Conducting interviews with the assessment’s target population enhanced the possibility of retaining items with greater content validity and represented a validity enhancing tactic rarely used by most researchers (Holliman & Ray, 2013; Vogt, King, & King, 2004). Through EFA, the CIRAA demonstrated acceptable factor loadings for its 30 items, no loadings of items on multiple factors and met the EFA standards for parsimonious factor structure (Henson & Robert, 2006; Holliman & Ray, 2013). The CIRAA showed concurrent validity by demonstrating strong correlations with the CBCL and the PSI. The CIRAA demonstrated reasonable reliability for total problems and its subscales via reliability analysis (Holliman & Ray, 2013). Inter-item correlations’ means and ranges were gathered for each subscale to assess the unidimensionality of each factor. All of these outcomes presented reasonable levels of reliability for both subscales and total score. Unlike many instruments, the CIRAA is quickly administered with only 30 multiple choice items for parents of children ages 3 to 10 to answer (Holliman & Ray, 2013). Lastly, the CIRAA does not require any particular training beyond a master’s degree in a mental health discipline that include courses in test administration and interpretation of statistics.
Identified and Defined Factors on the CIRAA

Holliman and Ray (2010) conducted EFA to ensure construct validity and construct equivalence (Laher, 2010; Pett, Lackey, & Sullivan, 2003). EFA enables the researcher to condense the instrument’s items into fewer constructs to represent the theory in mind more parsimoniously. Holliman (2010) identified four factors on the CIRAA in the context of the objectives of CCPT stated by Landreth (2012) and named them: Self-Control, Interpersonal Relationship, Coping Skills, and Internal Locus of Evaluation. These four factors represent variables that are not directly observable and are considered to be latent variables or constructs (Dimitrov, 2012). Dimitrov (2012) further stated that “typically, the definition of a construct is embedded into a more general theory and then operationalized in the context of the theory and practice in which inferences and decisions are to be made based on assessment scores” (p. 55). Therefore, it is crucial for assessment developers to define carefully the identifying constructs according to the theory on which an instrument is based. The following sections represent my attempt to further define and clarify the four constructs that Holliman and Ray (2013) proposed.

Self-Control

Many researchers have sought to define the concept of self-control (Bronson, 2000; Honing & Lansburg, 1991; Landreht, 2012; Logue, 1995; Mischel, Shoda, & Peake, 1988; Rodriguez, Mischel, & Shoda, 1989; Rosenbaum, 1983; Strayhorn, 2002). These individuals’ efforts to capture the meaning of self-control demonstrate the complexity of self-control as a construct. Moreover, the notion of self-control highly overlaps and closely relates to several other concepts. Logue (1995) defined self-control as “engaging in behaviors that result in delayed (but more) reward” (p. 3). Mischel et al. (1988) operationally defined the term of “delay
of gratification” as how long people wait in the course of gaining a more valued but postponed result. Bronson (2000) described that young children are expected to “delay, defer, and accept substitutions without becoming aggressive or disorganized by frustration, and [to] cope with arousal, whether due to environmental challenge of fatigue” (p. 71).

Kochanska (1993) used the term conscience to indicate the demonstration of self-control when individuals choose the delayed reward through obeying a moral rule and refusing temptation. Concurring with Kochanska’s belief, Honing and Lansburgh (1991) described self-control as “a good signpost that moral development is proceeding well for young children” (p. 22). Baumeister, Heatherton, and Tice (1994) preferred the term self-regulation. Carver and Scheier (2011) clarified that even though people have used self-regulation as an interchangeable term to self-control, they believed that self-regulation encompasses a broader meaning as follows:

The sense of purposive processes, the sense that self-corrective adjustments are taking place as needed to stay on track for the purpose being served (whether this entails overriding another impulse or simply reacting to perturbations from other sources), and the sense that the corrective adjustments originate within the person. (p. 3)

Honing and Lansburgh (1991) stated that young children express self-control through the ability to internalize standards, trust adults, control angry impulses, delay gratification, take turns, find ways to cheer up when feeling sad, and discover internal ways to be more patient regardless of frustrations. As a multifaceted concept, self-control can be synonymous with terms such as self-regulation, affect regulation, delay-of-gratification, and self-discipline.

More importantly, according to Mischel et al. (1988) and Mischel, Shoda, and Rodrigues’ (1989) longitudinal studies, the level of one’s self-control seems to be a consistent trait across time. Mischel et al. (1988, 1989) measured preschool children’s delay of gratification skills and followed up on the same participants as adolescents. Mischel et al. showed that individuals, who
were able to delay gratification on preschool tasks, demonstrated more self-control and were considered more planful as adolescents by their parents. Parker, Cowen, Work, and Wyman (1990) stated that self-control is related to problem-solving skills, high self-esteem, and higher empathy, which are qualities enabling children to cope while dealing with stressed lives. Tangney, Baumeister, and Boone (2004) showed that high self-control is strongly correlated with less pathology, higher grades, better adjustment, more optimal emotional responses, secure attachment, and interpersonal success.

Landreth (2012) noted “in response to the feeling of permissiveness established in the playroom, the safety to be fully one’s self, and the careful use of therapeutic limits…children learn self-control and responsible freedom of expression" (p. 88). He further clarified that children need opportunities to experience being in control before they can gradually learn self-control and self-direction. Honing and Lansburgh (1991) emphasized the importance of parents and caregivers in fostering secure and supportive relationships with children to better develop self-control. Children need to be able to exercise and practice self-control within secure relationships. This conclusion corresponds with Landreth’s belief that one of the main tasks of play therapists is to develop a secure, warm, caring, and accepting relationship with children so they can express themselves more freely. Children will then be able to develop self-control through limit setting and accepting choices presented by therapists. Instead of being controlled by therapists, children are permitted to control their own behaviors in play therapy. Self-control requires individuals to utilize delay of gratification skills, such as attention, which tends to increase as children grow older (Logue, 1995; Mischel et al., 1988, 1989). Children start to exhibit self-control as they establish their internal processes of evaluation (Holliman & Ray, 2013; Landreth, 2012).
Items on the CIRAA within the self-control factor were focused on behaviors such as impulsivity, arguments, fighting, sharing with others, and caring for others (Holliman & Ray, 2013). Other than addressing a child’s inability to control and postpone desires and behaviors, these items also describe the ability to share and care for others. Holliman’s (2010) items on the self-control factor are the following (and are numbered according to their item numbers on the CIRAA):

1. My child shares with other children
5. My child has verbal fights with other student at school
7. My child apologizes for hurting other’s feelings
8. My child blames others for mistakes
13. My child hits/kicks family members
15. My child gets along with other children
18. My child exhibits self-control
20. My child disrupts family events/outing
23. My child yells frequently
25. My child often receives reports of disruptive behavior from his/her teacher
26. My child often acts without thinking
29. My child has physical fights with other children.

Interpersonal Relationship

Various theorists postulated that human are social beings (Adler, 1956a, 1956b; Minuchin, 1974; Miller, 1976; Roger, 1951). Humans innately long for connections with others (Adler, 1956a, 1956b; Bonnie, 2008; Bowlby, 1969; Miller, 1976; Rogers, 1951). Children
mainly encounter three types of interpersonal relationships in their lives: parent-child, teacher-child, and peers. Miller (1993) stated that children carry and maintain their general behavioral patterns and ways of relating to others across interactions and individuals. Therefore, these patterns of interactions may lead to similar results in the different types of relationships children encounter. Because of innate needs of individuals to be a part of larger group, it is important for children to be able to develop healthy and functioning interpersonal relationships with their parents or main caregivers, teachers, and peers. The following discussion includes the reasons positive interpersonal relationships are crucial for children.

According Bowlby (1969), young children are entirely dependent on adults to take care of them, both physically and emotionally. The development of strong bonds between young children and their main caregivers is extremely important for their future development (Bowlby, 1969). Relationship with a parent is one of the first interpersonal relationships that children develop. The development of parent-child relationships has been shown to promote the development of a set of crucial social skills that will later enable children to comprehend their other social relationships (Freitage, Belsky, Grossmann, Grossmann, & Scheuerer-Englisch, 1993; Sroufe & Fleeson, 1986). Cox (1970) indicated that children’s self-perceptions are highly correlated with parental acceptance or rejection. When children have a positive self-perception, they have better chance of developing healthier interpersonal relationships in the future.

Piaget (1932) and Sullivan (1953) affirmed that child-to-child interaction is a central component in facilitating children’s development. Piaget (1952) further stated that peer interaction is one main factor impacting children’s cognitive changes, as the transition from preoperational to concrete operational representations occurs around seven years of age. From Piaget’s (1928) perspective, children in the preoperational stage are socially egocentric, lack of
introspection and rational justification, and can demonstrate inflexibility, which will not be overcome merely by experience with objects and events but mostly through social interaction with peers. Interpersonal experiences, especially with peers, are required to assist children in the preoperational stage for breaking down egocentrism, considering others’ perspectives, and gradually moving into the operational stage (Piaget, 1928; Rardin & Moan, 1971).

Empirical research has supported the positive impact of peer relationships on cognitive (Rardin & Moan, 1971), linguistic (Bates, 1975), gender role (Fagot, 1977), and moral (Berndt, McCartney, Caparulo, & Moore, 1984) development. Parker and Asher (1987) generated a reasonable assumption that low-accepted children may be more vulnerable to later life struggles because peers greatly contribute to the socialization of social competence. Parker and Asher further stated that “because low-accepted children experience limited opportunities for positive peer interaction, it follows that they would be relatively deprived of opportunities to learn normal, adaptive modes of social conduct and social cognition” (p. 358). Through a review of different studies, Parker and Asher found that peer relationships differ predictably by type of peer relationship. For instance, aggressiveness and low acceptance were more steady predictors than withdrawal/shyness. Little evidence was found about the connection between withdrawal/shyness and later struggles. At the same time, they cautioned that poor peer relationships are not necessarily predictive of later struggles and disturbances, even though these two phenomena may sometimes be related (Parker & Asher, 1987).

In play therapy, children experience a different kind of interpersonal relationship. The CCPT relationship is developed by play therapists through being real, warm, caring, accepting, sensitive, and understanding (Landreth, 2012). Children’s perceptions about the world can be changed through meaningful relationships. In CCPT, the therapeutic relationship is the key to
growth. When children experience consistent acceptance in the therapeutic relationship, they gradually develop enough inner freedom and security to express themselves in self-enhancing ways. Landreth (2012) addressed the importance of interpersonal relationships in the following:

Children sense the therapist's respect, feel respected, and as an absence of evaluation and an ever present acceptance exits, they internalize the respect; thus children learn to respect themselves. Once children have respect for themselves, they learn to respect others. (p. 87)

Even though Landreth (2012) did not explicitly list enhancing interpersonal relationships as one of the objectives of CCPT, it is reasonable to consider this concept as one of the outcomes of CCPT. According to Landreth (2012), CCPT may help children develop a more positive self-concept; assume greater self-responsibility; become more self-directed, self-accepting, and self-reliant; engage in self-determined decision making; experience a feeling of control; become sensitive to the process of coping; develop internal source of evaluation; and become more trusting of oneself.

The interpersonal relationships items on the CIRAA ask parents to report how their children are perceived by other children and how their children perceive themselves (Holliman & Ray, 2013). Holliman’s (2010) items on the interpersonal relationship factor are the following (and are numbered according to their item numbers on the CIRAA):

2. My child complains few people like him/her
3. My child is often sad for prolonged periods of time
9. My child is teased by other children
11. My child complains other children are mean to him/her
14. My child often compares him/herself unfavorably to others
17. My child is overly sensitive
22. My child complains no one likes him/her
Coping Skills

The concept of coping is a very broad construct. Lazarus and his colleagues (Folkman & Lazarus, 1980; Lazarus & Launier, 1978; Monat & Lazarus, 1977) made an attempt to define coping as the “cognitive and behavioral efforts made by individuals to master, tolerate, or reduce” (Folkman & Lazarus, 1980, p. 223) stressful demands while “a routine or automatic response is not available” (Monat & Lazarus, 1977, p. 8). Based on this definition, Curry and Russ (1985) stated that coping is an active purposeful process mediating adjustment outcome. Adjustment or adaptation is viewed as the result of “effective and well-practiced coping efforts” (Murphy & Moriarty, 1976, p. 77).

Folkman and Lazarus (1980) proposed the cognitive appraisal theory of stress with its associated coping model as an attempt to describe the coping process. This model has been a central guide to most studies of the coping process in children (Pincus & Friedman, 2004). According to the cognitive appraisal theory, coping is a multidimensional process influenced by temporal, situation-specific, and personal variables (Folkman & Lazarus, 1980; Pincus & Friedman, 2004). The coping process includes the three major components of cognitive appraisal, coping response, and coping outcome. Lazarus and Folkman (1984) identified two types of coping responses. One is problem-focused coping indicating individuals’ efforts to directly change or master stressors. The other is emotion-focused coping as part of the individual’s effort to regulate and manage the emotions related to the stressful situation.

Different researchers believed that individuals psychological, physical, and social well-being are impacted by the ways individuals cope with stressful situations (Antonovsky, 1979; Cohen & Lazarus, 1979; Moos, 1977) and suggested that individuals’ abilities to purposefully and effectively deal with various stressors and demands of everyday life are important skills for
healthy functioning (Pincus & Friedman, 2004). When thinking about children in relation to coping, Murphy (1962) described the nature of coping in the following:

Through his coping experiences the child discovers and measures himself, and develops his own perception of who and what he is and in time [what he] may become. We can say that the child creates his own identity through his efforts in coming to terms with the environment in his own personal way. (p. 374)

In children’s daily lives, they are consistently confronted by possible interpersonal challenges (Pincus & Friedman, 2004). Children’s capabilities to handle common stressors have been considered to be significantly correlated to psychological adjustment and adaption (Pincus & Friedman, 2004; Rutter, 1994; Spirito, Stark, Grace, & Stamoulis, 1991). Children’s ability to cope effectively with daily stress mediates the influence of major life events and relates to positive emotional and behavioral adjustment (Compas, 1987; Pincus & Friedman, 2004). Curry and Russ (1985) suggested that successful coping experiences in childhood improve resiliency and flexibility and enhance the likelihood for adaptive stress-management in the future. Weisz, McCabe, and Denning (1994) studied children undergoing medical procedures and found children capable of adapting their coping skills to fit the circumstances and flexible in their use of coping skills seem to have the best outcomes. Many researchers have suggested that having coping skills at a young age provides a “buffer” or “moderator” (Pincus & Friedman, 2004, p. 224) for the impacts of stressful life events on the development of psychological maladjustment (Dubow & Tisak, 1989; Spivack, Platt, & Shure, 1976).

Landreth (2012) argued that "becoming sensitive to the process of coping" is among the 10 broad objectives of CCPT (p. 89). When children learn that their feelings are acceptable, they feel freer to express these feelings. Once the feelings are expressed, the intensity of these feelings decreased. As a result, children can control their feelings more easily, which indicated better coping. Studies have showed that young children ages six to nine depend greatly on
problem-focused strategies to cope with different stressful instances; however, older children demonstrate better flexibility in changing between problem-focused and emotion-focused strategies (Caplan, Bennetto, & Weissberg, 1991; Compas, Malcarne, & Banez, 1992; Compas, Worsham, & Ey, 1992; Wertlieb, Weigel, Springer, & Feldstein, 1987). Additionally, young children do have the capability to utilize emotion-focused strategies (Altshuler & Ruble, 1989; Weisz et al., 1994). Children tend to use emotion-focused strategies to cope more effectively with relatively uncontrollable circumstances in which the use of straight action strategies could be ineffective (Altshuler & Ruble, 1989; Pincus & Friedman, 2004; Weisz et al., 1994). Pincus and Friedman (2004) believed the importance for young children to learn, develop, and apply emotion-focused strategies, especially when encountering uncontrollable situations. In CCPT, children may express their feelings freely while finding and developing appropriate behaviors to express those feelings (Landreth, 2012; Ray, 2011). Gradually, children experience the process of coping and move toward better adjustment. Holliman’s (2010) items on the coping skills factor are the following (and are numbered according to their item numbers on the CIRAA):

4. My child is often upset by minor things
6. My child becomes anxious over small matters
19. My child is able to express his/her feelings when he/she is feeling upset
21. My child can name things he/she likes about him/herself
27. For his/her age, my child is able to enter new situations with confidence
28. My child can calm down when upset

Internal Locus of Evaluation

The concept of locus of was developed fully within person-centered theory (Rogers,
1951, 1959, 1961). Rogers (1959) believed that individuals’ psychological health and function are critically influenced by evaluation of self. People develop self-evaluations according to internal perceptions as well as external feedback and responses from others. Rogers (1951) defined locus of evaluation as “the extent to which values and standards depend upon the judgments and expectation of others or are based on a reliance upon [one’s] own experience” (p. 156). Bucur (2007) emphasized that the attitudes toward the self or an individual’s self-esteem can be influenced by not only the judgment itself but also the source of the judgment. Rogers (1959) defined internal locus of evaluation as follows:

An internal locus of evaluation, within the individual himself, means that he is the center of the valuing process, the evidence being supplied by his own sense. When the locus of evaluation resides in others, their judgment as to the value of an object or experience becomes the criterion of value for the individual. (p. 210)

Rogers (1951, 1959) believed that an internal locus of evaluation is not merely an important indicator of successful treatment but also a predecessor to people’s ultimate goal of self-actualization. He stressed the importance of individuals giving credit to their own experiences, standards, and values when formulating an attitude toward the self in order to be genuine and accomplish purpose in life. Through this process, individuals can increase their self-understanding and attain congruence between how they perceive and experience themselves and how they present themselves to others (Bucur, 2007; Rogers, 1951, 1959). Moreover, Rogers (1948) viewed internal locus of evaluation as allowing for “growth and development” (p. 214) and facilitating an increased sense of self-respect and acceptance. The importance of developing internal locus of evaluation can be captured by Rogers (1959) description of a fully functioning person as one who “experiences himself as the locus of evaluation” (p. 234).

Raskin (1952) clarified that the locus of evaluation concept does not suggest total dependence or independence on other people. Individuals’ valuing process is the outcome of
some mixture of both external and internal factors, which means that individuals’ locus of evaluation exists along a continuum (Bucur, 2007; Raskin, 1952). Therefore, when individuals have an internal locus of evaluation, they take external feedback from others into consideration yet put more weight on the internal feedback they experience through the self. Rogers (1959) viewed individuals who encompass optimal psychological adjustment, maturity, and health as fully functioning people. Hence, individuals have a better chance to reach their full potential and sustain psychological health by placing greater emphasis on internal standards and values through an internal locus of evaluation (Bucur, 2007; Rogers, 1959).

In echoing Rogers’ (1951, 1959) belief about the importance of developing an internal locus of evaluation, Landreth (2012) emphasized that "as children experience being accepted just as they are with no conditional expectation from the therapist, they gradually, and in sometime imperceptible ways, begin to accept themselves as worthwhile" (p. 89). Through the experience of acceptance that is free from evaluation, children can develop the ability to evaluate what they like or dislike instead of depending on adults for confirmation (Landreth, 2012). Through nonjudgmental and accepting relationships, children can become who they truly are, develop their internal locus of evaluation, and accept themselves authentically (Landreth, 2012). When children develop an internal locus of evaluation, they learn to trust themselves and listen to their internal feedback (Landreth, 2012; Raskin, 1952). Increased self-acceptance is a major contributor to the development of a positive self-concept (Landreth, 2012). Holliman’s (2010) items on the internal locus of evaluation factor are the following (and are numbered according to their item numbers on the CIRAA):

10. My child enjoys doing things for him/herself

12. My child volunteers to help out around the house
Confirmatory Factor Analysis and the Assessment of Model Fit

Factor analysis is a statistical technique utilized to help researchers understand the structure of the particular phenomenon in which they are interested (Pett et al., 2003). Exploratory and confirmatory are two basic types of factor analysis. EFA is utilized when researchers do not have a preconceived structure or do not know how many factors are required to explain the interrelationship among a set of items, indicators, or characteristics (Gorsuch, 1983; Pett et al., 2003; Tabachnick & Fidell, 2007). In contrast, confirmatory factor analysis (CFA) is utilized to examine the extent to which the hypothesized construction of a set of identified factors fits the data (Nunnaly & Bernstein, 1994; Pedhazur & Schmelkin, 1991; Pett et al., 2003). CFA is used when researchers have some knowledge about the underlying structures of the interest constructs.

“Researchers typically use CFA after an instrument has already been assessed using EFA, and they want to know if the factor structure produced by EFA fits the data from a new sample” (Worthington & Whittaker, 2006, p. 815). CFA is often used after EFA to support the validity of a scale (Brown, 2006). CFA is a type of structural equation modeling (SEM), which handles specific measurement models, designed to show the relationships between observed measures or indicators, such as test items, and latent variables or factors (Brown, 2006). Unlike EFA, CFA is hypothesis-driven in nature. Therefore, researchers must specify all aspects of the CFA model beforehand. For CFA, researchers need to decide, “How many factors are present in
an instrument; which items are related to each factor; and whether the factors are correlated or uncorrelated” (Worthington, & Whittaker, 2006, p. 808). When conducting CFA, researchers use hypothesized models to gauge a population’s covariance matrices and to compare those with the observed covariance matrices (Schreiber, Stage, King, Nora, & Barlow, 2006). Theoretically, researchers would like to lessen the differences between the estimated and observed matrices. Bentler (1995) indicated that large samples sizes are required to provide stable parameter estimates. Bentler and Chou (1987) suggested using at least the 5:1 ratio of participants to number of parameters, with the ratio of 10:1 being ideal.

Two approaches for CFA are the traditional approach and the structural equation modeling (SEM) approach (Garson, 2010). More detail can be gathered when utilizing the traditional approach because it gives the factor loadings of the indicators to attest if they load on the factors as anticipated by the researcher’s model (Garson, 2010; Mehta, 2011). The researcher can compare the traditional approach with the SEM approach, which offers a single-coefficient goodness of fit measures. CFA represents the unique case of SEM, which is also known as linear structural relationship modeling (LISREL; Jöreskog & Sörbom, 2006) and the covariance structure (McDonald, 1978). A measurement model and a structural model are two types of SEM (Brown, 2006; Garson, 2010). A measurement model connects a set of observed variables/indicators to a smaller set of latent variables/factors, like CFA. A structural model ties the latent variables/factors through a series of recursive and non-recursive relationships (Albright & Park, 2009; Mehta, 2011).

The model fit of CFA can be addressed on three major areas: (a) overall goodness of fit; (b) the presence of absence of localized areas of strain in the solution; (c) the interpretability, size, and statistical significance of the model’s parameter estimates (Brown, 2006; Mehta, 2011).
A number of omnibus tests can be utilized to assess how well the model fits the observed data. The $\chi^2$ statistic is a classic goodness-of-fit measure used to decide overall fit and is the most commonly used statistic for this purpose. Nevertheless, $\chi^2$ can be problematic due to its sensitivity to sample size (Brown, 2006; Jöreskog, 1969). Other alternative fit statistics have been created because of the disadvantage of the $\chi^2$. These goodness-of-fit indices can usually be categorized as comparative fit, absolute fit, and adjusting for model parsimony. Brown (2006) suggested that at least one index from each fit category should be assessed because each index offers exclusive information about the fit of the CFA solution (Mehta, 2011).

Absolute fit statistics assess the model fit at a complete level (Brown, 2006; Mehta, 2011). The $\chi^2$ is an example of an absolute fit. Fit indices adjusting for model parsimony encompass a penalty function for poor model parsimony (Brown, 2006; Mehta, 2011). When the model is not perfect, the root mean square error of approximation ($RMSEA$) is the distribution of the fitting function, which is a population-based statistic that depends on the uniform $\chi^2$ distribution. $RMSEA$ is slightly insensitive to the sample size but is sensitive to the number of parameters estimated. Value of zero of $RMSEA$ indicates a perfect fit.

Brown (2006) stated that “comparative fit indices evaluate the fit of a user-specified solution in relation to a more restricted, nested baseline model” (p. 84). Nested baseline model usually refers to a null or independence model in which the covariances among all input indicators are fixed to zero. Both the comparative fit index ($CFI$; Bentler, 1990) and Tucker-Lewis index ($TLI$; Tucker & Lewis, 1973) are popular $CFIs$ (Brown, 2006). The $TLI$, also known as non-normed fit index ($NNFI$), has features that compensate for the impact of model complexity, and the $NNFI$ encompasses a penalty function for adding freely estimated parameters that do not noticeably enhance the fit of the model. Because $NNFI$ values are
nonnormed, the values of **NNFI** can exceed the range of zero to one. Still, **NNFI** values are explained similarly to **CFI** as values coming close to one are considered a good fit (Brown, 2006).

Brown (2006) stated that numerous features of the solution including sample size, normality of data, type of data, model complexity, amount and type of misspecification, and estimation method can influence fit indices. Hu and Bentler (1999) conducted a study on the cutoff criteria and recommended the following guidelines for an acceptably good fit between the identified model and the observed data, assuming maximum likelihood (ML) estimation, is attained in cases when (a) **REMSEA** values are near to 0.06 or below, and (b) **CFI** and **NNFI** values are close to 0.95 or above.

A drawback of goodness-of-fit indices is that they only offer a global indication of the capability of the model to reproduce the variance-covariance matrix (Brown, 2006). These indices fail to provide information on the sources of inadequate model fit. Residuals and modification indices are two most commonly utilized statistics to address local areas of misfit in a CFA solution. The sample variance-covariance matrix, predicted variance-covariance matrix, and residual variance-covariance matrix are three frequently seen matrices in CFA modeling (Brown, 2006). The residual variance-covariance matrix addresses the difference between the sample and predicated variance-covariance matrix and offers detailed information regarding how well each variance and covariance was reproduced by the parameter estimates of the model (Brown, 2006; Mehta, 2011). These residuals are standardized due to the difficulty to explain them, are similar to standardized scores in a sampling distribution, and can be explained correspondingly by **z** scores. Usually, larger sample sizes are correlated to bigger standardized
residuals because of an inverse relationship between the size of the standard errors of the fitted residuals and the sample size (Brown, 2006; Mehta, 2011).

Researchers have used modification indices to examine particular relationships in a CFA solution (Brown, 2006). An index is calculated for each fixed and constrained parameter in the model. Brown (2006) expressed modification index as reflecting “an approximation of how much the overall model $\chi^2$ would decrease if the fixed or constrained parameter was freely estimated” (p. 119). Typically, a model that fits well generates small size modification indices. Like standardized residuals, modification indices are sensitive to sample size. Other than evaluating different indices described previously, reviewing the direction, magnitude, and significance of the parameter estimates of a CFA model is a necessary next step (Brown, 2006; Mehta, 2011). Brown (2006) stated that parameter estimates ought to make “statistical and substantive sense” (p. 126).

In order to further develop and strengthen the validity of the CIRAA, conducting CFA is a reasonable next step. Worthington and Whittaker (2006) stated that “researchers typically use CFA after an instrument has already been assessed using EFA, and they want to know if the factor structure produced by EFA fits the data from a new sample” (p. 815). Holliman and Ray (2013) also suggested that a reasonable next step for developing the CIRAA is to conduct CFA. Affirmative findings for CIRAA using CFA will provide credibility to the theoretical factor structure and strengthen the construct validity of this instrument.

Conclusion

Assessments used in play therapy research typically focus on behavioral changes of children, negating CCPT's philosophy of emphasizing holistic and internal changes occurring in
children (Holliman, 2010; Landreth, 2012). Specifically, CBCL (Achenbach & Rescorla, 2001) and PSI (Abidin, 1995) are two assessments that have been used most frequently for CCPT evaluation (Holliman, 2010; Lin, 2011). Holliman and Ray (2013) developed a theoretically-based instrument in order to enhance the studies of CCPT and move this discipline toward empirically supported treatment.

Because of the limitations and restrictions of assessments utilized in measurement of the effectiveness of CCPT, Holliman (2010) decided to design a theory-based assessment with reasonable reliability and validity in order to provide an instrument for CCPT and to establish a global assessment of CCPT effectiveness that could allow play therapy to be more easily studied. Holliman and Ray (2013) established and reported reasonable content validity, factor validity, criterion validity, predictive validity, and reliability of the CIRAA. Holliman and Ray (2013) also established a clinical threshold score of the CIRAA.

The factor structure of Holliman's CIRAA (i.e., Self-Control, Interpersonal Relationship, Coping Skills, and Internal Locus of Control) was found through exploratory factor analysis. The factor structure indicates that the CIRAA appears to be an assessment aligned with the objectives of CCPT. Each factor represents qualities that are beneficial for children and directly linked with the intended outcomes of CCPT. In order to develop and strengthen the validity of the CIRAA, Holliman and Ray (2013) suggested that CFA be conducted on a new sample of participants. CFA is a theory-driven statistic approach (Brown, 2006), that, with affirmative findings, can not only provide credibility for the factor structure of CIRAA but also strengthen the construct validity of the instrument.
CHAPTER 3
METHODS AND PROCEDURES

The primary purpose of this study was to examine Holliman’s (2010) findings of the factor structure of the Child Interpersonal Relationship and Attitudes Assessment (CIRAA). Holliman (2010) developed the CIRAA as a parent-report instrument congruent with the child-centered play therapy (CCPT) philosophy in order to further CCPT research. Holliman (2010) utilized exploratory factor analysis (EFA) and found the CIRAA measured four factors, including Self-Control, Interpersonal Relationship, Coping Skills, and Internal Locus of Evaluation. This study will employ confirmatory factor analysis (CFA) to examine the relationship between CIRAA items and the four factors (Brown, 2006). Additionally, I examined the relationship between CIRAA scores and child characteristics, as well as the ability of the CIRAA to distinguish between clinical and non-clinical samples.

Research Questions

This study answered the following research questions:

1. Will CFA support the factor structure of CIRAA (Self-Control, Interpersonal Relationship, Coping Skills, and Internal Locus of Evaluation) previously defined by the original EFA?

2. How are children’s characteristics, if any, such as age and gender, related to factors and total score on the CIRAA?

3. Will children from non-clinical and clinical samples differ on the CIRAA mean subscale scores (Self-Control, Interpersonal Relationship, Coping Skills, and Internal Locus of Evaluation) and total score?

Participants

In deciding upon a sample size, researchers suggest different guidelines for factor
analysis (Comrey, 1973; Gorsuch, 1983). Based on Velicer and Fava’s (1998) suggestion, 
variable ratios lower than 3:1 are often inadequate in factor analysis. Typically, a higher 
standard is held for CFA, usually requiring higher ratios than EFA. Comrey (1973) reported 
samples sizes as ranging from very poor ($n = 50$) to excellent ($n = 1000$). Gorsuch (1983) 
proposed a minimum participant to item ratio of 5:1, up to10:1. For the purposes of the current 
study, the criterion ratio of 5:1 was used.

The sample consisted of 206 parents of children from four locations in the southwest 
United States, resulting in an approximate 6:1 participant to item ratio. These locations included 
local elementary schools, two local counseling clinics, graduate level courses, and parent 
education groups. The 5 participating local elementary schools were eligible for Title 1 funds 
with over 40% of enrolled children from low-income families. All participating schools reported 
a majority of Hispanic or Caucasian students, and minority populations of African-American, 
Asian-American and other ethnicities. The local university-based counseling clinics provide 
services for the local community, and the majority of clients receiving services were Caucasian. 
Two different samples were collected. A clinical sample of 105 children was gathered from 
children referred for counseling services through the local school district or at local mental health 
clinics. A non-clinical sample of 101 children was recruited from parents who responded to the 
request of participation in local schools, graduate level classes, and parent education groups. 
The ages of the children ranged from 3 to 10 with the mean age being 6.14 years old. The age 
distribution of current sample is compatible to the sample used in Holliman and Ray’s (2013) 
exploratory factor analysis ($M = 6.5$ years old). Ethnic distribution was as follows: 125 (61%) 
Caucasian, 29 (14%) Hispanic, 14 (7%) African-American, 8 (4%) Asian, 12 (6%) Bi-racial, and
Table 3.1 presents detailed information of the sample’s demographic characteristics.

**Table 3.1**

*Demographic Characteristics of Participants (n = 202)*

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>131</td>
<td>63.6</td>
</tr>
<tr>
<td>Female</td>
<td>75</td>
<td>36.4</td>
</tr>
<tr>
<td><strong>Ethnicity</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>125</td>
<td>60.7</td>
</tr>
<tr>
<td>Hispanic</td>
<td>29</td>
<td>14.1</td>
</tr>
<tr>
<td>African-American</td>
<td>14</td>
<td>6.8</td>
</tr>
<tr>
<td>Asian</td>
<td>8</td>
<td>3.9</td>
</tr>
<tr>
<td>Bi-Racial</td>
<td>12</td>
<td>5.8</td>
</tr>
<tr>
<td>Non-Identified</td>
<td>18</td>
<td>8.7</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>13</td>
<td>6.3</td>
</tr>
<tr>
<td>4</td>
<td>20</td>
<td>9.7</td>
</tr>
<tr>
<td>5</td>
<td>38</td>
<td>18.4</td>
</tr>
<tr>
<td>6</td>
<td>48</td>
<td>23.3</td>
</tr>
<tr>
<td>7</td>
<td>48</td>
<td>23.3</td>
</tr>
<tr>
<td>8</td>
<td>23</td>
<td>11.2</td>
</tr>
<tr>
<td>9</td>
<td>13</td>
<td>6.3</td>
</tr>
<tr>
<td>10</td>
<td>3</td>
<td>1.5</td>
</tr>
<tr>
<td><strong>Grade</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not in School</td>
<td>8</td>
<td>3.9</td>
</tr>
<tr>
<td>Pre-Kindergarten</td>
<td>48</td>
<td>23.3</td>
</tr>
<tr>
<td>Kindergarten</td>
<td>48</td>
<td>23.3</td>
</tr>
<tr>
<td>Grade 1</td>
<td>42</td>
<td>20.4</td>
</tr>
<tr>
<td>Grade 2</td>
<td>41</td>
<td>19.9</td>
</tr>
<tr>
<td>Grade 3</td>
<td>13</td>
<td>6.3</td>
</tr>
<tr>
<td>Grade 4</td>
<td>5</td>
<td>2.4</td>
</tr>
<tr>
<td>Grade 5</td>
<td>1</td>
<td>.5</td>
</tr>
</tbody>
</table>
Instrument

CIRAA

CIRAA is a parent report instrument for parents with children from 3 to 10 years old. Parents respond to each of the 30 items via a 5-point Likert scale (Holliman, 2010). Holliman and Ray (2013) identified and defined four factors in the CIRAA. These factors are Self-Control, Interpersonal Relationships, Coping Skills, and Internal Locus of Evaluation. The internal consistency reliability is .93, which meets acceptability standards for the reliability of an instrument (Dimitrov, 2012; Nunnaly & Bernstein, 1994). The Cronbach's alpha coefficients for the subscales ranges from .72 to .91. The internal consistency reliabilities are respectable for the whole scale and subscales; and the inter-item correlations are acceptable for each factor (Holliman, 2010).

Concurrent validity has been obtained through assessing the correlation between the CIRAA total score and the PSI and the CBCL (items 6-18). The Pearson’s $r$ coefficients are .75 and .74 for the CIRAA and the CBC and the PSI, respectively. With ROC curve analysis, Holliman (2010) identified a score of 2.5 as the clinical cut-off score on the CIRAA.

The CIRAA was originally scored with higher scores indicating more problematic behaviors or attitudes resulted in higher scores. Lower scores indicated that children were exhibiting less problematic behaviors or attitudes. The directionality of scores appeared counter-intuitive in interpretation. Hence, I chose to reverse the scoring system so that higher scores indicate the presence of more desirable behaviors and attitudes. I made this modification prior to collection of data for CFA.
Child Behavior Checklist (CBCL)

The CBCL is a 120-item instrument (Achenbach & Rescorla, 2001) measuring parents’ views of children’s behavioral and emotional problems based on children’s social relationships, social activities, school performance, and behavior. The CBCL can be used with parents of children 6 to 18 years old and requires approximately 15 minutes to complete. The CBCL is composed of three competence subscales including Activities, Social, and School, along with a Total Competence score. The CBCL is further composed of eight syndrome subscales that represent a parent’s perception of a child’s behavior in the areas of Anxious/Depressed, Withdrawn, Somatic Complaints, Social Problems, Thought Problems, Attention Problems, Rule-Breaking Behaviors, and Aggressive Behavior. The CBCL generates a Total Problems score using a combination of the syndrome subscale scores. Furthermore, an Internalizing Problems score, which refers primarily to problems within the self, is created using the Anxious/Depressed, Withdrawn, and Somatic Complaints subscale scores. An Externalizing Problems score, which refers primarily to outwardly expressed problematic behaviors, is created using Rule-Breaking and Aggressive subscale scores.

The norm sample for the CBCL was based on a diverse sample composed primarily of children referred for clinical services and children exhibiting a high level of problematic behaviors. The test-retest reliability through Cronbach’s alpha coefficients for all of the subscales of the CBCL range from .63 and .97, with most being fairly strong and falling between .80 and .90. Achenbach and Rescorla (2001) supported the validity of the assessment by noting the differentiation in scores between children referred for clinical services and non-referred children. Details regarding the validity and reliability of the CBCL were presented in the Chapter 2 literature review.
Data Collection Procedures

Before proceeding with the collection of data, approval was obtained to conduct this research study from the University of North Texas Institutional Review Board. Copies of the 30 item instrument and the university approved informed consent form were copied and prepared for distribution. Parents were required to read and understand English in order to participate in the study.

Clinical Sample Procedures

In the elementary schools and local clinics, children identified as demonstrating behavioral or/and emotional concerns by parents, teachers, or counselors and referred for play therapy were sent packets including an informed consent document (see Appendix A), a letter explaining the study (See Appendix B), and the CIRAA (See Appendix C). Parents, recruited from the counseling clinics were also asked to complete a CBCL. Data collection procedures ensured that no parents in the current sample participated in Holliman's (2010) study.

Non-Clinical Sample Procedures

I recruited a non-clinical sample in order to expand the diversity of the sample. I visited graduate classes in a large university counseling program and parent education groups to recruit participants for the non-clinical sample. Additionally, I recruited non-clinical participants through the local elementary schools and counseling clinics by seeking out parents of non-referred children, those who did not participate in play therapy. Parents interested in the study were sent a packet including an informed consent document (see Appendix A), a letter explaining the study (See Appendix B), and the CIRAA (see Appendix C). Data collection
procedures ensured that no parents in the current sample participated in Holliman's (2010) study. All data from both the clinical and non-clinical participants were coded to ensure anonymity of participants' names.

Data Analyses

Confirmatory Factor Analysis

Confirmatory factor analysis (CFA) was conducted on CIRAA data provided by 206 parents of children aged 3 to 10 years in order to address Research Question 1: Will CFA support the factor structure of CIRAA (Self-Control, Interpersonal Relationship, Coping Skills, and Internal Locus of Evaluation) previously defined by the original EFA? Items and factors of CIRAA are presented in Table 3.2.

A preliminary descriptive statistical analysis was conducted in order to understand possible issues around missing data, scaling, collinearity issues, outlier detection, and multivariate normality. These issues need to be addressed before conducting the CFA. A very small percentage of item values were missing spread across 11 cases (.29%). A regression imputation was conducted to replace missing values with means. Multivariate normality was evaluated using Mardia’s multivariate kurtosis coefficient, which suggests normal distribution when the value is close to zero (Liu, 2009).
Table 3.2

**CIRAA Items and Factors**

<table>
<thead>
<tr>
<th>Factor 1: Self-Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. My child shares with other children</td>
</tr>
<tr>
<td>5. My child has verbal fights with other student at school *</td>
</tr>
<tr>
<td>7. My child apologizes for hurting other’s feelings</td>
</tr>
<tr>
<td>8. My child blames others for mistakes *</td>
</tr>
<tr>
<td>13. My child hits/kicks family members *</td>
</tr>
<tr>
<td>15. My child gets along with other children</td>
</tr>
<tr>
<td>18. My child exhibits self-control</td>
</tr>
<tr>
<td>20. My child disrupts family events/outings *</td>
</tr>
<tr>
<td>23. My child yells frequently *</td>
</tr>
<tr>
<td>25. My child often receives reports of disruptive behavior from his/her teacher *</td>
</tr>
<tr>
<td>26. My child often acts without thinking *</td>
</tr>
<tr>
<td>29. My child has physical fights with other children *</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Factor 2: Interpersonal Relationship</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. My child complains few people like him/her *</td>
</tr>
<tr>
<td>3. My child is often sad for prolonged periods of time *</td>
</tr>
<tr>
<td>9. My child is teased by other children *</td>
</tr>
<tr>
<td>11. My child complains other children are mean to him/her *</td>
</tr>
<tr>
<td>14. My child often compares him/herself unfavorably to others *</td>
</tr>
<tr>
<td>17. My child is overly sensitive *</td>
</tr>
<tr>
<td>22. My child complains no one likes him/her *</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Factor 3: Coping Skills</th>
</tr>
</thead>
<tbody>
<tr>
<td>4. My child is often upset by minor things *</td>
</tr>
<tr>
<td>6. My child becomes anxious over small matters *</td>
</tr>
<tr>
<td>19. My child is able to express his/her feelings when he/she is feeling upset</td>
</tr>
<tr>
<td>21. My child can name things he/she likes about him/herself</td>
</tr>
<tr>
<td>27. For his/her age, my child is able to enter new situations with confidence</td>
</tr>
<tr>
<td>28. My child can calm down when upset</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Factor 4: Internal Locus of Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>10. My child enjoys doing things for him/herself</td>
</tr>
<tr>
<td>12. My child volunteers to help out around the house</td>
</tr>
<tr>
<td>16. My child shows me projects he/she is proud of</td>
</tr>
<tr>
<td>24. My child talks about accomplishments of which he/she is proud</td>
</tr>
<tr>
<td>30. My child tells me things he/she is proud of</td>
</tr>
</tbody>
</table>

*Note. * indicates negatively worded item where scoring was reversed.
**CFA Models**

The next step is to specify the model, that is, the factor structure proposed by Holliman and Ray (2013). Thompson (2004) suggested testing rival models when conducting CFA because a preferred model is more impressive when the fit happens in the context of evaluating several models, particularly when competitive models are theoretically plausible. The primary model (Model 1) for CIRAA with CFA contains four uncorrelated factors: self-control, interpersonal relationship, coping skills, and internal locus of evaluation (see Figure 3.1), which replicated the EFA model identified by Holliman and Ray (2013).

A competing model was generated, which allowed factors to correlate with each other (Model 2), based on the logic that some correlation is plausible among factors (see Figure 3.2). Thus, Model 1 is the primary model and Model 2 is the rival model.

Then, CFA was conducted in LISREL 8.8 to estimate the parameters. The main strength of CFA is its ability to estimate parameters, such as factor loadings, factor variances, and covariance, and to indicate error variances (Brown, 2006; Mehta, 2011), ensuring the predicted covariance matrix is as near as possible to the sample covariance matrix. Robust maximum likelihood (MLM) is the fitting function utilized for deciding the closeness of the inferred asymptotic covariance matrix to the sample asymptotic covariance matrix. Model convergence was gathered when the LISREL 8.8 demonstrated a set of parameter estimates that cannot be enhanced upon further to lessen the difference between the predicted asymptotic covariance matrix and the sample asymptotic covariance matrix.
Figure 3.1. CFA Model 1: Uncorrelated four factor solution.
Figure 3.2. CFA Model 2: Correlated four factor solution.

Model Evaluation

Brown (2006) suggested three stages for model evaluation: (a) overall goodness-of-fit; (b) localized areas of strain; and (c) interpretability, size, and statistical significance of the
parameter estimates. Brown (2006) recommended researchers utilize at least one index from the absolute fit, comparative fit, and adjusting for model parsimony categories for the overall goodness of fit. The overall fit indices utilized were $\chi^2$, RMSEA, CFI, and NNFI. When using MLM estimation, the following guidelines are recommended for acceptable goodness-of-fit: (1) RMSEA values are approximate to .06-.08 or below and (2) CFI and NNFI are close to 0.95 or above (Hu & Bentler, 1999; Mehta, 2011; Schreiber et al., 2006). Furthermore, residuals and modification indices were calculated to detect local areas of misfit in the CFA solution. Finally, the magnitude, direction, and significance of the parameter estimates were studied for making “statistical and substantive sense” (Brown, 2006, p. 126).

Multiple Regressions

In order to address Research Questions 2: How are children's characteristics, if any, such as age and gender, related to factors and total score on the CIRAA? I conducted multiple regression analysis to explore the relationship between children's characteristics, such as age and gender, and CIRAA’s four factor scores and total score. Preliminary assumption testing was conducted before the multiple regression analysis to check for normality, linearity, homoscedasticity, multicollinearity, singularity, and outliers in order to deal with possible concerns (Pallant, 2013). Basic descriptive data, histograms, and Q-Q plots were consulted to ensure the data met the requirement of normality (Newton & Rudestam’s, 1999). No major violations of these assumption occurred, five multiple regression analyses were conducted in order to understand the relationships between age and gender, and the four CIRAA factors (Self-Control, Interpersonal Relationship, Coping Skills, and Internal Locus of Evaluation) and the CIRAA total score.
Through the process of multiple regression analyses, the standardized regression coefficients (beta weights or $\beta$) were utilized for direct comparisons between coefficients. The beta weights were utilized in evaluating how effectively a predictor variable impacted the criterion variable (Thompson, 1992). The statistical significance value of $p$ must equal or be less than .05. Additionally, the structure coefficient ($r_s$) is the bivariate correlation indicating the degree of relationship between a given predictor variables and the synthetic variable, predicted $Y$ or $\hat{Y}$.

To be exact, the structure coefficient is a Pearson’s $r$ between observed variables and synthetic variables. The structure coefficient offers information concerning the practical significance, or utility, of the results (Courville & Thompson, 2001). The squared structure coefficient ($r_s^2$) shows the amount of variability that each predictor variable can explicate in the presence of the other composed predictor variables. Because reporting beta weights ($\beta$) alone might, in some cases, lead to insufficient reporting of significant results, use of both beta weights ($\beta$) and structure coefficients ($r_s$) are crucial to accurately interpreting regression results (Courville & Thompson, 2001; Thompson, 1992).

**Independent-Sample $t$-Tests**

In order to address Research Question 3: Will children from non-clinical and clinical samples differ on the CIRAA mean subscale scores (Self-Control, Interpersonal Relationship, Coping Skills, and Internal Locus of Evaluation) and total score? I conducted independent sample $t$-tests for the clinical and non-clinical sample groups.

Before conducting the independent-sample $t$-test, the sample scores were evaluated for normal distribution, homogeneity of variance through Levene’s test for equality of variance, and
missing data. Five independent sample t-tests were conducted to see if non-clinical and clinical samples scored differently on four subscales (Self-Control, Interpersonal Relationship, Coping Skills, and Internal Locus of Evaluation) and total scores.

Pintea (2010) stated that “statistical significance does not provide information about the magnitude of change” (p. 102). In addition, it is highly plausible, if a study includes a large sample size, to have a statistically significant result yet small magnitude of change, such as small effect size. APA Publication Manual (2010) asserted that in order for the reader “to appreciate the magnitude or importance of a study’s findings, it is almost always necessary to include some measure of effect size in the Results section” (p. 34). I used eta squared, ranging from 0 to 1, to interpret practical significance of results. Eta squared represents “the proportion of variance in the dependent variable that is explained by the independent (group) variable” (Pallant, 2013, p. 242).
CHAPTER 4

RESULTS

In this chapter, I present the results of the CFA, multiple regression analyses, and t-tests to address research questions. This study offered the opportunity to ensure the theoretical soundness of CIRAA as a validated play therapy assessment. Holliman (2010) utilized exploratory factor analysis (EFA) and found the instrument measured the four factors of Self-Control, Interpersonal Relationship, Coping Skills, and Internal Locus of Evaluation of CIRAA. This study incorporated confirmatory factor analysis (CFA) to examine the relationship between CIRAA items and the four factors (Brown, 2006). I also explored the relationships between child characteristics and CIRAA scores, as well as differences between clinical and non-clinical samples.

Research Question 1: Will CFA Support the Factor Structure of CIRAA (Self-Control, Interpersonal Relationship, Coping Skills, and Internal Locus of Evaluation) Previously Defined by Original EFA?

For Research Question 1, a CFA was conducted. Before conducting CFA, item values were examined for missing data, normality, and multivariate normality. The sample consisted of 206 parents who completed the CIRAA for the analyses. Eleven cases with missing values (.29%) were identified. Regression imputation was used to replace missing data. All available data were utilized in the later analysis through this approach. Then, item means and standard deviations were computed to examine for normality. The item means ranged from 2.78 to 4.55, and their standard deviations ranged from 0.660 to 1.362.

Multivariate normality was assessed using Mardia's kurtosis coefficient. Mardia's coefficient of kurtosis suggested multivariate non-normality at 182.816. The critical ratio was
Liu (2009) suggested that the assumption of multivariate normality is met when the kurtosis and critical ratio values approach zero due to Likert scaling measurements being more likely to yield data that violate the multivariate normality assumption.

Because the maximum likelihood (ML) estimator is appropriate only when the data are interval-type and achieve multivariate normality (Brown, 2006), the robust maximum likelihood (MLM) estimator was used to compensate for the data's multivariate non-normality. Chou and Bentler (1995) stated that MLM provides a “well-behaved” estimator when dealing with different levels of non-normality, sample size, and model complexity (as cited by Brown, 2006, p. 379). Brown (2006) declared that “the MLM provided ML parameter estimates with standard errors and a mean-adjusted $X^2$ test statistic that are robust to non-normality” (p. 379). The mean-adjusted $X^2$ test statistic is also known as the Santorra-Bentler scaled $X^2$ (SB $X^2$; Santorra & Bentler, 1994). A covariance matrix and an asymptotic covariance matrix were generated with PRELIS software (Jöreskog & Sörbom, 2006) for the subsequent CFA analysis.

Brown suggested that “the acceptability of the fitted CFA solution should be examined on the basis of three major components: (1) overall goodness of fit; (2) the presence or absence of localized areas of strain in the solution (i.e., specific points of ill fit); and (3) the interpretability, size, and statistical significance of the model's parameter estimates” (p. 113). Goodness-of-fit indices only offer a global descriptive summary of the capability of the model to reproduce the input covariance matrix. The other two components of fit evaluation supply more explicit information about the “acceptability” and “utility” of the solution (Brown, 2006, p. 113).
Model 1: Uncorrelated Four Factor Solution

*Overall Goodness of Fit*

Brown (2006) suggested at least one index from each fit category, such as absolute fit, parsimony correction, and comparative fit should be consulted because each category offers different information about the fit of the CFA solution. Based on this suggestion, the goodness of fit indices utilized to examine the model fit were $X^2$, root mean square error of approximation ($RMSEA$), the comparative fit index ($CFI$; Bentler, 1990), and non-normed fit index ($NNFI$).

For Model 1, the absolute fit index was $X^2(405) = 1,049.66$, $p < .01$, indicating statistical significance. Although $X^2$ is the classic goodness-of-fit index, it is inflated by sample size (Brown, 2006). Therefore, large sample size solutions are consistently rejected on the basis of $X^2$ even when differences between predicted covariance matrix and input matrix are negligible. Moreover, $X^2$ is interpreted according to a very stringent hypothesis that predicted covariance matrix equals input matrix. While $X^2$ is consistently reported in CFA research, other fit indices often take on more credibility in the evaluation of model fit (Brown, 2006). $RMSEA$ evaluates the degree to which a CFA model fits reasonably well in the population.

The $RMSEA$ of the Model 1 was .088 and slightly higher than .06 to .08 criteria of fit, with the 90% confidence interval ($CI$) ranging from .082 to .095. The $RMSEA$ analysis suggested borderline model fit. For comparative fit, the $CFI$ and $NNFI$ were both .95 and suggested acceptable model fit (Bentler, 1990; Brown, 2006; Hu & Bentler, 1999).

*Localized Areas of Strain*

When checking the localized areas of strain using modification indices for Model 1, $X^2$ decreased 113.9, and the new estimate was .88 when Self-Control and Coping Skills were
allowed to correlate. Brown (2006) stated that “the modification index reflects an approximation of how much the overall model $X^2$ would decrease if the fixed or constrained parameter was freely estimated” (p. 119). Because the modification index provides a $X^2$ statistic with 1 degree of freedom ($df$), indices of 3.84 or greater imply that the overall fit of the model could be significantly enhanced ($p < .05$) if the fixed or constrained parameter is estimated freely. At the same time, expected parameter change ($EPC$) values for each modification index need to be consulted. $EPC$ is used to interpret how much the parameter is presumed to change in a positive or negative direction when estimating the parameter freely in the analysis. When allowing the Self-Control and Coping Skills to correlate, the goodness of model fits were $X^2$ of 928.59; $RMSEA$ of .08, with the 95% $CI$ ranging from .073 to .086; $NNFI$ of .96; and $CFI$ of .96, which showed negligible improvement. Moreover, Brown (2006) cautioned that the parameter “should not be freed with the sole intent of improving model fit” (p. 123).

Brown emphasized that revisions of a model should always focus exclusively on parameters justified by prior evidence or theory. Self-control can be considered as one type of coping skill (Kurman, Hui, & Dan, 2010; Mezo, 2009; Smith, 1989). Mezo (2009) stated that a cognitive-behavioral coping skills model signifying self-control and self-management skills has been successfully utilized in assessment and treatment. With this concept in mind, allowing the Self-Control and Coping Skills to correlate can be a revision to the Model 1 solution. However, Brown (2006) stressed that “even when revisions to the initial solution can be defended with compelling empirically or theoretically based arguments, it is important to note that by pursuing respecifications of the initial model, one has moved out of a confirmatory analytic framework” (p. 124). Because the purpose of the study was to confirm the uncorrelated four-factor model, I found the borderline to acceptable goodness of model fits on three out of four indices other than
the \( X^2 \) for Model 1, and the minor improvements to the goodness of model fits when allowing two factors to correlate; therefore, I did not develop a revised model following the production of the modification indices.

Brown (2006) stated that models can be revised by deleting statistically nonsignificant parameters, and the necessity of existing parameters can be assessed by evaluating their statistical significance. Parameters with \( z \) values equal to or greater than \( \pm 1.96 \) are statistically significant. All parameters were statistically significant; therefore, it was reasonable to conclude that no parameters were unnecessary for the model.

**Interpretability, Size, and Statistical Significance of the Model's Parameter Estimates**

With acceptable goodness-of-fit and the rationale for retaining Model 1 after consulting with the modification indices, the model evaluation can be used to address direction, magnitude, and significance of parameter estimates (Brown, 2006). The first step is to decide if the parameter estimates make both substantive and statistical sense. Brown (2006) stated that the parameter estimates “should not take on out-of-range values such as completely standardized factor correlations that exceed 1.0, negative factor variances, or negative indicator error variances” (p. 126). From a statistical perspective, if out-of-range values, known as Heywood cases or offending estimates, are found, model specification errors, problems with sample, or problems with model-implied matrices may have occurred. In the results, no out-of-range values were found as part of reviewing the completely standardized factor parameter correlations, factor variance, and indicator variances of each item. Results indicated that all parameter estimates from the Model 1 solution were statistically viable (Brown, 2006). From a substantive
perspective, the direction of the parameter estimates was consistent with the prediction. The items and their identified factors were positively correlated.

The next step requires evaluating the standard errors of the parameter estimates to decide if their magnitude is suitable (Brown, 2006). Standard errors signify estimates of how much sampling error is working in the model's parameter estimates. No specific guidelines are obtainable in deciding if the magnitude of standard errors is too large or too small (i.e., problematic) in a given data set. However, Brown (2006) recommended being alarmed about observing standard errors with standout or close to zero values and parameter estimates that seem reasonably large but lack statistical significance.

The standard errors among the results ranged from .032 to .082. Brown (2006) stated that even though small standard errors might suggest substantial precision in parameter estimates, the significance test of the parameter cannot be calculated if the standard error is close to zero. Because significance test (z test) of the parameter was calculated for each item, it was reasonable to conclude that these standard errors were not too small. All parameter estimates were statistically significant. Therefore, no problematic standard errors were found in the data, suggesting considerable precision for the parameter estimates.

In CFA models with no cross-loading indicators, the pattern coefficients for completely standardized factor loading can be explained as the correlation between the indicator and the latent factor. Therefore, when squaring the pattern coefficients, the proportion of variance of the indicator explained by the factor is obtained and is known as communality (Brown, 2006). This information along with the factor correlations provides information about the data's reliability and discriminant validity. (A more detailed discussion of the reliability and validity are
presented in the discussion section.) The information regarding completely standardized parameter estimate, test statistic (z value), standard error, and $R^2$ are presented in the Table 4.1.

Table 4.1

*Parameter Estimate from the CFA Model 1*

<table>
<thead>
<tr>
<th>Model Result</th>
<th>S.E.</th>
<th>Est./S.E</th>
<th>Errovar</th>
<th>StdYX</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Self-Control</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CI1</td>
<td>.063</td>
<td>9.51</td>
<td>.64</td>
<td>.60*</td>
<td>.36</td>
</tr>
<tr>
<td>CI5</td>
<td>.045</td>
<td>16.43</td>
<td>.45</td>
<td>.74*</td>
<td>.55</td>
</tr>
<tr>
<td>CI7</td>
<td>.072</td>
<td>5.36</td>
<td>.85</td>
<td>.39*</td>
<td>.15</td>
</tr>
<tr>
<td>CI8</td>
<td>.055</td>
<td>11.13</td>
<td>.63</td>
<td>.61*</td>
<td>.37</td>
</tr>
<tr>
<td>CI13</td>
<td>.054</td>
<td>10.92</td>
<td>.65</td>
<td>.59*</td>
<td>.35</td>
</tr>
<tr>
<td>CI15</td>
<td>.058</td>
<td>11.62</td>
<td>.55</td>
<td>.67*</td>
<td>.45</td>
</tr>
<tr>
<td>CI18</td>
<td>.050</td>
<td>14.84</td>
<td>.45</td>
<td>.74*</td>
<td>.55</td>
</tr>
<tr>
<td>CI20</td>
<td>.040</td>
<td>19.35</td>
<td>.40</td>
<td>.77*</td>
<td>.60</td>
</tr>
<tr>
<td>CI23</td>
<td>.051</td>
<td>13.21</td>
<td>.55</td>
<td>.67*</td>
<td>.45</td>
</tr>
<tr>
<td>CI25</td>
<td>.037</td>
<td>20.30</td>
<td>.44</td>
<td>.75*</td>
<td>.56</td>
</tr>
<tr>
<td>CI26</td>
<td>.037</td>
<td>20.83</td>
<td>.39</td>
<td>.78*</td>
<td>.61</td>
</tr>
<tr>
<td>CI29</td>
<td>.050</td>
<td>14.42</td>
<td>.47</td>
<td>.73*</td>
<td>.53</td>
</tr>
<tr>
<td><strong>Interpersonal Relationship</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CI2</td>
<td>.032</td>
<td>27.04</td>
<td>.24</td>
<td>.87*</td>
<td>.76</td>
</tr>
<tr>
<td>CI3</td>
<td>.048</td>
<td>14.93</td>
<td>.50</td>
<td>.71*</td>
<td>.50</td>
</tr>
<tr>
<td>CI9</td>
<td>.056</td>
<td>11.58</td>
<td>.57</td>
<td>.65*</td>
<td>.43</td>
</tr>
<tr>
<td>CI11</td>
<td>.037</td>
<td>20.68</td>
<td>.41</td>
<td>.77*</td>
<td>.59</td>
</tr>
<tr>
<td>CI14</td>
<td>.053</td>
<td>13.13</td>
<td>.51</td>
<td>.70*</td>
<td>.49</td>
</tr>
<tr>
<td>CI17</td>
<td>.060</td>
<td>8.43</td>
<td>.75</td>
<td>.50*</td>
<td>.25</td>
</tr>
<tr>
<td>CI22</td>
<td>.040</td>
<td>22.36</td>
<td>.22</td>
<td>.88*</td>
<td>.78</td>
</tr>
<tr>
<td><strong>Coping Skills</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CI4</td>
<td>.070</td>
<td>8.14</td>
<td>.68</td>
<td>.57*</td>
<td>.32</td>
</tr>
<tr>
<td>CI6</td>
<td>.074</td>
<td>7.74</td>
<td>.67</td>
<td>.57*</td>
<td>.33</td>
</tr>
<tr>
<td>CI19</td>
<td>.075</td>
<td>6.99</td>
<td>.72</td>
<td>.53*</td>
<td>.28</td>
</tr>
<tr>
<td>CI21</td>
<td>.063</td>
<td>10.19</td>
<td>.58</td>
<td>.64*</td>
<td>.42</td>
</tr>
<tr>
<td>CI27</td>
<td>.070</td>
<td>8.91</td>
<td>.61</td>
<td>.62*</td>
<td>.39</td>
</tr>
<tr>
<td>CI28</td>
<td>.060</td>
<td>11.93</td>
<td>.48</td>
<td>.72*</td>
<td>.52</td>
</tr>
<tr>
<td><strong>Internal Locus of Evaluation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CI10</td>
<td>.059</td>
<td>9.53</td>
<td>.69</td>
<td>.56*</td>
<td>.31</td>
</tr>
<tr>
<td>CI12</td>
<td>.082</td>
<td>4.32</td>
<td>.88</td>
<td>.35*</td>
<td>.12</td>
</tr>
<tr>
<td>CI16</td>
<td>.071</td>
<td>11.29</td>
<td>.36</td>
<td>.80*</td>
<td>.64</td>
</tr>
<tr>
<td>CI24</td>
<td>.066</td>
<td>12.32</td>
<td>.34</td>
<td>.81*</td>
<td>.66</td>
</tr>
<tr>
<td>CI30</td>
<td>.042</td>
<td>21.61</td>
<td>.18</td>
<td>.90*</td>
<td>.82</td>
</tr>
</tbody>
</table>

*Note. S.E. = standard error; Est./S.E. = test statistic (z value); Errovar= Error Variance; StdYX = completely standardized parameter estimate (pattern coefficient). *p < .001.*
Model 2: Correlated Four Factor Solution

Model 2 allowed the four factors of the CIRAA to correlate with each other. For the solution, $X^2$ equaled 828.05, $p < .05$; $RMSEA$ was .072, with the 95\% CI ranging from .065 to .079; $NNFI$ equaled .96; and $CFI$ was .97. The goodness-of-fit indices for Model 1 and 2 are presented in Table 4.2. Model 2 showed minor improvements on goodness-of-fit indices; however, the differences were negligible.

Table 4.2

<table>
<thead>
<tr>
<th></th>
<th>$X^2$</th>
<th>RMSEA</th>
<th>NNFI</th>
<th>CFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>1049.66</td>
<td>.088</td>
<td>.95</td>
<td>.95</td>
</tr>
<tr>
<td>Model 2</td>
<td>828.05</td>
<td>.072</td>
<td>.96</td>
<td>.97</td>
</tr>
</tbody>
</table>

* $p < .001$

Thompson (2004) stated that, in EFA, researchers often use orthogonal rotations to provide sample structure. Nevertheless, correlated factors are usually expected in CFA and “almost always” offer a better fit to data (p. 118). Therefore, the Model 2 was expected to have a better goodness-of-fit in comparison to the Model 1. Thus, Model 1 was chosen and confirmed as the primary measurement model based on the theoretical (Axline, 1974; Landreth, 2012) and empirical (Holliman & Ray, 2013) support for the four-factor model of the CIRAA.

Research Question 2: How are Children's Characteristics, if any, such as Age and Gender, related to Factors and Total Score on the CIRAA?

For Research Question 2, five multiple regressions were conducted to see how children's ages and genders predicted performance on the average scores from each of the four factors of Self-Control, Interpersonal Relationship, Coping Skills, and Internal Locus of Evaluation and for CIRAA total score. Tabachnick and Fidell (2007) provided a formula for estimating sample size
requirements when conducting social science research. Based on the number of predictors, this analysis required a minimum of 66 participants. From the power analysis, the study needed 107 participants for obtaining a power of .95 with an effect size of .15, given the number of predictors summed to two. The sample included 206 parents who filled out the CIRAA. Eighteen missing values were found in the original data across 11 cases. Missing values were replaced with the same procedure used for the CFA, using means from regression imputation. The sample of 206 parents was utilized in the following analyses.

Before performing any statistical analysis for the research question, preliminary assumptions testing was conducted to check for normality, linearity, homoscedasticity, multicollinearity, singularity, and outliers (Pallant, 2013). Continuous variables were age; average scores of Self-Control, Interpersonal Relationship, Coping Skills, Internal Locus of Evaluation; and total scores. All variables’ data showed reasonably straight lines in the normal Q-Q plots and presented no strong violations of kurtosis and skewness. The variables demonstrated normal distributions. Tabachnick and Fidel (2007) also suggested that with reasonably large sample size, more than 200 cases, the skewness and kurtosis statistics fail to show a substantive influence in an analysis.

The variables’ boxplots displayed no outliers. When comparing the differences between the 5% trimmed mean and the original mean, the two mean values were similar to each other for all of the variables of interest. Therefore, no extreme values were found among the data as likely to impact the results. The normal probability plot of the regression standardized residuals and the scatterplots between age and the scores for Self-Control, Interpersonal Relationship, Coping Skills, Internal Locus of Evaluation, and whole total scores were consulted to determine normality, linearity, homoscedasticity, and outliers. The data met the necessary assumptions
without any serious violations. The result of collinearity diagnostics did not indicate the presence of multicollinearity, and the correlations between age and gender \( (r = .011) \) was low. Therefore, the original data were used for the multiple regression analyses (see Table 4 for the results of the five regression analyses).

The first multiple regression analysis showed that age and gender together explained only 2.1% of the variance for Self-Control scores \( (F[2, 203] = 2.219, p = .111) \). Results did not yield statistical significance between the age and gender predictors and Self-Control (see Table 4.3).

Table 4.3

<table>
<thead>
<tr>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>2.729</td>
<td>2</td>
<td>1.365</td>
<td>2.219</td>
</tr>
<tr>
<td>Residual</td>
<td>124.846</td>
<td>203</td>
<td>.615</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>127.575</td>
<td>205</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Predictors are age and gender. Dependent variable is self-control. *p < .05.

The second regression analysis demonstrated age and gender as explaining 6.1% of the variance for Interpersonal Relationship scores \( (F[2, 203] = 6.290, p < .05) \). Even though the result was statistically significant between the two predictors and Interpersonal Relationship scores, the two predictors only accounted for 6.1% of the variance of Interpersonal Relationship scores, with even smaller adjusted \( R^2 \) of .052. When consulted with beta weights and squared structure coefficients of age and gender, age alone explained 5.3% of the variance of the Interpersonal Relationship scores and gender explained .8%. From a practical perspective, the relationship between the two predictors and Interpersonal Relationship scores was small, and the statistically significant result might be attributed to a large sample size, indicating that age and gender were not strong predictors of Interpersonal Relationship Scores.
Table 4.4

Summary of the Regression of Two Predictors and Interpersonal Relationship.

<table>
<thead>
<tr>
<th></th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>9.027</td>
<td>2</td>
<td>4.513</td>
<td>6.59</td>
<td>.002*</td>
</tr>
<tr>
<td>Residual</td>
<td>139.034</td>
<td>203</td>
<td>.685</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>148.061</td>
<td>205</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Predictors are age and gender. Dependent variable is interpersonal relationship. *p < .01.

The third analysis indicated that age and gender together accounted for 2.8% of the variance in Coping Skills scores ($F[2, 203] = 2.873, p = .059$; see Table 4.5). This result was not statistically significant.

Table 4.5

Summary of the Regression of Two Predictors and Coping Skills.

<table>
<thead>
<tr>
<th></th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>3.005</td>
<td>2</td>
<td>1.502</td>
<td>2.873</td>
<td>.059</td>
</tr>
<tr>
<td>Residual</td>
<td>106.152</td>
<td>203</td>
<td>.523</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>109.156</td>
<td>205</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Predictors are age and gender. Dependent variable is coping skills.

The fourth regression analysis showed that age and gender together accounted for 3.7% of the variance in Internal Locus of Evaluation scores ($F[2, 203] = 3.935, p < .05$; see Table 4.6). Even though the result was statistically significant, the two predictors only explained 3.7% of the variance for Internal Locus of Evaluation scores. The smaller adjusted $R^2$ of .028 indicated the result lacked practical significance. When consulted with beta weights and squared structure coefficients of two predictors, age and gender accounted for 1.6% and 2.1% of the variance of the Internal Locus of Evaluation scores, respectively, indicating that age and gender were not strong predictors of Internal Locus of Evaluation.
Table 4.6

Summary of the Regression of Two Predictors and Internal Locus of Evaluation

<table>
<thead>
<tr>
<th></th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>2.900</td>
<td>2</td>
<td>1.450</td>
<td>3.935</td>
<td>.021*</td>
</tr>
<tr>
<td>Residual</td>
<td>74.809</td>
<td>203</td>
<td>.369</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>77.709</td>
<td>205</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Predictors were age and gender. Dependent variable was internal locus of evaluation. *p < .05.

The last regression analysis showed age and gender as explaining 4.2% of the variance of the total scores ($F[2, 203] = 4.436, p < .05$; see Table 4.7). Once again, although the result was statistically significant, age and gender only accounted for 4.2% of the variance of the total scores. The adjusted $R^2$ of .032 indicated the result lacked practical significance. When consulted with beta weights and squared structure coefficients of two predictors, age and gender accounted for 2.1% and 2.1% of the variance of the total scores, respectively, indicating that age and gender were not strong predictors of CIRAA total scores.

Table 4.7

Summary of the Regression of Two Predictors and Total Score.

<table>
<thead>
<tr>
<th></th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>3.541</td>
<td>2</td>
<td>1.770</td>
<td>4.421</td>
<td>.013*</td>
</tr>
<tr>
<td>Residual</td>
<td>81.284</td>
<td>203</td>
<td>.400</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>84.825</td>
<td>205</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Predictors are age and gender. Dependent variable is total score. *p < .05.

Even though the regression models for the two predictors of age and gender with Interpersonal Relationship, Coping Skills, and total scores yielded statistical significance, these results were not practically significant due to the small magnitude of variance accounted for by each of the model’s $R^2$. The adjusted $R^2$ (Ezekiel, 1932) for Self-Control, Interpersonal Relationship, Coping Skills, Internal Locus of Evaluation, and total scores were very small 0.9%,
0.9%, 1%, 0.9%, and 1% less than model $R^2$ values, respectively. Each difference between the $R^2$ and adjusted $R^2$ indicated a small reduction due to theoretical correction for sampling error in each model. The small difference between $R^2$ and adjusted $R^2$ might have occurred because of large sample size and indicated a minimal level of sampling error.

In sum, age and gender together only accounted for a negligible amount of variance in the four subscales scores and the total scores, indicating the applicability of the CIRAA for both genders and range of ages from 3 to 10 years old.

Research Question 3: Will Children from Non-Clinical and Clinical Samples Differ on the CIRAA Mean Subscale Score (Self-Control, Interpersonal Relationship, Coping Skills, and Internal Locus of Evaluation and Total Score)?

In order to answer Research Question 3, five independent samples $t$-tests were conducted to examine differences between four subscales and total scores for non-clinical and clinical sample. From the power analysis, the study needed 70 participants for power of .95 to yield a large effect size of .08. Sample scores’ normality, homogeneity of variance, and missing data were evaluated before conducting the independent-samples $t$-tests. Missing data were replaced by means, as with the prior statistical tests of the data. The five dependent variables of Self-Control, Interpersonal Relationship, Coping Skills, Internal Locus of Evaluation, and total scores met the assumptions for normality and homogeneity. Because the Levene’s tests for equality of variance were statistically non-significant ($p < .05$) for all five analyses, the variance between the non-clinical and clinical groups that were tested against each dependent variable was deemed equal.

The results showed that scores for non-clinical and clinical samples yielded statistically significant differences on three subscales and on total score; however the Internal Locus of
Evaluation did not show a statistical difference between these groups. Table 4.8 presents the means, standard deviations, and $t$-test results for the two groups (non-clinical/clinical) on the four subscales and total score scale.

Table 4.8

*Independent Sample $t$-Test on Four Subscales and Total Score on Two Groups*

<table>
<thead>
<tr>
<th>Variables</th>
<th>$n$</th>
<th>$M$</th>
<th>$SD$</th>
<th>$t$ score</th>
<th>$\eta^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-Control</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Clinical</td>
<td>101</td>
<td>3.81</td>
<td>.683</td>
<td>-5.322*</td>
<td>.120</td>
</tr>
<tr>
<td>Clinical</td>
<td>105</td>
<td>3.26</td>
<td>.792</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interpersonal Relationship</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Clinical</td>
<td>101</td>
<td>3.94</td>
<td>.750</td>
<td>-6.022*</td>
<td>.150</td>
</tr>
<tr>
<td>Clinical</td>
<td>105</td>
<td>3.28</td>
<td>.818</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coping Skills</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Clinical</td>
<td>101</td>
<td>3.65</td>
<td>.693</td>
<td>-5.305*</td>
<td>.120</td>
</tr>
<tr>
<td>Clinical</td>
<td>105</td>
<td>3.15</td>
<td>.678</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Internal Locus of Evaluation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Clinical</td>
<td>101</td>
<td>4.21</td>
<td>.599</td>
<td>-0.803</td>
<td>.003</td>
</tr>
<tr>
<td>Clinical</td>
<td>105</td>
<td>4.14</td>
<td>.633</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Clinical</td>
<td>101</td>
<td>116.24</td>
<td>17.600</td>
<td>-5.851*</td>
<td>.140</td>
</tr>
<tr>
<td>Clinical</td>
<td>105</td>
<td>101.65</td>
<td>18.170</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*$p < .001$

Eta squared ($\eta^2$) coefficients were computed in order to understand the magnitude of change (Pintea, 2010). All subscales, except the Internal Locus of Evaluation ($\eta^2 = .003$), showed large effect sizes. The magnitude of the differences in the means between the clinical and non-clinical samples was very large for Self-Control ($\eta^2 = .12$), Interpersonal Relationship ($\eta^2 = .15$), Coping Skills ($\eta^2 = .12$), and total score ($\eta^2 = .14$). In sum, non-clinical and clinical samples do score differently on three of the subscales and total score scale of the CIRAA, with
the one exception to these results being the Internal Locus of Evaluation subscale.

**Internal Consistency Reliability**

Nunnally and Bernstein (1994) stated that establishing reliability is crucial to any measurement method, and an instrument’s reliability should be examined when developing new instruments. Researchers want to discern that items within the instrument are consistent (Gliner, Morgan, & Harmon, 2001; Nunnally & Bernstein, 1994). Previously, Holliman and Ray (2013) utilized Cronbach’s alpha (α) to investigate CIRAA’s overall reliability. Results suggested reasonable reliabilities for both total score and subscales. I calculated the Cronbach’s α again and analyzed the inter-item correlations with the sample of 206 cases.

The alpha for the CIRAA was established to be .928. Nunnally and Bernstein (1994) reported the Cronbach’s α of greater than .9 not only as an adequate level of reliability but also as appropriate for making clinical decisions about individuals (Gliner et al., 2001). The Cronbach’s α for the Self-Control subscale yielded .88, indicating an appropriate level of reliability. Inter-item correlations for Self-Control ranged from .043 to .629. It is worth noting that two lowest inter-item correlations occurred for Item 7 with Item 5 and Item 7 with Item 25. Other than these two number pairs, the lowest inter-item correlation was .238 (see Table 4.9). Clark and Watson (1995) suggested that the average inter-item correlation should fall in the range of .15 to .50 (Briggs & Cheek, 1986). Generally, Cronbach’s α reliability coefficients for subscales that are lower than for the whole instrument’s reliability are usually considered acceptable because of the general relationship between instrument length and α values (Clark & Watson, 1995; Holliman & Ray, 2013).
Table 4.9

*Inter-item Correlations for CIRAA Factor 1: Self-Control*

<table>
<thead>
<tr>
<th>Item</th>
<th>1</th>
<th>5</th>
<th>7</th>
<th>8</th>
<th>13</th>
<th>15</th>
<th>18</th>
<th>20</th>
<th>23</th>
<th>25</th>
<th>26</th>
<th>29</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>.322</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>.412</td>
<td>.089</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>.349</td>
<td>.445</td>
<td>.280</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>.252</td>
<td>.301</td>
<td>.310</td>
<td>.307</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>.518</td>
<td>.485</td>
<td>.277</td>
<td>.245</td>
<td>.236</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>.448</td>
<td>.484</td>
<td>.310</td>
<td>.488</td>
<td>.339</td>
<td>.453</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>.372</td>
<td>.466</td>
<td>.274</td>
<td>.359</td>
<td>.437</td>
<td>.434</td>
<td>.458</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>.315</td>
<td>.420</td>
<td>.326</td>
<td>.426</td>
<td>.518</td>
<td>.238</td>
<td>.384</td>
<td>.554</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>.285</td>
<td>.603</td>
<td>.043</td>
<td>.354</td>
<td>.173</td>
<td>.378</td>
<td>.475</td>
<td>.529</td>
<td>.387</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>.323</td>
<td>.442</td>
<td>.193</td>
<td>.406</td>
<td>.400</td>
<td>.347</td>
<td>.588</td>
<td>.540</td>
<td>.413</td>
<td>.629</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>.296</td>
<td>.560</td>
<td>.209</td>
<td>.304</td>
<td>.481</td>
<td>.447</td>
<td>.358</td>
<td>.453</td>
<td>.410</td>
<td>.514</td>
<td>.451</td>
<td>1</td>
</tr>
</tbody>
</table>

The Cronbach’s α on Interpersonal Relationships subscale was .86 indicating a reasonable level of reliability for a factor. In addition, the inter-item correlation of this factor ranged from .268 to .734 (see Table 4.10), which approached the criteria for appropriate inter-item correlations (Clark & Watson, 1995).

Table 4.10

*Inter-item Correlations for CIRAA Factor 2: Interpersonal Relationship*

<table>
<thead>
<tr>
<th>Item</th>
<th>2</th>
<th>3</th>
<th>9</th>
<th>11</th>
<th>14</th>
<th>17</th>
<th>22</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>.510</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>.537</td>
<td>.336</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>.562</td>
<td>.419</td>
<td>.574</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>.555</td>
<td>.423</td>
<td>.268</td>
<td>.459</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>.327</td>
<td>.439</td>
<td>.277</td>
<td>.366</td>
<td>.338</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>.734</td>
<td>.552</td>
<td>.436</td>
<td>.574</td>
<td>.538</td>
<td>.385</td>
<td>1</td>
</tr>
</tbody>
</table>
The Cronbach’s α for Coping Skills was .689. According to Nunnally and Bernstein (1994), the alpha value greater than .6 is usually acceptable. The inter-item correlations ranged from .133 to .406 (see Table 4.11), which were close to the criteria Clark and Watson suggested (1995).

Table 4.11

*Inter-item Correlations for CIRAA Factor 3: Coping Skills*

<table>
<thead>
<tr>
<th>Item</th>
<th>4</th>
<th>6</th>
<th>19</th>
<th>21</th>
<th>27</th>
<th>28</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>.512</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>.265</td>
<td>.133</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>.251</td>
<td>.303</td>
<td>.260</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>.267</td>
<td>.347</td>
<td>.297</td>
<td>.406</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>.330</td>
<td>.298</td>
<td>.441</td>
<td>.397</td>
<td>.313</td>
<td>1</td>
</tr>
</tbody>
</table>

The Cronbach’s α for Internal Locus of Evaluation was .699, indicating acceptable reliability. The inter-item correlations ranged from .192 to .60 (see Table 4.12), and approximated the criteria Clark and Watson recommended (1995).

Table 4.12

*Inter-item Correlations for CIRAA Factor 4: Internal Locus of Evaluation*

<table>
<thead>
<tr>
<th>Item</th>
<th>10</th>
<th>12</th>
<th>16</th>
<th>24</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>.318</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>.369</td>
<td>.242</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>.370</td>
<td>.192</td>
<td>.446</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>.308</td>
<td>.201</td>
<td>.612</td>
<td>.630</td>
<td>1</td>
</tr>
</tbody>
</table>
Concurrent Validity with Child Behavioral Checklist

To assess the validity of the CIRAA, a concurrent validity study was conducted to evaluate the correlation of the CIRAA total score with the Child Behavioral Checklist (CBCL; Achenbach & Rescorla, 2001). Concurrent validity was conducted with a clinical sample of 47 parents whose children were receiving services in the two local counseling clinics. The Pearson $r$ correlation coefficient for assessing the relationship between CIRAA total scores and CBCL Total Problems scores was generated and reached statistical significance. A medium negative correlation ($r = -0.417$, $n = 47$, $p < .01$) was found between the two variables. High levels of adjustment behaviors on the CIRAA correlated with low levels of clinical behaviors on the Total Problem score of the CBCL (Cohen, 1988).
CHAPTER 5
DISCUSSION

One imperative task for researchers is the evaluation of their selected instruments' abilities to measure the named construct in order to establish a body of evidence for any mental health intervention’s effectiveness (APA, 2008; Heppner et al., 2008). After investigating the play therapy studies, most instruments used did not truly capture the objectives of child-centered play therapy (CCPT) and only focused on children's behavioral changes (Holliman, 2010; Lin, 2011). Therefore, Holliman and Ray (2013) decided to develop an instrument based on the objectives of CCPT and believed that developing this rigorous theoretically-based instrument would move the discipline toward empirically-supported treatment. With this intention in mind, they used a meticulous assessment development process to develop the Child Interpersonal Relationships and Attitudes Assessment (CIRAA) with strong empirical support (Holliman, 2010). They recommended conducting confirmatory factor analysis (CFA) as the next validation process. Obtaining affirmative findings in the CFA might not only provide credibility for the factor structure of CIRAA but might also strengthen the construct validity of the instrument.

The results from the present study showed acceptable goodness of model fits and interpretable factor loadings. Therefore, the results provided confirmation and further evidence of the factorial validity of the CIRAA’s four-factor model. Additionally results indicated acceptable internal consistency reliability and concurrent validity, and supported the applicability of the instrument for both genders and a range of ages, suggesting its appropriateness use in making clinical decisions.
Validity

The CFA results suggested overall reasonable factor/structure validity of the CIRAA with acceptable goodness model fits and interpretable factor loadings. In addition to the CFA, the instrument demonstrated concurrent validity through a medium-sized negative correlation with the CBCL, one instrument often utilized in play therapy research. Holliman (2010) also used the CBCL to assess concurrent validity of the CIRAA’s subscales and total scores. He found strong, positive correlation between the CIRAA and CBCL ($r = .75$). The direction difference from positive to negative correlation was based on the adjustment I made to CIRAA scoring for this study, reversing the scores to denote positive behaviors or attitudes with higher scores. Because higher scores on the CBCL indicate more problematic behavior and higher scores on the adjusted CIRAA indicate more desirous behaviors or attitudes, the relationship correlation was expected to be negative. In the current study, the correlation coefficient demonstrated a statistically significant relationship between the CIRAA and CBCL but with a moderate correlation. In comparison to Holliman’s (2010) concurrent validity results ($r = .75$), the current study demonstrated a lesser relationship between CIRAA and CBCL. These results might be explained by smaller sample size used in the current study ($n = 47$) as compared to larger sample in Holliman’s study ($n = 80$).

Although CFA results supported the factor/structure validity of the CIRAA, one question arose from the correlation results obtained from the CFA analysis in Model 2. Unlike Model 1, Model 2 allowed four factors to correlate with each other. The correlation between Self-Control and Coping Skills yielded $r = .9$, indicating an especially high correlation between these two constructs’ subscales. Brown (2006) stated that factor correlations close to 1.0 provide apparent evidence for the notion that the latent factors stand for distinct constructs. A factor correlation
that exceeds .85 is usually utilized as the criterion to suggest poor discriminant validity (Brown, 2006). Further investigation of the factors’ structures is needed in order to enhance the instrument’s discriminant validity. A more detailed discussion of this issue is presented in the recommendation section.

Reliability

The results of the reliability analysis of the CIRAA indicated reasonable internal consistency reliability for the subscales and the total scores. Both the Cronbach’s α value and inter-item correlation ranges were calculated for each subscale. All of these analyses suggested reasonable levels of reliability for the total scores and four subscales. Moreover, the internal consistency coefficients of the CIRAA (α = .928) met Nunnally and Bernstein’s (1994) criteria for reliability standards for making clinical decisions for individuals based on treatment effects (Holliman, 2010).

Generalizability of Sample

The data were collected from local university-based counseling clinics, elementary schools, and graduate level courses in the Southwest United States. In addition, the cultural make-up of the participants encompassed mainly Caucasian, Hispanic, and African-American individuals. Although the sample represented the three major ethnicity groups reported in the United States, the results should not be generalized to populations that do not match the demographics of the sample (Holliman, 2010). Because the sample was a convenience sample, the results of the study may not generalize to different populations or different conditions.
Utility of Instrument

The mean age for the children recruited in this study was 6.14 years old. In the attempt to design an instrument that is feasible for use in play therapy, the mean age of the current study’s sample matches the reported mean age of children studied in play therapy research which was reported at 6.5 years old (Bratton et al., 2005), indicating the sample for assessment development matched the typical sample in play therapy studies. Such findings suggest greater utility for the instrument in play therapy research. The mean age of the independent sample in this study also replicates the examination of the CIRAA structure in Holliman's study (2010) with a similar sample age ($M = 6.5$ years). Hence, the CIRAA appears appropriate for utilizing with younger children, a unique characteristic of the instrument relative to other frequently-used child instruments.

The CIRAA offers practitioners important information for making clinical decisions. Holliman (2010) established the clinical cutoff score which offers clinicians a simple method of evaluating a child’s progress in therapy based on CCPT objectives. Unlike many assessments, the CIRAA contains only 30 Likert-scale items and can be administered very quickly. The instrument does not call for any specialized training beyond a master’s degree in a mental health discipline with training in statistics and instrument administration and interpretation. Moreover, the instrument does not need specific administration procedures. Instead, the CIRAA is essentially self-administered by the parent (Holliman, 2010). Overall, the CIRAA is easy and quick to administer and can provide meaningful information for clinicians making clinical decisions and assessing client progress.
Recommended Modifications to the Instrument

Based on results of the current study, there are recommendations that would improve the practicality and structure of the original CIRAA. On the original version of the CIRAA, the items were presented within the factor clusters. In order to decrease the possibility of respondent fatigue on factor items, I rearranged items to distribute them across factors. In addition to rearranging items, I also reversed scoring on items to match higher scores with more desirous behaviors and attitudes. After review of Holliman’s (2010) research, I believed that reversing the scoring method would better fit the original philosophy and rational for developing this instrument (Holliman & Ray, 2013). For example, after reversing the score, if a child had a high score on the Self-Control subscale, this meant that the child demonstrated better self-control.

Holliman and Ray (2013) developed this instrument based on objectives of child-centered play therapy (CCPT) and intended to focus on children’s strengths, which is consistent with CCPT philosophy. Therefore, reversing the scores was more theoretically consistent with a more positive view of children.

In reviewing the individual items in relation to the factor labels, I also noted discrepancies between the focus of items within a factor and an apparent mismatch to the factor label. I recommend using Self-Regulation as a more appropriate name for the Self-Control factor. The term self-regulation portrays the purposive process of self-corrective adjustments made by a child demonstrating the ability to adjust the intensity and frequency of inappropriate behaviors, delay the need of immediate gratification, take responsibility for behavior, show care and empathy for other individuals’ feelings, and demonstrate responsible freedom (Carver & Scheier, 2011; Holliman & Ray, 2013; Honig & Lansburgh, 1991; Landreth, 2012; Mischel et al., 1988/1989).
Additionally, the item "my child exhibits self-control" falls into the Self-Control factor. The circular pattern of the subscale name with an item using the same term is problematic. Pett et al. (2003) cautioned researchers to avoid naming a factor with a term similar to one of the factor’s items. Because of the difficulty of measuring a latent construct, researchers use different methods to compile observable items together in order to capture certain constructs. If the label of a factor is identical to wording of an item within that factor, the item is considered redundant and unnecessary because the purpose of items within a factor is to capture latent constructs. Hence, the original item “my child exhibits self-control” already captures the proposed latent factor of self-control. When consulting with the communality of this item ($R^2 = .55$), the factor only accounted for 55% of the variance in this item. Therefore, it is reasonable to conclude that the factor actually captures a broader concept than self-control, hence, another justification to rename this particular factor to a latent construct, self-regulation.

Limitations

Even though the instrument demonstrated acceptable model fit, a few limitations are worthy of discussion. First, the RMSEA of the primary model was .088, a value slightly bigger than .06-.08, with 90% CI ranging from .0082 to .095, suggesting minor degree of model misspecification. When investigating the modification indices, the results indicated that $\chi^2$ would increase 113.9, and the new estimate would be .88 if Self-Control and Coping Skills were allowed to correlate. Even though a competing model was not developed due to the purpose of this study, the relationship between Self-Control and Coping Skills are worth noting ($r = .9$ in Model 2). This high correlation between Self-Control and Coping Skills led to poor discriminate
validity. Suggestions for further study in regard to this issue are presented in the recommendation section of this chapter.

When conducting CFA, researchers must address the important question of how many samples should be obtained in order to gain “an acceptable level of precision and statistical power of the model’s parameter estimates, as well as reliable indices of overall model fit” (Brown, 2006, p. 412). Different guidelines for sample size suggest a minimum sample size of \( n \) greater than or equal to a range of 100 to 200 cases including a minimum number of cases per freed parameter (5:1; Bentler & Chou, 1987; Brown, 2006; Gorsuch, 1983). Even though the sample size of 206 was adequate for conducting the CFA and met the minimum ratios of five participants to each item, as suggested by Gorsuch (1983), a larger sample size would have provided an increase in the statistical power and precision of the model’s parameter estimates.

One phenomenon was noted in regard to the item-factor relationships. In the Interpersonal Relationship construct, all items were negatively worded (e.g., “My child complains few people like him/her”). Yet, in the Internal Locus of Evaluation construct, all items were positively worded. Brown (2003) stated that when scales are developed including a mixture of positively and negatively worded items, "factor analyses will frequently produce distinct factors whereby positively worded items load on one factor and reverse-worded items on another" (p. 1412; Bagozzi, 1993; Baumgartner & Steenkamp, 2001). The combination of positively and negatively worded items within a scale often leads to the presence of method effects (Bagozzi & Heatherton, 1994; Brown, 2003; Marsh, 1996). A method effect occurs as a result of the differential covariance among items due to the measurement approach instead of due to the substantive latent factors. The use of positively and negatively worded items might have affected the factor structure and should be explored further.
The lack of generalizability can be attributed to the use of convenience sample, as discussed in the generalizability section. Clinicians need to be cautious when interpreting the results of the CIRAA for clients representing ethnic groups other than Caucasian, Hispanic, and African-American and other places, languages, and cultures of the world.

Implications

In both Holliman and Ray’s (2013) and the present study, the results suggest that the instrument meets psychometric internal consistency reliability guidelines for both treatment effect research and for individual clinical decision making. The instrument can be utilized to specify when children need more therapy, attend to the concerns encompassed in the CIRAA, or identify when children have progressed to a degree of adjustment for which termination might be suggested (Holliman, 2010). This instrument can also be utilized to recognize children who might profit from counseling services.

This instrument depends on parents’ observations. Because most clinicians do not have the opportunity to observe children outside of therapy sessions or in school, parental reporting through this instrument can provide valuable information to clinicians about a child’s behavior outside of the therapy sessions. This instrument offers therapists opportunities to examine if their perceptions of their child clients are congruent with the clients’ parental perceptions. School-based mental health professionals, who offer interventions, can also utilize this instrument to determine if a certain intervention impacts a child’s at-home behavior (Holliman, 2010).

The CIRAA is also the first parent-report instrument based on the objectives of CCPT, with both clinical and research applications (Holliman & Ray, 2013). Assessment criteria for the
establishment of empirically supported treatments (EST) need valid and reliable measurement of real-world change associated with the constructs of the specified mental health intervention (Kazdin & Weisz, 1998). This instrument was designed to measure the objectives of CCPT. Therefore, researchers can utilize the CIRAA in play therapy studies to capture the constructs of interest.

Recommendations for Further Studies

Further studies are recommended to reevaluate the factor structure of the CIRAA, such as testing the two-factor model using the positively versus negatively worded items. Also, researchers may consider modifying the wording of the items or allow the instrument to represent the constructs with an equal number of negatively and positively worded items to avoid the current type of method effect found in my study (Brown, 2003; March, 1996).

Regarding poor discriminant validity, Brown (2006) suggested when two factors are highly overlapping, one general research strategy is to “respecify the model by collapsing the dimensions into a single factor and determine whether this modification results in a significant degradation in model fit” (p. 131). If the respecified model presents an adequate fit to the data, it is often preferred due to its superior parsimony. Therefore, future studies are needed to explore the possibility of modifying the CIRAA into a three-factor instrument and to see if this modification strengthens model fits and discriminant validity.

Holliman (2010) used receiver operating characteristics (ROC) curve analysis to determine the clinical threshold score for the CIRAA as 2.5. Therefore, further studies regarding the clinical threshold score are needed to validate Holliman’s (2010) finding. According to Nunnally and Bernstein (1994), the internal consistency reliability of the CIRAA total scores ($\alpha$
implied that this instrument is appropriate for making clinical decisions about individuals. Therefore, the evidence of CIRAA demonstrating stable clinical cut-off score allows practitioners to make appropriate clinical judgments.

The non-clinical and clinical samples scored statistically significantly different and yielded large effect sizes on all subscales and total score, except the Internal Locus of Evaluation subscale. The mean difference between these two groups on the Internal Locus of Evaluation subscale was very small and yielded overall higher scores than on the other subscales, indicating non-clinical and clinical samples scored well and did not score differently from each other on this subscale. This result suggested that this factor needs to be examined more closely through future research to see if internal locus of evaluation is not a differentiating factor for children who are clinical or not clinical or explore other possibilities of this occurrence.

After modifying the CIRAA according to the recommendations, researchers are encouraged to conduct a study with a population not only larger but also with a broad array of ethnic groups represented. The multiple regression results suggested that the CIRAA is applicable across age and gender. Therefore, exploring the relationship between ethnicity and CIRAA scores may improve the generalizability of the instrument.

Conclusion

Holliman and Ray (2013) developed the CIRAA according to the objectives of CCPT and to accurately measure the constructs of interest and promote CCPT as an empirically supported treatment. The CIRAA is the first instrument to be developed based on the philosophy of CCPT. The CIRAA demonstrates an overall reasonable factor/structure validity and acceptable concurrent validity. The instrument can be utilized to assess treatment effects and make clinical
decisions. Clinicians may utilize this instrument to monitor children’s progress in therapy, recognizing that the instrument modifications yield higher scores as indicating better adjustment. The instrument is applicable to across ages (3-10 years of age) and between genders, and appears to demonstrate sensitivity in distinguishing between clinical and non-clinical samples.

Based on the results of this study, the CIRAA demonstrates a confirmed factor structure yielding four factors and a total score. Thorough review of the CIRAA resulted in the recommendation to modify the Self-Control factor/subscale by renaming the factor to better represent the latent construct, Self-Regulation. Further research and refinement of the CIRAA is encouraged to enhance the instrument’s generalizability, discriminant validity, and factor validity.
APPENDIX A

INSTITUTIONAL REVIEW BOARD INFORMED CONSENT DOCUMENTS
University of North Texas Institutional Review Board

Informed Consent Form

Before agreeing to participate in this research study, it is important that you read and understand the following explanation of the purpose, benefits and risks of the study and how it will be conducted.

Title of Study: Confirming the Constructs of the Child Interpersonal Relationship and Attitudes Assessment

Student Investigator: Chiao-Feng Chung, a doctoral candidate at the University of North Texas (UNT) Department of Counseling and Higher Education.

Supervising Investigator: Dr. Dee Ray, a faculty member in the University of North Texas (UNT) Department of Counseling and Higher Education.

Purpose of the Study: You are being asked to participate in a research study which involves investigation of your perception of your child’s behaviors at home and school. This information gathered through this study will be used to further develop a psychometric instrument.

Study Procedures: You will be asked to answer background question such as age, race, gender, and relationship to the child, as well as questions about your experience of your child’s behavior. This should take about 10 minutes of your time.

Foreseeable Risks: There are no significant foreseeable risks to any participant but the possibility exists that participants may experience discomfort during completing instruments, such as length of time to complete or acknowledgement of child behavior problems. Any instruments completed by the parent are considered confidential, meaning that the researchers will not reveal anything that is said or written during the administration process. However, if the parent discloses child abuse, neglect, exploitation or intent to harm another person, the therapist is required by law to report to the appropriate authority.

Benefits to the Subjects or Others: Although this study may not be of any direct benefit to you, this study is expected to benefit others by providing information to assist the researcher in measuring the effect of play therapy to better enable mental health professionals to meet the need of parents who are experiencing difficulties with their children.

Compensation for Participants: None

Procedures for Maintaining Confidentiality of Research Records: This researcher will take all necessary precautions to protect your confidentiality by coding your signed consent form, your demographic information, and your responses, and maintaining them in separate locations. When the results of this study are presented, only general demographic information will be disclosed, and therefore, your confidentiality will be maintained. The confidentiality of your
individual information will be maintained in any publications or presentations regarding this study.

**Questions about the Study:** If you have any questions about the study, you may contact Chiao-Feng Chung at 940-565-2066 or email at Chiao-Feng.Chung@unt.edu or Dr. Dee Ray, UNT Department of Counseling Development and Higher Education at telephone number 940-565-2006.

**Review for the Protection of Participants:** This research study has been reviewed and approved by the UNT Institutional Review Board (IRB). The UNT IRB can be contacted at (940) 565-3940 with any questions regarding the rights of research subjects.

**Research Participants’ Rights:** Your signature below indicates that you have read or have had read to you all of the above and that you confirm all of the following:

- Ms. Chung has explained the study to you and answered all of your questions. You have been told the possible benefits and the potential risks and/or discomforts of the study.
- You understand that you do not have to take part in this study, and your refusal to participate or your decision to withdraw will involve no penalty or loss of rights or benefits. The study personnel may choose to stop your participation at any time.
- Your decision whether to participate or to withdraw from the study will have no effect on your grade or standing in this course or services provided to your child.
- You understand why the study is being conducted and how it will be performed.
- You understand your rights as a research participant and you voluntarily consent to participate in this study.
- You have been told you will receive a copy of this form.

__________________________
Printed Name of Participant

__________________________                                ____________
Signature of Participant                                      Date
For the Student Investigator or Designee:

I certify that I have reviewed the contents of this form with the subject signing above. I have explained the possible benefits and the potential risks and/or discomforts of the study. It is my opinion that the participant understood the explanation.

____________________________________  ____________
Signature of Student Investigator    Date
APPENDIX B

LETTER TO PARENTS
Dear Parents,

We would like to ask you to complete a short survey regarding each of your children between the ages of 3 – 10 years old. The survey is titled, *Child Interpersonal Relationship and Attitudes Assessment (CIRAA)*, and is under development as a research instrument to help with child services. This assessment includes 30 items and will only require **5-10 minutes** of your time to complete. Your decision to participate is completely voluntary. Because the instrument is still in development, the information you share will be used for statistical analyses purposes only and will not be used to evaluate your individual child’s behaviors. Please complete an individual assessment for each of your children between the ages of 3-10 years.

If you have any questions, please contact Chiao-Feng Chung at 940-565-2066. Thank you!!

Sincerely,

Chiao-Feng Chung
Doctoral Candidate

Dee C. Ray
Professor, Counseling Program
APPENDIX C

CIRAA
CHILD INTERPERSONAL RELATIONSHIP AND ATTITUDE ASSESSMENT

Child name_____________________________________Today’s date:
_____/_____/____________
Child’s birthdate: _____/_____/____________Child’s age _________   Child’s grade_____________
Child’s gender: □ Male □ Female   Child’s ethnic group or race:_________________________
Name of person completing this form:_________________________ Your gender: □ Male □ Female   
Your relationship to the child:
□ Biological Parent □ Step Parent □ Adoptive Parent □ Foster Parent □ Grandparent □ Other (specify) _______
Has your child been in counseling during the past six months: □ YES □ NO

Instructions: Please read each statement about your child carefully. Please circle a response for each item indicating the degree to which you agree or disagree with the statement. Please be sure to answer all items.

SD=Strongly Disagree   D=Disagree   N=Neutral   A=Agree   SA=Strongly Agree

1. My child shares with other children...........................   SD D N A SA
2. My child complains few people like him/her.....................   SD D N A SA
3. My child is often sad for prolonged periods of time.............   SD D N A SA
4. My child is often upset by minor things..........................   SD D N A SA
5. My child has verbal fights with other student at school..........   SD D N A SA
6. My child becomes anxious over small matters...................   SD D N A SA
7. My child apologizes for hurting other’s feelings.................   SD D N A SA
8. My child blames others for mistakes...............................   SD D N A SA
9. My child is teased by other children...............................   SD D N A SA
10. My child enjoys doing things for him/herself.....................   SD D N A SA
11. My child complains other children are mean to him/her........   SD D N A SA
12. My child volunteers to help out around the house................   SD D N A SA
13. My child hits/kicks family members...............................   SD D N A SA
14. My child often compares him/herself unfavorably to others.....   SD D N A SA
15. My child gets along with other children..........................   SD D N A SA
16. My child shows me projects he/she is proud of...................   SD D N A SA
17. My child is overly sensitive........................................   SD D N A SA
18. My child exhibits self-control.......................................   SD D N A SA
19. My child is able to express his/her feelings when he/she is feeling upset...........................................   SD D N A SA
20. My child disrupts family events/outings............................   SD D N A SA
21. My child can name things he/she likes about him/herself........ SD  D  N  A  SA
22. My child complains no one likes him/her......................... SD  D  N  A  SA
23. My child yells frequently........................................... SD  D  N  A  SA
24. My child talks about accomplishments of which he/she is proud.. SD  D  N  A  SA

25. My child often receives reports of disruptive behavior from his/her teacher.................................................. SD  D  N  A  SA
26. My child often acts without thinking............................... SD  D  N  A  SA
27. For his/her age, my child is able to enter new situations with confidence........................................................ SD  D  N  A  SA
28. My child can calm down when upset................................ SD  D  N  A  SA
29. My child has physical fights with other children................. SD  D  N  A  SA
30. My child tells me things he/she is proud of....................... SD  D  N  A  SA
REFERENCES


http://www.indiana.edu/~statmath/stat/all/cfa/cfa.pdf


discussions on children's moral decisions. *Social Cognition, 2*(4), 343-359. doi:
http://dx.doi.org/10.1521/soco.1984.2.4.343


* Dissertations Abstracts International: Section A. Humanities and Social Sciences, 61**(07)**, 2603.


Brown, T. A. (2003). Confirmatory factor analysis of the Penn State Worry Questionnaire:


discounting of delayed hypothetical rewards in cocaine-dependent individuals.


Holliman, R. P., & Ray, D. C. (2013). Development of the child interpersonal relationship and
doi:10.1037/a0032311

Effects in child behavior, parent-child relationship stress, and parental empathy.
*Dissertation Abstracts International Section B: Sciences and Engineering*, 71(8).


doi:10.1037/h0030773


Miller, J. B. (1976). *Toward a new psychology of women.* Boston, MA; Beacon Press.


Ray, D. C. (2011). *Advanced play therapy: Essential conditions, knowledge, and skills for child*


Smith, R. E. (1989). Effects of coping skills training on generalized self-efficacy and locus of


adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality, 72*(2), 271-322.


