UC San Diego
UC San Diego Electronic Theses and Dissertations

Title

Psychometric properties of the parent daily report and its potential for use in child welfare settings

Permalink

Author

Keil, Vivien

Publication Date

2007

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

SAN DIEGO STATE UNIVERSITY

Psychometric Properties of the Parent Daily Report

and Its Potential for Use in Child Welfare Settings

A Dissertation submitted in partial satisfaction of the requirements for the degree

Doctor of Philosophy

in

Clinical Psychology

by

Vivien Keil

Committee in charge:

    University of California, San Diego

        Professor Ann F. Garland
        Professor Susan F. Tapert

    San Diego State University

        Professor Joseph M. Price, Chair
        Professor Alan J. Litrownik
        Professor May Yeh

2007

The dissertation of Vivien Keil is approved, and it is

acceptable in quality and form for publication on microfilm:

_____

_____

_____

_____

_____
                                                                    Chair

University of California, San Diego

San Diego State University

2007

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

VITA

EDUCATION

2007
Doctor of Philosophy, Clinical Psychology
San Diego State University (SDSU)/University of California, San Diego (UCSD)
Joint Doctoral Program in Clinical Psychology (JDP)
Cumulative GPA: 3.90
Dissertation: *Psychometric Properties of the Parent Daily Report and its Potential for Use in Child Welfare Settings*
　　　Successful Defense Completed in July 2006
Dissertation Chair: Joseph M. Price

2004
Master of Science, Clinical Psychology
SDSU/UCSD Joint Doctoral Program in Clinical Psychology
Thesis: *Social Information-Processing Patterns and Peer Relations among Maltreated Children*
Thesis Chair: Joseph M. Price

2002
Bachelor of Science, Cognitive Neuroscience
Brown University, Providence, RI
Cumulative GPA: 3.90, Magna Cum Laude

AWARDS AND RESEARCH GRANTS

2007 – present
Phi Kappi Phi Honor Society

2004 – 2007
Individual National Research Service Award (NRSA)
NIMH F31 MH068863 ($112,184 over four years)
Training on the Etiology and Treatment of Behavior Disorders

2003
Research Travel Award, SDSU/UCSD

2002 – 2003
NIH Scholars Award, UCSD

2002
Baker Fellowship for Graduate Studies, Brown University

2002
Magna Cum Laude, Brown University

SUPERVISED CLINICAL EXPERIENCE

APA-Accredited Clinical Internship, Children's Hospital, Orange County, Department of Pediatric Psychology
September 2006 – present
Supervisors: Heather Huszti, Ph.D; Julie Reeker, Ph.D; Marni Nagel, Ph.D
- Conduct intake and diagnostic assessments for inpatients and outpatients
- Administer neuropsychological assessment batteries and write integrated reports
- Provide consultation and liaison services to medical teams throughout pediatric hospital
- Provide emergency and after-hours consults for high-risk patients (e.g., SI/HI/psychosis)
- Conduct outpatient psychotherapy with children, adolescents, and families presenting with significant psychopathology and comorbid medical conditions such as cancer and diabetes
- Participate in weekly supervision and multidisciplinary treatment team meetings with attending physicians and medical residents, psychologists, social workers, fellows
- Administer Bayley Scales of Infant Development and serve as psychological consultant for Young Child Clinic and Neonatal Intensive Care Unit (NICU)
- Completed rotation in conducting medication appointments at Psychiatry Clinic under supervision of Dr. Wayne Nguyen, M.D.
- Completed rotation as psychological consultant for Hematology Specialty Clinic
- Co-facilitated 10-week parenting group targeting children with symptoms of ADHD
- Will complete minor rotation in and Eating Disorders Clinic beginning May 2007

Advanced Psychology Practicum, Children's Hospital, San Diego, Outpatient Psychiatry
July 2005 – June 2006
Supervisor: Ann Garland, Ph.D
- Conducted intake and diagnostic assessments
- Administered psychological assessment batteries and wrote integrated reports
- Conducted outpatient psychotherapy with children, adolescents, and families
- Participated in weekly supervision and multidisciplinary treatment team meetings with psychiatrists, psychologists, social workers, marriage and family therapists, trainees

- Presenting problems included behavior disorders, anxiety, mood disorders, learning disabilities, schizophrenia, attachment disorders, developmental disorders, substance abuse
- Treatment approaches included psychoeducation, supportive therapy, ecological/systems, cognitive behavioral therapy, behavioral parent training, play therapy, behavioral analysis

Advanced Psychology Practicum, UCSD Veterans Affairs Medical Center (VAMC), Department of Psychiatry
February 2005 – December 2005
Supervisors: Sandy Brown, Ph.D; John McQuaid, Ph.D
- Conducted intake and diagnostic assessments
- Engaged in structured, evidence-based Cognitive-Behavioral Therapy (CBT)
- Conducted psychotherapy with veterans with complex co-occurring medical conditions
- Presenting problems included anxiety, depression, PTSD, personality disorders, psychotic disorders, substance abuse
- Treatment approaches included CBT, behavioral activation, exposure therapy, motivational interviewing, case management
- Attended weekly individual supervision, group supervision, and CBT seminars

Advanced Psychology Practicum, UCSD Child and Adolescent Psychiatry Services (CAPS)
July 2004 – August 2005
Supervisor: Sandy Brown, Ph.D
- Conducted intake and diagnostic assessments in inpatient hospital for children and adolescents exhibiting severe psychopathology
- Administered comprehensive psychological assessment batteries and wrote 21 integrated reports, including cognitive functioning, academic achievement, personality, emotional/ behavioral functioning, and neuropsychological domains such as executive functioning
- Created and implemented behavior modification programs for youth with severe behavior problems and developmental disorders
- Conducted individual and family therapy in crisis situations
- Led child and adolescent groups focusing on milieu therapy, coping and social skills training
- Participated in weekly multidisciplinary team meetings including psychiatrists, nurses, psychologists, social workers, occupational and recreational therapists, and marriage and family therapists, to discuss case conceptualization, treatment and disposition planning
- Attended weekly individual and group supervision meetings, and intern seminars

- Trained incoming students in screening process and administration and scoring of intake measures, including WISC-IV, WASI, WRAT-3

Psychology Practicum, SDSU Psychology Clinic
July 2003 – June 2004
Supervisors: Linda Gallo, Ph.D; Alan Litrownik, Ph.D
- Conducted intake interviews and administered diagnostic evaluations using the Structured Clinical Interview for DSM-IV-TR (SCID) Axis I disorders
- Conducted psychotherapy with individuals and couples
- Presenting problems included anxiety, mood disorders, eating disorders, marital discord, psychotic symptoms, substance abuse
- Treatment approaches included cognitive-behavioral therapy, interpersonal therapy, behavioral couples therapy, case management
- Attended weekly individual and group supervision meetings

Psychology Practicum, San Diego Office of the District Attorney
July 2003 – January 2004
Supervisor: Alan Litrownik, Ph.D
- Conducted training groups for Kids In Court, a program for maltreated children and adolescents testifying against their perpetrators in court
- Led group therapy for children primarily between the ages of 5 and 14
- Provided psychoeducation, supportive therapy, relaxation therapy, and coping skills training in the courts of San Diego

Psychology Practicum, Chadwick Center for Children and Families, Children's Hospital
July 2003 – December 2003
Supervisors: Alan Litrownik, Ph.D; Iby Kantor, Ph.D
- Trained in conducting intake assessments with children and adolescents with a history of trauma, including sexual abuse, physical abuse, and/or domestic violence
- Participated in weekly multidisciplinary treatment team meetings focusing on diagnosis and treatment planning
- Training at this site was ended prematurely due to budget cuts

Clinical Group Facilitator, Child and Adolescent Services Research Center, Project KEEP
September 2002 – April 2004
Supervisors: Joseph M. Price, Ph.D; Patricia Chamberlain, Ph.D
- Served as a co-facilitator for parent management training and support groups for foster and kin parents, as part of a Project KEEP, a large-scale effectiveness study
- Administered manualized version of parent management training to culturally diverse populations in San Diego County

- Conducted weekly assessments of child behavior problems
- Attended weekly training and group supervision meetings

RESEARCH EXPERIENCE

Predoctoral Fellow, Child and Adolescent Services Research Center
(CASRC), Children's Hospital, San Diego
August 2002 – present
Supervisor: Joseph M. Price, Ph.D
- Participate in data collection, data coding, data analysis, and grant writing
- Analyze data from a longitudinal investigation on the social adjustment of maltreated children and social information-processing patterns using SPSS and HLM software
- Analyze data from a treatment outcome study on parent management training for foster parents using SPSS and M Plus software
- Prepare original manuscripts for submission to peer-reviewed journals
- Participate in weekly supervision, research meetings, and statistics workshops
- Attend annual meeting of the Child and Adolescent Interdisciplinary Research Network (CAIRN), a NIMH-funded network of researchers focusing on the dissemination of evidence-based, parent-mediated interventions into community settings
- Currently working on research projects remotely from clinical internship site

Research Assistant, Brown University, Education Department
March 1999 – May 2002
Supervisor: Jin Li, Ed.D
- Recruited and interviewed elementary-aged children and college students
- Coded and analyzed data for three research projects examining learning and motivation among young children and college students
- Served as leader of research teams, coordinated weekly meetings, and trained new assistants

Research Assistant, Miriam Hospital, Center for Behavioral and Preventive Medicine
September 2001 – April 2002
Supervisor: Laura Stroud, Ph.D
- Recruited and interviewed subjects for study of stress responses in children and adolescents
- Collected heart rate, blood pressure, and saliva samples
- Served as experimental confederate

OTHER CLINICAL EXPERIENCE

Mental Health Worker, Butler Psychiatric Hospital, Providence, RI
April 2001 – May 2002
Supervisor: Barbara Ostrove
- Facilitated milieu therapy in inpatient child and adolescents units
- Monitored behavior problems of youth for treatment planning and created individualized behavior programs

Treatment Teacher, The Groden Center for Autism and Developmental Disabilities, Providence, RI
October 1999 – August 2000
Supervisor: Denise Bowman
- Coordinated and supervised excursions for children with pervasive developmental disabilities
- Conducted relaxation training and implemented behavior modification programs
- Monitored therapy goals and progress and adjusted behavior plans accordingly

TEACHING EXPERIENCE

Guest Lecturer, SDSU Department of Psychology
November 2005
- Presented lecture entitled, *Language and Learning Disorders,* in an undergraduate course in Developmental Psychopathology

Guest Lecturer, SDSU Department of Psychology
October 2005
- Presented lecture entitled, *Disorders of Childhood and Adolescence,* in an undergraduate course in Abnormal Psychology

Graduate Teaching Assistant, SDSU Department of Psychology
September 2004 – December 2004
Supervisor: Joseph M. Price, Ph.D
- Conducted office hours to aid students struggling in an undergraduate course in Developmental Psychology
- Assisted professor in developing, proctoring, and grading exams and essays

Guest Lecturer, SDSU Department of Psychology
October 2004
- Presented lecture entitled, *Information-Processing and Social Behavior,* in an undergraduate course in Developmental Psychology

Guest Lecturer, SDSU Department of Psychology
October 2004
- Presented lecture entitled, *Neurophysiologic Effects of Maltreatment,* in an undergraduate course in Developmental Psychology

Guest Lecturer, SDSU Department of Psychology
July 2004
- Presented lecture entitled, *Attachment,* in an undergraduate course in Developmental Psychology

Teaching Assistant, Brown University, Department of Psychology
January 2002 – May 2002
Supervisor: John Wincze, Ph.D
- Aided professor in developing curriculum for seminar course in Behavior Modification
- Held office hours, developed, proctored, and graded exams and written assignments

Teaching Assistant, Brown University, Department of Psychology
September 2001 – December 2001
Supervisor: Brian Hayden, Ph.D
- Developed lesson plans and ran independent discussion sections for course in Abnormal Psychology
- Held office hours, developed, proctored, and graded exams and written assignments

Teaching Assistant, Brown University, Education Department
September 2000 – December 2000
Supervisor: Jin Li, Ed.D
- Developed lesson plans and ran discussion sections for course in Child Development
- Held office hours, developed, proctored, and graded exams and written assignments

PROFESSIONAL AFFILIATIONS

2004 – present   Society for Research in Child Development, Student Affiliate
2004 – present   American Psychological Association, Student Affiliate

PROFESSIONAL ACTIVITIES

2005 – present   Joint Doctoral Program Graduate Student Mentor
2005   Joint Doctoral Program Faculty Selection Committee
2005   Joint Doctoral Program Student Selection Committee
2003   Joint Doctoral Program Student Selection Committee

PROFESSIONAL PUBLICATIONS

Keil, V., & Price, J.M. (2006). Externalizing behavior disorders in child welfare settings: Definition, prevalence, and implications for assessment and treatment. *Children and Youth Services Review, 28,* 761-779.

Keil, V., & Price, J.M. (in press). Social information-processing patterns of maltreated children in two social domains. *Journal of Applied Developmental Psychology.*

MANUSCRIPTS UNDER REVIEW OR IN PREPARATION

Keil, V., & Price, J.M. (under review). Treatment of externalizing behavior disorders in child welfare settings: Integrating evidence-based practice.

Price, J.M., & **Keil, V.** (under review). Social adjustment of maltreated children in early elementary school.

Price, J.M., **Keil, V.**, & Roesch, S.C. (in preparation). Developmental patterns of social problem-solving among maltreated children.

Keil, V. (in preparation). Psychometric properties of the Parent Daily Report and its potential for use in child welfare settings.

PROFESSIONAL PRESENTATIONS

Keil, V., Price, J.M., & Chamberlain, P. (2007, March). *Latent class analysis of the Parent Daily Report.* In V. Keil (Co-Chair, Organizer), Applications of latent class analysis to developmental psychopathology. Symposium presented at the biennial meeting of the Society for Research in Child Development, Boston, Massachusetts.

Keil, V., & Price, J.M. (2005, March). *Developmental patterns of social problem-solving among maltreated children.* In V. Keil (Organizer), Adjustment of school-aged maltreated children: Assessment issues, risk factors, and outcomes in two longitudinal investigations. Symposium presented at the biennial meeting of the Society for Research in Child Development, Atlanta, Georgia.

Keil, V., & Price, J.M. (2004, July). *Social information processing patterns among maltreated children.* Poster presented at the annual meeting of the American Psychological Association conference, Honolulu, HI.

Keil, V., & Price, J.M. (2004, April). *Social information processing patterns and peer conflict among maltreated children.* Poster presented at the annual meeting of the Western Psychological Association conference, Phoenix, AZ.

CLINICAL/DIDACTIC COLLOQUIA

Clinical Internship Seminars, Children's Hospital, Orange County
September 2006 – present
- Attend weekly presentations in pediatric psychology, focusing on topics relevant to pediatric medical settings, such as oncology, endocrinology, and palliative care
- Participate in weekly discussions about multicultural and ethical issues in psychology
- Attend biweekly supervision and training in intake interviews and diagnostic formulations
- Participate in weekly consultation and liaison rounds with psychiatrist with emphasis on case presentations and differential diagnoses
- Attend weekly training in assessment with emphasis on neuropsychological evaluations

Statistical Methods, Child and Adolescent Services Research Center
January 2005 – July 2006
- Attended workshops by leading statisticians to learn new data analytic methods
- Acquired proficiency in SPSS, HLM and M Plus software

Cognitive-Behavioral Interventions Seminars, Veteran Affairs Medical Center
February 2005 – December 2005
- Attended weekly, 60-minute workshops by VA faculty and staff on topics related to CBT
- Emphasis on adapting CBT to address comorbid conditions, Axis II symptomatology, and co-occurring medical conditions
- Special topics included PTSD and exposure therapy, personality disorders, CBT for insomnia, CBT for pain and other medical conditions, bipolar disorder, and psychosis

Assessment Didactics, UCSD Child and Adolescent Psychiatry Services
July 2004 – July 2005
- Attended weekly 60-minute seminars on assessment with psychology interns
- Learned how to administer, score, and interpret several instruments such as WISC-IV, MMPI-A, WMS-III, CMS, CVLT, DKEFS, WCST, NEPSY
- Presented challenging assessment cases to illustrate and discuss obstacles to testing

Trauma Counseling Seminars, Chadwick Center for Children and Families, Children's Hospital, San Diego
September 2003 – December 2003
- Attended weekly 60-minute workshops focusing on working with victims of sexual trauma
- Special topics included assessment of trauma, mandated reporting, cultural competency, domestic violence, and comorbid substance abuse
- Trained in administration and scoring of screening measures used by Chadwick Center for assessment of child maltreatment

Privacy and Ethical Issues, Child and Adolescent Services Research Center
February 2003
- Attended a full-day workshop presented by Celia Fisher, Ph.D on ethics and HIPAA
- Received certificate of training in HIPAA regulations

ABSTRACT OF THE DISSERTATION

Psychometric Properties of the Parent Daily Report

and Its Potential for Use in Child Welfare Settings

by

Vivien Keil

Doctor of Philosophy in Clinical Psychology

University of California, San Diego, 2007

San Diego State University, 2007

Professor Joseph M. Price, Chair

Youth in child welfare settings exhibit high rates of externalizing behavior disorders, which place them at increased risk for placement disruptions. These placement disruptions are linked to further increases in externalizing problems. Given these mental health concerns and also the economic constraints of child welfare settings, there is a need for reliable, valid, and efficient measures of externalizing behavior problems. The Parent Daily Report (PDR) is one particularly promising

measure of problematic child behaviors. It is a 31-item parent observation report of behavior problems such as defiance, fighting, and arguing. The present investigation represents the first large-scale, systematic examination of the psychometric properties of the PDR using a diverse child welfare sample. The sample consisted of 700 children (52% female) between the ages of 5 and 12 years ($M = 8.8$ years) who were placed in relative or nonrelative foster care in San Diego County between 1999 and 2004, and whose caregivers participated in an effectiveness trial of a parent management training and support intervention. The first study aim was to examine the factor structure of the PDR using exploratory factor analysis, which uncovered a two-factor structure. The second and third aims were to compare the factor structure across language and gender. Multiple group analysis across language revealed partial equivalence between the English and Spanish versions of the PDR. Multiple group analysis across gender revealed an equivalent factor structure for males and females. The final study aim was to determine whether the PDR could be used to identify distinct subgroups of children. Latent class analysis supported a three-class solution, based on the severity of behavioral disturbance (i.e., low, moderate, severe). These findings collectively suggest that the PDR has sound psychometric properties when used with a diverse child welfare sample. Further, the measure is clinically valuable in that it can be used to identify youth that are at high risk for placement disruptions and escalation of externalizing problems. Additional implications for assessment, treatment, and future research in child welfare settings are discussed.

CHAPTER 1

Introduction

Of the many challenges facing the field of psychology, perhaps the greatest obstacle is the accurate assessment of behavior. Since its inception, biased and inappropriate measures have threatened to undermine the legitimacy of psychology. For example, in his book entitled *The Mismeasure of Man*, Stephen Jay Gould (1996) traced the innumerable mistakes that have been made in an attempt to quantify intelligence. Some of the more astounding approaches to measuring intelligence included comparing the shape of the human skull to that of apes and other simians, examining the anatomical dominance of the forehead region, and simply calculating the size of skulls. These misguided approaches may seem absurd in the present day, but during their times these theories were regarded as legitimate scientific findings. These mistakes have left an indelible mark on psychology such that accurate assessment has become one of the foremost goals of the field. The current emphasis on assessment issues can be seen in several recent articles (e.g., De Los Reyes & Kazdin, 2005; Snowden, 2003), texts (e.g., Groth-Marnat, 2003; Sattler, 2002), and national practice guidelines (e.g., American Psychological Association, 2002; U.S. Department of Health & Human Services [USDHHS], 2001b).

Given the recent findings about biases in assessment and the importance of accurate assessment, it is imperative that the reliability and validity of measures be systematically evaluated. The present investigation represents one such attempt in that it is an evaluation of the psychometric properties of a measure of child behavior

1

problems. More specifically, the goal is to examine the factor structure of the Parent

Daily Report (PDR; Chamberlain & Reid, 1987) and its potential for use in child

welfare settings.

Prevalence of Behavior Problems in Child Welfare Settings

Within the context of child welfare settings, valid and reliable measures are

needed for externalizing problems in particular. In his review of the literature

published between 1978 and 1994, Pilowsky (1995) concluded that externalizing

disorders were the most predominant and problematic types of disorders among foster

children. The prevalence of externalizing problems has been found to range from

approximately 20% (Thompson & Fuhr, 1992) to as high as 78% (Silver et al., 1992),

with a mean of about 42% across several investigations (see Keil & Price, 2006a,

for a more thorough discussion of prevalence rates). At the low end of the range,

Thompson and Fuhr (1992) examined 50 children in out-of-home care, including

foster care, kinship care, and residential settings, and found that 20% of the sample

scored above the 98th percentile on the Total Problems and Externalizing scales of the

Child Behavior Checklist (CBCL). In contrast, Silver et al. (1992) found much higher

rates of behavior disorders using the Diagnostic Interview Schedule for Children

(DISC) with 382 children in residential settings across 4 states (Colorado, Wisconsin,

New Jersey, Mississippi). They determined that 77.4% of their sample had some form

of conduct disorder (CD) and 39.1% had severe CD. The sharp contrast in prevalence

rates between these two studies is likely due to a number of factors such as definitions

and measures used. It should also be noted that the latter study involved only children

in residential settings while the former study included children from foster care, kinship care, and residential settings. Given that children in residential homes tend to have more severe behavioral profiles, as compared to children in other types of out-of-home care (Helfinger, Simpkins, & Combs-Orme, 2000), the higher rates found in the Silver et al. (1992) study are not surprising.

For the vast majority of the studies reviewed, prevalence rates fell between the 20-78% range previously described. For example, Urquiza, Wirtz, Peterson, and Singer (1994) examined 167 children entering child protective custody in Sacramento, California and found that the rates of CD, oppositional defiant disorder (ODD), and attention-deficit/hyperactivity disorder (ADHD) ranged from 38-58%. In a multi-site California study (San Diego, Monterey, Santa Cruz) with 140 children entering foster care, Clausen, Landsverk, Ganger, Chadwick, and Litrownik (1998) found that an average of about 40% of the children scored in the clinical range on the Externalizing scale of the CBCL. Thus, findings from several studies with diverse sample characteristics suggest that, on average, approximately one-half of youth entering child welfare settings exhibit externalizing behavior problems.

High rates of externalizing behavior problems are cause for concern because they are linked with placement stability and disruptions. Foster children with externalizing problems have been found to be half as likely to be reunified with their birth parents within 18 months of entry into foster care, as compared to those without significant externalizing problems (Landsverk, Davis, Ganger, Newton, & Johnson, 1996). As a result, children exhibiting externalizing behaviors are spending longer

amounts of time in child welfare settings. Further, results from recent investigations by Chamberlain and colleagues (2006) and Newton, Litrownik, and Landsverk (2000) suggest that foster children with high rates of externalizing problems are more likely to experience placement disruptions, and these placement disruptions are related to further increases in externalizing problems. Thus, high levels of externalizing problems place the children at risk for future disruptions in their relationships with caregivers, which in turn serves to perpetuate and increase maladaptive behavior patterns. Given the high prevalence of behavior problems in child welfare settings and its related adverse consequences, the question of appropriate assessment is paramount.

Assessment of Behavior Problems in Child Welfare Settings

There are a number of instruments available for assessing the externalizing behavior problems of children and adolescents. For example, there are behavior rating scales, diagnostic interviews, observational methods, and sociometric ratings. At present, the use of behavior rating scales such as the CBCL is the most predominant approach in child welfare settings (e.g., Clausen et al., 1998; Helfinger et al., 2000; Hulsey & White, 1989; Keil & Price, 2006a; McIntyre & Keesler, 1986; Thompson & Fuhr, 1992; Urquiza et al., 1994). Such measures are typically completed by a parent or teacher and provide a standardized format for assessing a child's or adolescent's behavioral characteristics. Although there are several other assessment approaches, the practical appeal of behavior ratings scales for child welfare settings is that they are relatively time efficient and cost effective.

The most commonly used rating scale is the Child Behavior Checklist (CBCL; Achenbach, 1991a) and Teacher's Report Form (TRF; Achenbach, 1991b) for children from age 4 to 18. These instruments were designed to assess behaviors in the past 6 months and thus require that raters have known the child for at least this duration of time. The CBCL and TRF include 120 problem behavior items that are rated on a 3-point scale: 0 = not true, 1 = sometimes true, 2 = often true. These items form nine subscales: Schizoid or Anxious, Depressed, Uncommunicative, Obsessive-Compulsive, Somatic Complaints, Social Withdrawal, Hyperactive, Aggressive, and Delinquent. These narrow-band scales are grouped into two broad-band scales, the Internalizing Scale and Externalizing Scale. The psychometric properties of the CBCL and TRF have been extensively researched and range from adequate to excellent (Achenbach, 1991a, 1991b). The CBCL has been shown to distinguish between clinical and normative samples. Achenbach (1991a) also found that the CBCL correlates highly with related measures such as the Revised Behavior Problem Checklist and Connor's Rating Scales. Test-retest reliability taken at 1-week intervals is in the .80 to .90 range and remains strong at 3-, 6-, and 18-month intervals.

The Revised Behavior Problem Checklist (RBPC; Quay & Peterson, 1987, 1996) is similar to the CBCL in its structure and purpose. The RBPC assesses problem behaviors for youth between the ages of 5 and 16. Raters should be adults who are sufficiently familiar with the child in question, although no explicit time frame is specified. It is a relatively short instrument, with 89 problem behaviors that are rated according to a 3-point scale: 0 = not a problem, 1 = mild problem, 2 = severe problem.

The items are grouped into 6 empirically derived scales: Conduct Disorder, Socialized Aggression, Attention Problems-Immaturity, Anxiety-Withdrawal, Psychotic Behavior, and Motor Excess. The instrument has adequate to excellent internal consistency, with all scales having alphas greater than .70. Quay and Peterson (1987) found that the RBPC can differentiate between clinical and normative groups and also special education groups and normative groups, providing evidence for its construct validity. The manual also indicates that the Conduct Disorder scale correlates significantly ($r = .60$) with playground observations of aggressive behavior.

The Eyberg Child Behavior Inventory (ECBI; Robinson, Eyberg, & Ross, 1980) is a 36-item behavior rating scale that was explicitly designed to assess behavior problems in children via parental reports. This measure has two components: the Intensity score, which examines the frequency of behaviors on a 7-point Likert scale ranging from 1 ("never") to 7 ("always"), and the Problem score, which indicates whether or not a behavior is a problem on a 2-point "yes" or "no" scale. This measure must be completed by an adult familiar with the child in question. Studies examining the psychometric properties of the ECBI suggest that it has adequate to strong reliability and validity. Eyberg and Robinson (1983) reported an internal consistency alpha value of .98 for both sections of the ECBI, and it has strong convergent validity with the CBCL, particularly the Externalizing scale (Boggs, Eyberg, & Reynolds, 1990). Robinson and Eyberg (1981) also found that the ECBI relates significantly to direct observational measures of noncompliance and negative parent-child interactions.

The Conners' Parent Rating Scale and Conners' Teacher Rating Scale

(Conners, 1990) was created to help identify behavior problems in children between

the ages of 3 and 17. The 93-item full version of the parent scale consists of eight

factors: Conduct Disorder, Fearful-Anxious, Restless-Disorganized, Learning

Problem-Immature, Psychosomatic, Obsessional, Antisocial, and Hyperactive-

Immature. Each item is rated on a 4-point problem scale from 0 ("not at all") to 3

("very much"), typically by parents and teachers who are familiar with the child. The

teacher version has fewer items and factors, but items are scored in the same manner.

Studies indicate that both the parent and teacher versions have adequate reliability

and validity (Conners, 1990) and test-retest reliability for the teacher version was .72

over a 1-month interval (Conners, 1969). Margalit (1983) concluded that the scales

differentiated between behavior disordered and nonspecial education children,

providing support for the construct validity of the Conners' scales.

The Behavior Assessment System for Children (BASC; Reynolds & Kamphaus,

1992) is a more comprehensive instrument that combines aspects of the CBCL, RBPC,

and ECBI. The BASC is designed to assess a variety of problem behaviors, school

problems, and adaptive skills through parent and teacher rating scales and also self-

report forms for children and adolescents. As with most of the behavior rating scales

described, the instructions do not contain a specific time frame necessary for the raters

to have known the child in question. The various forms have anywhere from 126 to

148 items and the items are rated by circling adjacent letters indicating how frequently

each behavior is perceived to occur, ranging from never to almost always. Like the

CBCL, the BASC has both externalizing and internalizing components. The BASC

also includes a Behavioral Symptoms Index, which is a composite problem behavior

score with an internal consistency ranging from .88 to .97. Test-retest reliability ranges

from .70 to .80, with time intervals ranging from 2 to 8 weeks. The BASC had

produced strong internal consistency values, typically ranging from .80 to .90. Lett and

Kamphaus (1997) concluded that the BASC has sound factor structure and good

construct validity.

In summary, there are several behavior rating scales that may be useful in

assessing behavior problems among youth in child welfare settings. The CBCL is by

far the most commonly used instrument, although there is evidence to suggest that

the RBPC, ECBI, Conners', and BASC may also be useful (Keil & Price, 2006a).

However, it is important to remember that the measures described were not developed

specifically for use in child welfare settings, nor have they been evaluated with this

population. Thus, the reliability and validity of these measures may be in question

when administered in these settings.

Characteristics of Child Welfare Settings

Clearly, there are several potential behavior rating scales that can be used to

assess child behavior problems. When choosing from these measures, it is imperative

to consider the unique characteristics and practical constraints that exist when working

within child welfare settings. These settings present a unique challenge for mental

health workers attempting to assess problematic child behaviors for several reasons.

First, there are economic constraints that must be considered. Funding for child

welfare services has been stretched thin as the number of youth in child welfare settings has increased dramatically in the last two decades. From 1982 to 1995, the total number of children in foster care increased by 63.2%, from 435,000 to 710,000 (Curtis, 1999). During that same period, federal expenditures for foster care maintenance and administration skyrocketed from $309 million to $3.05 billion. These staggering numbers have continued to increase (Courtney, 1999). This growth in the number and costs of the child welfare population, coupled with inadequate federal finding, has resulted in an extremely overwhelmed social system that Curtis (1999) described as "in crisis."

Given these conditions, it is not surprising that caseworkers have expressed that it is extremely difficult, if not impossible, to keep up with their constantly increasing caseloads. In their qualitative analysis of caseworkers' experiences, B. D. Smith and Donovan (2003) identified resource limitations and time pressures as the two central challenges facing caseworkers. They point out that the need for efficient and effective practices was underscored in 1997, when the federal Adoption and Safe Families Act (ASFA, U.S. Public Law 105-89) established decision-making time frames for child welfare caseworkers. Further, given the recent move towards a managed care model for Medicaid and mental health services, there is an even greater push for cost effectiveness and service efficiency (Embry, Buddenhagen, & Bolles, 2000). Collectively, these real-world constraints suggest that measures used in child welfare settings must be affordable, effective, and easily administered and interpreted.

Beyond these more straight-forward economic constraints, there are a number of other challenges that arise. Youth in child welfare represent a population typically characterized by a great deal of instability. This instability is partially due to the inadequate care of their primary caregivers, which has resulted in their placement into the care of child welfare. Furthermore, this is often compounded by experiences that occur after being placed into care, such as multiple foster care placements and school transfers. These common experiences of youth in child welfare settings have significant implications for assessment. In particular, most assessments require that the respondent be familiar with the behavior of the child or adolescent in question (e.g., 6 months in the case of the CBCL). This necessary level of familiarity may be difficult or even impossible to achieve given the transitory experiences of many of the children in care and the structure of child welfare settings. Thus, there is a need for a measure of behavior problems that does not require a long period of familiarity with the child or adolescent.

The diversity of youth in child welfare settings also impacts which measures should be chosen to assess child behavior problems. Children and adolescents in child welfare are an extremely heterogeneous population with regards to their ethnic/racial backgrounds, cultures, and languages. National statistics have consistently shown that minority groups are overrepresented in the child welfare system. A recent census indicated that the ethnic composition of youth in child welfare settings is approximately 42% African American, 36% Caucasian, 15% Latino, 2% Native American, and 1% Asian/Pacific Islander (USDHHS, 2001a). In the state of

California, the Latino population is particularly dominant. Approximately one-third of the children in California's child welfare system are identified as Latino, and a large number of these children come from primarily Spanish-speaking families (Needell et al., 2001; U.S. Census Bureau, 2000). In some counties in California, Latinos comprise more than half of the child welfare population (Hines, Lemon, Wyatt, & Merdinger, 2004), which is not surprising given that Latino children and families constitute the fastest growing ethnic group in the child welfare system (Rivera, 2002). In a recent review of service use among minority groups in California, Ayon and Lee (2005) emphasized that minority children constitute a majority in the child welfare system. Given these findings, any measure of child behavior problems used in child welfare settings must be appropriate for use with these diverse groups.

Taken together, these findings suggest that practical utility must play a role in determining which measures are used for assessment of child behavior problems in child welfare settings. Any measure widely disseminated into such a large social system must be time-efficient and cost-effective, and must possess sound psychometric properties across the diverse groups served by child welfare.

The Parent Daily Report: A Promising Measure for Child Welfare Settings

Description of the PDR. Given the contextual characteristics of child welfare settings, the Parent Daily Report (PDR) checklist is a particularly promising measure of child behavior problems (Chamberlain & Reid, 1987; see Figure 1). It is a parent observation measure consisting of 31 items describing a variety of problem behaviors (e.g., complaining, lying, fighting) drawn from items previously developed by

Child Name:	#	Behavior	__/__/__ Mon	__/__ Tue	__/__ Wed	__/__ Thurs	__/__ Fri	__/__ Sat	__/__ Sun
_____	1	Argue							
	2	Back-talk							
	3	Wet							
Week of:	4	Competitive							
__ _/__ _/__ _to	5	Complain							
__ _/__ _/__ _	6	Defiant							
	7	Destructive/vandalism							
Caller Initials:	8	Fight							
___ ___	9	Irritable							
	10	Lie							
	11	Negative							
Notes:	12	Boisterous/rowdy							
	13	Not mind							
_____	14	Stay out late							
_____	15	Skip meals							
_____	16	Run away							
_____	17	Swear/use bad language							
_____	18	Soil							
_____	19	Tease/provoke							
_____	20	Depressed/sad							
_____	21	Sluggish							
_____	22	Jealous							
_____	23	Truant (skipped school)							
_____	24	Steal							
_____	25	Nervous/jittery							
_____	26	Short attention span							
_____	27	Daydream							
_____	28	Irresponsible							
_____	29	Drugs/alcohol							
_____	30	School problem							
_____	31	Inappropriate sexual Activity							

Figure 1: Parent daily report telephone log

Patterson (1964). To administer the PDR, the parent is asked via telephone which, if any, of the listed behaviors occurred during the previous 24-hour period. During each call, a trained interviewer asks the parent, "Thinking about (child's name) during the past 24 hours, did any of the following behaviors occur?" The parent is not asked about frequency of behavior, only occurrence or nonoccurrence during the previous 24-hour period (i.e., "yes" or "no"). This format makes the PDR distinct from the other behavior measures reviewed in that frequency and intensity are not of interest. Instead, the PDR essentially provides a sampling of problematic behaviors that are occurring. The PDR Total Score is simply the sum of the "yes" responses. The PDR checklist takes about 5 to 10 minutes to complete and is typically repeated on 3 to 5 separate occasions to get a stable estimate of child behavior problems, as experienced by the caregiver (Chamberlain et al., 2006). Chamberlain and Reid (1987) note that the PDR can be administered reliably after only 1 hour of training.

The structure of the PDR (i.e., requiring parents to recall only the past 24 hours) is intended to reduce systematic and random sources of measurement error in order to increase the reliability and validity of the parents' reports of child behavior problems (Chamberlain et al., 2006). Prior research has shown that the reporter's current emotional state may lead to biased estimates of such reports (Bower, 1981; Webster-Stratton & Hammond, 1990), and that reporters tend to give more weight to peak and recent levels of experiences, rather than giving equal weight to each instance (Stone, Broderick, Kaell, DelesPaul, & Porter, 2000). Given that retrospective reports tend to be biased (Tourangeau, 2000), the PDR was designed to avoid the need for aggregate

recall over a number of days or for estimates of the frequency and intensity of behaviors.

The specific item content of the PDR has changed over time and settings since its initial development. Researchers have essentially added or deleted items depending upon the unique characteristics of their sample. For example, the Chamberlain and Reid (1987) version contained 34 items, and was not initially used with child welfare samples. This version included "firesetting," "fighting with siblings," "police contact," and "parents spank," but did not include "inappropriate sexual activity." The PDR version used by Webster-Stratton and colleagues (Webster-Stratton & Hammond, 1997; Webster-Stratton, Hollinsworth, & Kolpacoff, 1989; Webster-Stratton, Kolpacoff, & Hollinsworth, 1988) contained 38 items, although they did not explicitly list the items that they included or describe the changes that they made. Their version was used with families from low socioeconomic backgrounds, but not child welfare samples. The only study that previously used the PDR with a child welfare sample is that of Chamberlain, Moreland, and Reid (1992). In this instance, the PDR included 36 items, with the additional items of "alcohol," "whining," "temper tantrums," "inappropriate laughing," and "talks to self." The version of the PDR that is of particular interest in the current study includes 31 items and was utilized recently with a large child welfare sample (Chamberlain et al., 2006; Price, Chamberlain, Reid, & Landsverk, 2006).

Existing research on the PDR. At present, examinations of the reliability and validity of the PDR are limited. Jones (1974) found 85% inter-caller agreement when

using a second interviewer on an extension phone. Reid and Patterson (1976) reported

no significant differences among weekly PDR behavior scores over a 4-week baseline

period, providing evidence for its temporal stability. Weinrott, Bauske, and Patterson

(1979) reported that the PDR has adequate reliability, with an inter-interviewer

correlation of .98, and test-retest values ranging from .60-.82. Patterson (1976)

provided support for the concurrent validity of the PDR in that he found that the

measure was significantly associated ($r = .69$) with child behavior problems recorded

by home observers. The concurrent validity of the PDR has also been demonstrated in

association with observational measures of family functioning (Forgatch & Toobert,

1979; Patterson, 1976) and parents' global ratings of child behavior (i.e., the Becker

Adjective Checklist; Becker, Madsen, Arnold, & Thomas, 1967).

In the primary study of the psychometric properties of the PDR, Chamberlain

and Reid (1987) found favorable results. They audiotaped 9% of the PDR calls so

that a second caller could score the interview and found that the mean inter-caller

agreement was 97.6%. In order to examine temporal stability, they compared the PDR

scores for the first 6 days versus the last 6 days of their 4-week normative study and

found a high correlation ($r = .82$). Collectively, these findings support the inter-caller

and temporal stability of the PDR.

Chamberlain and Reid (1987) also conducted cluster analyses and found two

primary factors with good internal consistency values: aggression (alpha = .88,

including items such as "argue," "talk back," and "noncompliance") and immaturity

(alpha = .80; including items such as "complain," "irritable," and "negative"). The

analyses also uncovered two secondary factors with moderate internal consistency values: unsocialized (alpha = .65; including items such as "stealing" and "lying") and retaliation (alpha = .65; including items such as "destructive" and "not eating meals"). Despite these favorable findings, there are several limitations that must be addressed. The sample was composed of only 81 children, with 62% of them being male and 92% of them being Caucasian. These sample characteristics severely limit generalizability and suggest that there is a need for additional evaluations of the PDR with a more diverse population. Given that the study is quite dated, the need for additional research is further reinforced.

Current use of the PDR. Despite the limited number of systematic evaluations of the PDR, major research groups have utilized the measure to assess child behavior problems, particularly in treatment outcome studies. The PDR has a long history of use, beginning in the 1970s and extending to the present day. Patterson (1976) suggested that the PDR is a useful treatment outcome measure for behavioral interventions targeting conduct problems. Weinrott et al. (1979) used it as a measure of treatment outcome for children referred for conduct problems. They found that the PDR was sensitive to change after a behavioral intervention for parents. Reid, Kavanagh, and Baldwin (1987) used the PDR in order to examine abusive parents' perceptions of child behavior problems. They found that abusive parents endorsed more problematic child behaviors than nonabusive parents, although independent home observations indicated no such differences in child behaviors.

Since those earlier studies, researchers have continued to utilize the PDR in order to monitor change and improvement in treatment outcome studies. Chamberlain and colleagues have utilized the PDR in several studies involving treatment of high-risk youth. Patterson, Chamberlain, and Reid (1982) used the PDR in their evaluation of a training program for parents of aggressive children. They averaged baseline and termination PDR data and found that the treatment group showed a significant decline in child behavior problems. Chamberlain and Reid (1991) examined youth that had recently been discharged from psychiatric hospitals and were subsequently placed in treatment foster care. They used the PDR as a measure of treatment outcome at baseline, 3 months, and 7 months posttreatment. They found that the youth in treatment foster care showed significant reductions in behavior problems at both 3 months and 7 months posttreatment. Chamberlain et al. (1992) used the PDR for similar purposes in an evaluation of training for foster parents. They compared baseline PDR data to that obtained 3 months later and found that the foster parents that received additional support and training reported the greatest decline in problematic child behaviors. Collectively, the Chamberlain studies suggest that the PDR is sensitive to treatments targeting child behavior problems.

Webster-Stratton and colleagues have also utilized the measure frequently in their intervention studies targeting aggressive children. Webster-Stratton et al. (1988) used the PDR to examine the effectiveness of their behavioral parent training program in reducing child behavior problems. Their treatment groups included a videotaped modeling condition, a videotaped modeling and group discussion condition, and a

group discussion condition. They found that all three treatment conditions showed decreases in problematic behaviors, as measured by the PDR. The same pattern of findings was replicated in a subsequent long-term effectiveness study of the same three treatment conditions (Webster-Stratton et al., 1989). In a more recent investigation, Webster-Stratton and Hammond (1997) used the PDR in their study of a child training intervention, a parent training intervention, and a combined child and parent training intervention, all designed to reduce conduct problems. They again found that all three of their treatment groups exhibited a significant decline in behavior problems on the PDR. The findings of Webster-Stratton and colleagues are consistent with those of the Chamberlain research group in that they provide evidence that the PDR is a sensitive and useful measure of child behavior problems.

The treatment foster care model espoused by Chamberlain and colleagues has recently been adapted to meet the needs of the increasing number of preschool-aged foster children. The Early Intervention Foster Care (EIFC) program described by Fisher, Gunnar, Chamberlain, and Reid (2000) was designed as a preventive intervention for high-risk preschool-aged foster children. In their preliminary evaluation of the EIFC program, these researchers used the PDR as a measure of child behavior problems and foster parents' stress as a result of the behaviors. They found that the parents in the treatment group showed a decrease in stress in response to child behavior problems, although it should be noted that the findings were only marginally significant ($r = .08$). Posthoc analyses indicated that while the EIFC group showed a decrease in stress levels, those in the foster care as usual group exhibited an increase in

stress levels. Since this pilot study, there had been one subsequent study of the EIFC program focusing on placement stability. Fisher, Burraston, and Pears (2005) found that the intervention resulted in significantly lower rates of placement disruptions. However, since the focus of their investigation was placement stability, child outcomes were not specifically addressed.

Most recently, Price et al. (2006) found that the level of behavior problems among foster children, as measured by the PDR, significantly decreased after foster parents received support and training in parent management skills. In addition, Chamberlain et al. (2006) provided evidence that the PDR may have great clinical value in child welfare settings. These researchers emphasized the need to identify risk factors for costly placement disruptions, and they found that the baseline PDR data was a significant predictor of placement stability. More specifically, the baseline PDR scores increased the risk of disruption by 17% for every child behavior problem reported. Further, for each additional behavior over 6, the risk of placement disruption increased by 25%. Child gender, child ethnicity, child age at baseline, foster parent ethnicity, and the total number of children in the foster home were not linearly related to risk of placement disruption. These findings suggest that the PDR may be a particularly valuable measure for child welfare settings, especially given the recent emphasis on improving placement stability in foster care (*Federal Guidelines*, 2001).

In summary, the PDR has become an increasingly popular measure, especially for treatment outcome studies. Its brevity and sensitivity to change undoubtedly contribute to its appeal. Further, although there are limited systematic evaluations of

the PDR, the existing evidence suggests that the measure has adequate psychometric properties. Given that the PDR is being used as an outcome measure in several treatments studies, there is a crucial need to understand its psychometric properties and factor structure in particular. An examination of the factor structure of the PDR will help provide insight into what construct the instrument is measuring and also how the construct operates across groups of interest. In addition, since there is reason to believe that the PDR has clinical value in child welfare settings (i.e., it predicts placement disruption), its reliability and validity needs to be examined using a large, diverse child welfare sample. The present investigation embarks upon this necessary line of research.

Current Project: Filling in the Missing Pieces

The broad goal of the present investigation is to examine the psychometric properties of the PDR using a large and diverse child welfare sample and to determine its potential for use in child welfare settings. This goal is represented by four specific study aims:

- *Aim 1:* Examine the factor structure of the PDR using exploratory factor analysis.

- *Aim 2:* Compare the factor structure of the PDR across English and Spanish versions of the PDR using multiple group analysis.

- *Aim 3:* Compare the factor structure across gender (i.e., males and females) using multiple group analysis.

- *Aim 4:* Determine whether there are distinct subgroups of youth based on the 31 items of the PDR using latent class analysis.

The first aim is to examine the factor structure of the PDR. The cluster analysis conducted by Chamberlain and Reid (1987) yielded two primary clusters (i.e., aggression, immaturity) and two secondary clusters (i.e., unsocialized, retaliation). Given that this is the only evaluation of the structure of the measure, it is hypothesized that there will be two factors (*Hypothesis 1*). The first factor will be composed of items describing oppositional and disruptive behaviors such as "argue" and "fight." The nature of the second factor is less predictable, although the cluster analysis of Chamberlain and Reid (1987) suggests that it may consist of developmentally immature and nuisance behaviors such as "complain" and "irritable."

Beyond the basic factor structure, it is necessary to know whether there is factor equivalence across groups of interest. The second and third study aims are designed to determine if the factor structure of the PDR is similar across the diverse groups represented in child welfare settings. Given the large number of Latino and Spanish-speaking families involved with child welfare, the second study aim will address the crucial question of whether or not the factor structure is equivalent across language groups. Cultural or linguistic variables may result in differences in response patterns on the PDR, which would affect the psychometric properties of the Spanish version. It is possible that Spanish-speaking populations may perceive child behavior problems or the measure in a different manner, perhaps resulting in discrepant factor structures. Thus, before the translated PDR is widely disseminated into child welfare

settings, there must be a systematic evaluation of its psychometric properties. This study represents the first examination of the factor structure of both the English and Spanish versions of the PDR.

Examining the factor structure of different linguistic versions has been used previously to establish the construct validity of translated measures. Rio, Quay, Santisteban, and Szapocznik (1989) evaluated the Spanish translation of the Revised Behavior Problem Checklist (RBPC). These researchers used the data from 144 Latino males to determine if the factor structure of the Spanish RBPC was comparable to that of the original English version. They found congruence for five of the six scales of the RBPC: Conduct Disorder, Socialized Aggression, Attention Problems-Immaturity, Anxiety-Withdrawal, and Motor Excess. They did not find a Psychotic Behavior factor on the Spanish version. Since the findings generally supported factor equivalence, Rio et al. (1989) concluded that the RBPC is suitable for use with Latino males. Similarly, it is hypothesized that there will be factor equivalence across the English and Spanish versions of the PDR (*Hypothesis 2*).

The Rio et al. (1989) study did not include females for unspecified reasons. However, the researchers acknowledged that their findings could not generalize to Latino females. The present investigation strives to generalize to both genders and thus the third study aim will examine whether the factor structure of the PDR varies across gender. This analysis is needed particularly because females are often underrepresented (e.g., Chamberlain & Reid, 1987) or even excluded (e.g., Rio et al., 1989) in studies examining measures of behavior problems, perhaps because males are

perceived as the more aggressive gender and are more often referred for behavioral

disturbance (Gorman-Smith & Loeber, 2005). Recently, there has been sharp criticism

from researchers about the lack of attention on aggression among females (e.g.,

Crick, Grotpeter, & Bigbee, 1995; Maccoby, 2004). Given that females have been so

commonly overlooked in studies of behavior problems, the present investigation will

examine the factor structure of the PDR among females to determine if it is equivalent

to the factor structure found for males. Since there are no specific data to suggest that

there will be gender differences on the PDR, it is hypothesized that there will be factor

equivalence for males and females (*Hypothesis 3*).

The present investigation will also contribute to the existing literature on the

PDR by examining variables that may impact the caregivers' report of problematic

child behaviors. Given that retrospective reports of behavior tend to be quite biased

(e.g., Bower, 1981; Sattler, 2002; Stone et al., 2000; Tourangeau, 2000), the format of

the PDR, with it focus on occurrence or nonoccurrence of specific behaviors, was

intended to reduce the subjectivity of reporting (Chamberlain et al., 2006). In order to

examine the impact of reporter characteristics, it is necessary to empirically examine

whether there is invariance across groups of interest. The present investigation will

consider characteristics of the caregivers that may impact their reporting of child

behaviors, specifically caregiver age, caregiver ethnicity, kinship status (i.e., kin vs.

nonkin), and the number of days the child was in the home at baseline. The ethnic

match between caregiver and foster child (i.e., match or mismatch) will also be

explored. These particular variables were chosen because they could impact the

familiarity of the foster parent with the child and thus result in biased reporting of child behavior.

The final aim of this investigation is to determine whether there are distinct, clinically useful categories of children, as measured by the PDR. This aim is largely exploratory in that no previous investigations have examined potential subgroups or classes of children using the 31 PDR items. Child characteristics that may impact class membership, specifically child age, gender, and ethnicity, will be explored. This aim may reveal distinct typologies of children based on the presence or absence of certain behavior problems, or perhaps severity of behavioral disturbance.

In summary, the present investigation represents the first large-scale, systematic evaluation of the psychometric properties of the PDR. In addition to an in depth examination of the factor structure of the PDR, this investigation will also provide insight into the clinical value of the PDR and its potential for use in child welfare settings.

CHAPTER 2

Method

Participants

    Recruitment. The sample for this investigation consisted of 700 kinship (relative) and nonkinship (nonrelative) substitute caregivers who were recruited between 1999 and 2004 in San Diego County for an effectiveness study of a parent training and support intervention. Data systems made available from the social service agency were reviewed on a weekly basis to identify eligible children and foster families. The eligibility requirements were as follows: (a) the child had been in either a relative or nonrelative foster care placement for a minimum of 30 days; (b) the child was between the ages of 5 and 12; and (c) the child was not considered medically fragile. Children in group homes were not included in the study.

    Once deemed eligible, families were randomly assigned to either the parent training condition or to the "services as usual" condition. The foster parents were contacted by phone and presented with a brief overview of the project and their assigned condition. Of the eligible foster parents in San Diego County who were contacted, 62% agreed to participate and 38% declined. Reasons given for declining to participate included: too busy, too much work, too many children (50%); not interested (43%); family health problems (2%); and concerns about participating in research (5%).

    Intervention and control procedures. Participants randomly assigned to the intervention group received 16 weeks of foster/kinship family support and training in

behavior management methods, which was named *Project KEEP* ("Keeping Foster and Kin Parents Supported"). Intervention groups consisted of 3 to 10 foster parents and were conducted by a trained facilitator and co-facilitator team. Curriculum topics included encouraging cooperation, using behavioral contracts, limit setting, dealing with difficult problem behaviors, school involvement, encouraging positive peer relationships, and strategies for managing stress.

Each session was structured so that the curriculum content was integrated into group discussions. Illustrations of primary concepts were presented via role-plays and videotapes. At the end of each meeting, a home practice assignment was given that related to the topics covered during the session. The purpose of these assignments was to assist parents in specific ways to implement the behavioral procedures reviewed in the group meeting. In addition, foster parents were telephoned each week by either the facilitator or co-facilitator to trouble-shoot problems the foster parent was having in implementing the assignment and to administer the PDR. If foster parents missed a parent training session, the material from the missed session was delivered during a home visit at a time convenient for the foster parent.

Caregivers in the control condition received standard services and participated in regular and required in-service trainings as offered by the social service agency. For yearly licensure, foster parents were required to participate in 8 hours of in-service parent training and 8 hours of participation in a foster parent support group. In addition, foster parents may take a variety of different types of training offered for free at local community colleges. Twenty-three percent of the control group participated in

some type of parent or child development training during the 4 months between the baseline and termination interviews.

Sample characteristics. Sixty-six percent of the 700 participating caregivers were nonrelative, while the remaining 34% were some type of relative caregiver (e.g., grandparent, aunt). The mean age of all caregivers at baseline was 48.6 years, with ages ranging from 19 to 81. The ethnic backgrounds of the caregivers were extremely diverse, which can be seen in Table 1: Latino (37.4%), Caucasian (27.4%), African-American (25.7%), Asian/Pacific Islander (2.9%), Native American (1.1%), and mixed ethnic background (5.5%). Sixty percent of the caregivers spoke only English, 8% spoke only Spanish, and 32% spoke both English and Spanish. Table 2 contains the demographic information of the participating caregivers.

The mean age of the target children in this investigation was 8.8 years, with ages ranging from 5 to 12. Fifty-two percent of the sample was female, while 48% was male. The ethnic backgrounds of the target children were also diverse, as seen in Table 3: Latino (32.6%), Caucasian (22.4%), African-American (21.1%), Native American (1.1%), Asian/Pacific Islander (.9%), and mixed ethnic background (21.2%). Sixty-nine percent of the children spoke only English, 2% spoke only Spanish, and 29% spoke both English and Spanish. Table 4 contains the demographic information of the target children.

Given that it is a goal of child welfare to place foster children with substitute caregivers of the same ethnic background, the degree to which an ethnic match occurred was examined. A caregiver-child dyad was considered a match if their ethnic

Table 1: Foster Parent Ethnicity

Ethnicity	Frequency	Percent
Latino	262	37.4
Caucasian	192	27.4
African American	180	25.7
Asian	20	2.9
Caucasian/Latino	9	1.3
Native American	8	1.1
African American/ Native American	7	1.0
Caucasian/Native American	5	.7
Caucasian/African American	4	.6
Caucasian/Asian	4	.6
African American/Latino	2	.3
African American/Asian	2	.3
Native American/Latino	2	.3
Middle Eastern	2	.3
Middle Eastern/Latino	1	.1

Table 2: Demographic Information of Participating Caregivers

Variable	Percent
Age at baseline	48.6 (11.8)
Gender	
Female	95
Male	5
Primary language spoken	
English only	60
Spanish only	8
English and Spanish	32
Employment	
Currently employed	50
Number of hours per week	17.0
Education level	
High school/GED or less	40
Some college	46
Vocational or technical degree	1
Bachelor's degree	7
Graduate degree	6
Household income	
Less than $35,000	35
$35,000-$64,999	34
$65,000-$94,999	13
Over $95,000	5
Refused/Don't know	13

Table 3: Target Child Ethnicity

Ethnicity	Frequency	Percent
Latino	228	32.6
Caucasian	157	22.4
African American	148	21.1
Caucasian/Latino	49	7.0
Caucasian/African American	27	3.9
African American/Latino	22	3.1
Latino/Asian	12	1.7
African American/Asian	10	1.4
Caucasian/Native American	10	1.4
Caucasian/Asian	9	1.3
Native American	8	1.1
Asian	6	.9
African American/Native American	6	.9
Native American/Latino	2	.3
Native American/Asian	1	.1
Middle Eastern	1	.1
Missing/Do not know	4	.6

Table 4: Demographic Information of Target Children

Variable	Percent
Age at baseline	8.8 (2.2)
Number of days in home	180 (126)
Gender	
Female	52
Male	48
Primary language spoken	
English only	69
Spanish only	2
English and Spanish	29

backgrounds overlapped to any degree (e.g., a Caucasian parent and Caucasian/Latino biracial child would be considered an ethnic match, whereas a Caucasian parent and a Latino child would be considered an ethnic mismatch). Of the 700 participants, 80.6% were ethnically matched and 18.8% were considered mismatched, which can be seen in Table 5. Both caregiver and child ethnicity were reported by the foster parents.

Table 5: Ethnic Match Statistics

Variable	Frequency	Percent
Ethnic match	564	80.6
Ethnic mismatch	132	18.8
Missing/Not reported	4	.6

Assessment

All measures were collected by telephone, record searches, or mail by an assessment team that was blind to the group assignment of the participants. This method of assessing the outcomes was designed to maintain separation between the

intervention staff and the assessment staff; and also minimize the burden to foster

parents in both the intervention and control conditions. At baseline, family background

occurred was examined. A caregiver-child dyad was considered a match if their ethnic

backgrounds overlapped to any degree (e.g., a Caucasian parent and Caucasian/Latino

information and assessment of child behavior problems were collected via three phone

interviews. Two to three phone interviews were also used to collect outcome

information at termination of the intervention. All caregivers were paid $25 for the

baseline interviews and $35 for the termination interviews.

During the 2 weeks prior to the start of the study, caregivers completed the

PDR at three separate times to obtain a baseline level of behavior problems. Each of

the three baseline administrations had to be separated by at least 24 hours. Using

approximately one-third of the total Project KEEP sample, Chamberlain et al. (2006)

reported that the average inter-call correlation across the three baseline PDR calls was

.64. They also reported that the internal consistency of the measure was strong, with a

Cronbach's alpha of .84. After the baseline period, foster parents completed the PDR

weekly via the telephone in order to monitor change in child behavior.

Spanish Procedures

Given that approximately one-third of the foster parents in San Diego County

are of Latino decent, it was necessary to offer Spanish-speaking groups to those who

preferred communicating in Spanish. Of the 700 foster parents, 200 (29%) chose to

participate in Spanish-speaking groups. Bilingual group facilitators and co-facilitators

were hired and all materials and assessments were translated into Spanish by

experienced translators. Videotapes of parenting sessions were transcribed into English for supervision purposes. In addition, group facilitators and co-facilitators provided continuous feedback on the cultural appropriateness of the procedures and materials. Several wording changes were made in the curriculum and assessment procedures as a result of this feedback.

A professional Spanish translator was hired by project staff in order to create a conceptually equivalent Spanish version of the PDR. The translator examined the PDR and then translated the 31 items. After this initial translation, the primary Spanish group facilitator of Project KEEP evaluated the translated version for conceptual equivalence. When there were disparities, she conferred with the translator to make any necessary adjustments to the translated measure. This resulted in the 31-item Spanish PDR, which both the translator and group facilitator agreed matched the conceptual meaning of the English PDR.

CHAPTER 3

Results

Overview of Analyses

The broad goal of this investigation was to gain a better understanding of the psychometric properties and factor structure of the PDR. Thus, in order to gather the most information regarding the psychometric properties of the PDR, the factor structure at the initial baseline assessment of the PDR (referred to as the *Time 1 PDR*), and also the factor structure averaged across the three baseline assessments (referred to as the *Time Composite PDR*) were both examined. This more comprehensive approach to the analyses was intended to help understand how the factor structure may vary according to the administration schedule (i.e., administered once versus administered repeatedly), and also to provide insight into how often the PDR should be administered to obtain a reliable and stable estimate of child behavior problems.

To begin, preliminary analyses were conducted in order to examine the descriptive statistics (i.e., mean, standard deviation, frequency, skewness, kurtosis) of the 31 PDR items. Next, the analyses of the core study aims proceeded in the following sequence: exploratory factor analysis; multiple group analysis of language; multiple group analysis of gender; exploratory tests of invariance examining kinship status, ethnic match status, caregiver ethnicity, caregiver age, and the number of days the child was in the home at baseline; latent class analysis; and logistic regression analyses of foster child variables (i.e., age, gender, ethnicity) related to class membership. This sequence of analyses was followed for both the Time 1 PDR (i.e.,

single initial administration of the PDR) and the Time Composite PDR (i.e., average

of the three baseline administrations of PDR). The findings from the Time 1 PDR will

be presented in their entirety first, followed by the results from the Time Composite

PDR. MPlus version 3.0 (Muthen & Muthen, 1998) was utilized for all of the study

aims, with the exception of the logistic regression analyses, which were conducted

using SPSS version 11.0.

Descriptive Analyses

Although a 31-item PDR was utilized in Project KEEP (see Figure 1), two PDR

items, "soil" and "drugs/alcohol," were not included in any of the analyses because

they were not endorsed by any participant in the present study during the 2-week

baseline period. Correlational analyses conducted on the three separate baseline

administrations of the PDR revealed significant and positive correlations: Time 1,

Time 2: $r = .61, p < .01$; Time 2, Time 3: $r = .65, p < .01$; Time 1, Time 3: $r = .62$,

$p < .01$.

Table 6 contains the means and standard deviations of the 29 PDR items at the

three baseline time points. Each PDR item is a binary variable, with "1" indicating a

"yes" response, and "2" indicating a "no" response. Thus, higher frequency items

such as "competitive" and "complain" had means closer to "1," while low-frequency

items such as "run away" and "truant" had means closer to "2." In general, the low-

frequency items tended to have non-normal distributions, as evidenced by the

skewness and kurtosis statistics listed in Table 7.

Table 6: Descriptive Statistics for PDR Items[a]

	Time 1		Time 2		Time 3	
Item	Mean	*SD*	Mean	*SD*	Mean	*SD*
Argue	1.57	.50	1.64	.48	1.61	.49
Back-talk	1.68	.47	1.74	.44	1.72	.45
Wet	1.90	.31	1.91	.28	1.89	.32
Competitive	1.52	.50	1.59	.49	1.58	.49
Complain	1.51	.50	1.55	.50	1.51	.50
Defiant	1.75	.43	1.78	.42	1.76	.43
Destructive, vandalize	1.93	.25	1.94	.23	1.93	.25
Fight	1.79	.41	1.82	.38	1.79	.41
Irritable	1.60	.49	1.66	.47	1.63	.48
Lie	1.69	.46	1.78	.42	1.75	.43
Negative	1.68	.47	1.73	.44	1.69	.46
Boisterous, rowdy	1.67	.47	1.70	.46	1.69	.46
Not mind	1.71	.45	1.73	.44	1.70	.46
Stay out late	1.99	.12	1.98	.13	1.99	.09
Skip meals	1.95	.22	1.94	.24	1.94	.24
Run away	2.00	.04	2.00	.04	2.00	.00
Swear, bad language	1.94	.24	1.95	.22	1.93	.26
Tease, provoke	1.67	.47	1.73	.45	1.68	.47

(*table continues*)

Table 6: Continued

Item	Time 1		Time 2		Time 3	
	Mean	*SD*	Mean	*SD*	Mean	*SD*
Depressed, sad	1.81	.39	1.82	.38	1.78	.42
Sluggish	1.84	.37	1.83	.37	1.83	.38
Jealous	1.70	.46	1.79	.41	1.76	.43
Truant	1.99	.09	1.99	.12	1.99	.11
Steal	1.97	.18	1.98	.15	1.97	.17
Nervous, jittery	1.88	.32	1.88	.33	1.88	.32
Short attention span	1.67	.47	1.73	.45	1.67	.47
Daydream	1.74	.44	1.79	.41	1.77	.42
Irresponsible	1.74	.44	1.79	.41	1.72	.45
School problem	1.90	.30	1.92	.27	1.92	.28
Sexual activity	1.98	.13	1.98	.14	1.98	.13

[a]PDR items are binary variables, with 1 indicating a "yes" response and 2 indicating a "no" response.

Table 7: Skewness and Kurtosis Statistics

	Time 1		Time 2		Time 3	
Item	Skewness	Kurtosis	Skewness	Kurtosis	Skewness	Kurtosis
Argue	-.29	-1.92	-.58	-1.67	-.44	-1.81
Back-talk	.76	-1.43	-1.10	-.79	-.96	-1.09
Wet	-2.59	4.71	-2.92	6.53	-2.43	3.91
Competitive	-.10	-2.00	-.36	-1.88	-.33	-1.90
Complain	-.06	-2.00	-.21	-1.96	-.03	-2.01
Defiant	-1.18	-.62	-1.32	-.25	-1.19	-.59
Destructive, vandalize	-3.50	10.28	-3.81	12.54	-3.46	9.98
Fight	-1.42	.02	-1.71	.93	-1.43	.04
Irritable	-.41	-1.84	-.67	-1.55	-.52	-1.74
Lie	-.84	-1.30	-1.32	-.26	-1.15	-.69
Negative	-.79	-1.38	-1.06	-.88	-.83	-1.31
Boisterous, rowdy	-.70	-1.51	-.89	-1.22	-.80	-1.36
Not mind	-.93	-1.14	-1.06	-.88	-.90	-1.20
Stay out late	-8.19	65.19	-7.78	58.63	-11.58	132.39
Skip meals	-4.06	14.53	-3.75	12.11	-3.71	11.79
Run away	-26.40	697.00	-26.36	695.00	Not endorsed	
Swear, bad language	-3.76	12.16	-4.12	15.03	-3.37	9.36
Tease, provoke	-.71	-1.50	-1.04	-.93	-.77	-1.41
Depressed, sad	-1.58	.48	-1.68	.84	-1.36	-.16
Sluggish	-1.88	1.54	-1.78	1.16	-1.76	1.10

(*table continues*)

Table 7: Continued

Item	Time 1		Time 2		Time 3	
	Skewness	Kurtosis	Skewness	Kurtosis	Skewness	Kurtosis
Jealous	-.90	-1.20	-1.40	-.03	-1.19	-.59
Truant	-10.66	111.99	-8.17	64.99	-8.55	71.32
Steal	-5.37	26.92	-6.60	41.66	-5.75	31.16
Nervous, jittery	-2.40	3.77	-2.33	3.45	-2.41	3.80
Short attention span	-.73	-1.48	-1.02	-.96	-.72	-1.49
Daydream	-1.12	-.74	-1.42	.03	-1.30	-.30
Irresponsible	-1.09	-.81	-1.40	-.05	-.98	-1.04
School problem	-2.66	5.11	-3.13	7.79	-3.02	7.12
Sexual activity	-7.78	58.72	-7.12	48.84	-7.36	52.24

Time 1 Parent Daily Report (Single Baseline Assessment)

Exploratory Factor Analysis (EFA). The initial EFA conducted on the 29-item PDR revealed a two-factor solution (Factor 1: eigenvalue = 8.73, variance accounted for = .30; Factor 2: eigenvalue = 2.85, variance accounted for = .10). A promax rotation was used in order to aid interpretation while allowing correlations among factors. Table 8 contains the factor loadings of the 29 PDR items across the three baseline time points.

In determining what items should be included in the Time 1 PDR, the following two criteria were used: (a) items were retained if the primary loading exceeded .45 in magnitude, and (b) the secondary loading did not exceed .25 (Comrey & Lee, 1992). The following 6 items were excluded because they violated these criteria: wet, skip

Table 8: Baseline Factor Loadings for 29-Item PDR

Item	Time 1		Time 2		Time 3	
	Factor 1	Factor 2	Factor 1	Factor 2	Factor 1	Factor 2
Argue	.820	-.134	.878	-.170	.884	-.014
Back-talk	.751	-.028	.835	-.232	.637	.115
Wet	.131	.313	.041	.406	.119	.278
Competitive	.448	-.059	.624	-.119	.730	-.147
Complain	.804	-.253	.707	-.094	.554	.253
Defiant	.827	-.050	.847	-.040	.667	.256
Destructive, vandalize	.495	.251	.544	.279	.495	.243
Fight	.530	.101	.642	.056	.775	-.156
Irritable	.838	-.180	.765	-.051	.557	.259
Lie	.479	.232	.488	.242	.280	.400
Negative	.736	-.018	.760	-.039	.549	.299
Boisterous, rowdy	.517	.024	.674	-.027	.671	.031
Not mind	.679	-.087	.652	.086	.530	.345
Stay out late	-.242	.692	.359	.339	.310	.196
Skip meals	.036	.306	.186	.309	.163	.089
Run away	-.266	1.170	-.165	1.129	n/a	n/a
Swear, bad language	.243	.393	.267	.325	.736	-.132
Tease, provoke	.568	.085	.666	-.064	.770	-.119
Depressed, sad	.360	.017	.292	.178	.084	.427
Sluggish	.262	.127	.033	.247	-.020	.470

(*table continues*)

Table 8: Continued

Item	Time 1		Time 2		Time 3	
	Factor 1	Factor 2	Factor 1	Factor 2	Factor 1	Factor 2
Jealous	.568	-.031	.545	-.015	.638	.073
Truant	-.066	.558	-.426	.800	-.013	-.247
Steal	.200	.509	-.104	.506	.068	.366
Nervous, jittery	.288	.132	.218	.159	-.045	.656
Short attention span	.528	-.079	.585	-.020	-.047	.874
Daydream	.408	-.037	.334	.114	-.116	.602
Irresponsible	.528	.192	.369	.005	.377	.414
School problem	.294	.188	.368	.251	-.168	.638
Sexual activity	.036	.547	.463	.341	.427	.354

meals, depressed/sad, sluggish, nervous/jittery, and school problems, resulting in a total of 23 items being included in the Time 1 PDR.

The EFA conducted on this 23-item PDR yielded a two-factor solution, with factor 1 (eigenvalue = 8.13, variance accounted for = .35) yielding high internal consistency (Cronbach's alpha = .84), and factor 2 (eigenvalue = 1.93, variance accounted for = .08) yielding surprisingly poor internal consistency (Cronbach's alpha = .18). Factor 1, entitled Negative Behaviors, included the following 17 items: argue, back-talk, competitive, complain, defiant, destructive, fight, irritable, lie, negative, boisterous/rowdy, not mind, tease/provoke, jealous, short attention span, daydream, and irresponsible. Factor 2, entitled Low Frequency Behaviors, included the following 6 items: stay out late, run away, swear/bad language, truant, steal, and inappropriate

sexual activity. Given the extremely poor internal consistency of factor 2, and the low frequency of the items comprising the factor, it was excluded from subsequent analyses. The single factor solution depicted in Figure 2 was utilized for all analyses conducted on the Time 1 PDR.

Multiple Group Analysis (MGA): Language. For all multiple group analyses, there were several a-priori criteria established for significance and good model fit. Most researchers suggest that in addition to the chi-square test of model fit, descriptive fit indices should be examined in order to make determinations of model fit (e.g., Hoyle, 2000). In the current study, the following measures were utilized: comparative fit index (CFI; Bentler, 1990), root mean square error of approximation (RMSEA; Steiger, 1990), and standardized root mean square residual (SRMR; see Hu & Bentler, 1999). In general, CFI values greater than .90 and RMSEA and SRMR values less than .08 were considered evidence of good model fit (Hu & Bentler, 1999; Maruyama, 1998). When the CFI, RMSEA, and SRMR were not in agreement, decisions regarding model fit were based on the majority (i.e., two of three indices). For all multiple group analyses, a chi-square difference statistic was calculated to compare statistically nested models ($\Delta\chi^2$; see Maruyama, 1998).

The one-factor model fit well for both the English version, χ^2 (119, $n = 509$) = 368.74, $p < .01$, CFI = .86, RMSEA = .06, SRMR = .06, and the Spanish version, χ^2 (119, $n = 185$) = 251.69, $p < .01$, CFI = .80, RMSEA = .08, SRMR =.07. Although the CFI values did not quite reach the .90 cut-off, the chi-square tests and RMSEA and SRMR values suggested that the model fit well. The factor loadings for the English
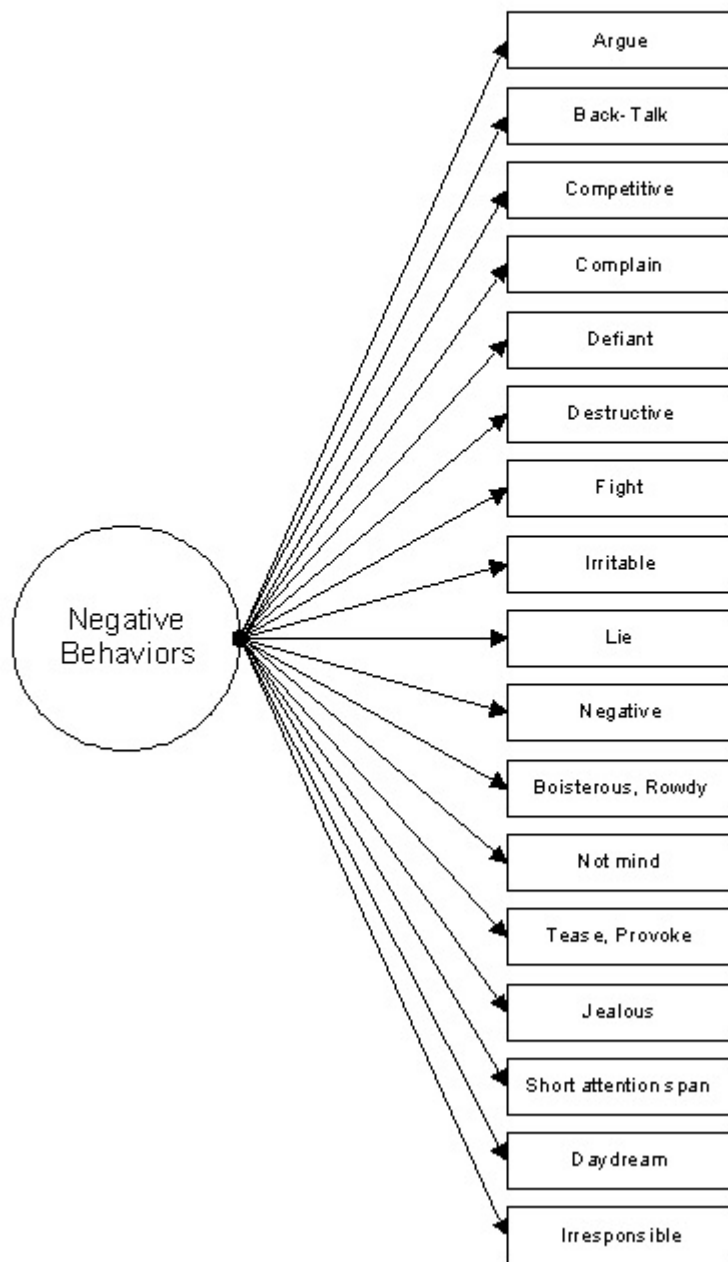
Figure 2: Factor Structure of Time 1 PDR

Figure 2: Factor structure of Time 1 PDR

version were all positive and significant (values ranging from .40 to 1.00), with the

exception of the item "destructive" (factor loading = .25) and "daydream" (factor

loading = .31). The factor loadings for the Spanish version were all positive and

significant (values ranging from .42 to 1.10) with the exception of the item

"destructive" (factor loading = .29). Given that the items with nonsignificant factor

loadings differed for the English and Spanish versions ("destructive" and "daydream"

for English and "destructive" for Spanish), there was not configural invariance across

language and thus tests of metric invariance were not conducted. These findings

indicate that there is not factor equivalence across language for the Time 1 PDR.

Multiple Group Analysis: Gender. The one-factor model fit well for both males,

χ^2 (119, n = 329) = 272.15, p < .01, CFI = .87, RMSEA = .06, SRMR = .06, and

females, χ^2 (119, n = 365) = 294.06, p < .01, CFI = .85, RMSEA = .06, SRMR = .06.

Again, although the CFI values did not quite reach the .90 cut-off, the chi-square tests

and RMSEA and SRMR values collectively suggested that the model fit well. The

factor loadings for males and females were generally positive and significant (values

ranging from .48 to 1.00 for males; from .51 to 1.00 for females), with the exception

of the items "destructive" (factor loading = .25 for both males and females) and

"daydream" (factor loading = .34 for males and .35 for females). Given that the non-

significant factor loadings were consistent across gender, configural invariance was

established.

To test for metric invariance across gender, the factor loadings of the one-factor

model were constrained to be equal across groups. This model also fit well, χ^2 (254,

$n = 694) = 577.34$, $p < .01$, CFI = .86, RMSEA = .06, SRMR = .06. Moreover, this model did not significantly differ from the configural invariance model, $\Delta\chi^2$ ($\Delta df =$ 16) = 11.13, $p > .05$. Since there is no statistical difference between the two models, the more parsimonious model fits better because fewer parameters are estimated (i.e., there are more degrees of freedom). Thus, the metric invariance model was deemed the better fitting model. All of the factor loadings for the metric invariance model were equal for both males and females, suggesting that there is factor equivalence across gender for the Time 1 PDR.

Exploratory tests of invariance. MGA was used in order to examine additional categorical variables of interest to determine if factor loadings were group invariant. Specifically, kinship status (kin caregiver vs. nonkin caregiver), ethnic match status (ethnic match vs. ethnic mismatch), and caregiver ethnicity (Caucasian, Latino, and African American) were examined. Ethnicity comparisons were limited to the three identified groups due to the ethnic composition of the sample, which is represented in Table 1.

The one-factor model of the Time 1 PDR fit well for both kinship caregivers, χ^2 (119, $n = 235$) = 251.48, $p < .01$, CFI = .84, RMSEA = .07, SRMR = .07, and non-kinship caregivers, χ^2 (119, $n = 459$) = 313.56, $p < .01$, CFI = .87, RMSEA = .06, SRMR = .05. The metric invariance model, which constrained the factor loadings between kin and nonkin caregivers to be equal, also fit well, χ^2 (254, $n = 694$) = 579.64, $p < .01$, CFI = .86, RMSEA = .06, SRMR=.06. Moreover, this constrained model did not significantly differ from the configural invariance model, $\Delta\chi^2$ ($\Delta df =$

16) = 14.60, $p > .05$, and thus is deemed the more parsimonious and better fitting

model. All of the factor loadings for the metric invariance model were equal across

kinship status, indicating that there is factor equivalence.

The one-factor model fit well for ethnically matched caregiver-child dyads,

χ^2 (119, $n = 558$) = 354.48, $p < .01$, CFI = .87, RMSEA = .06, SRMR = .05, and

mismatched dyads, χ^2 (119, $n = 459$) = 214.00, $p < .01$, CFI = .80, RMSEA = .08,

SRMR = .08. The constrained metric invariance model also fit well, χ^2 (254, $n = 690$)

= 583.01, $p < .01$, CFI = .86, RMSEA = .06, SRMR = .06. Moreover, this constrained

model did not significantly differ from the configural invariance model, $\Delta\chi^2$ ($\Delta df =$

16) = 14.62, $p > .05$, and thus is deemed the better fitting model. All of the factor

loadings for the metric invariance model were equal for ethnically matched and

mismatched dyads, indicating that there is factor equivalence across ethnic match

status for the Time 1 PDR.

The one-factor model fit well for Caucasian caregivers, χ^2 (119, $n = 190$) =

255.55, $p < .01$, CFI = .78, RMSEA = .08, SRMR = .07, African American caregivers,

χ^2 (119, $n = 178$) = 202.54, $p < .01$, CFI = .85, RMSEA = .06, SRMR = .07, and

Latino caregivers, χ^2 (119, $n = 260$) = 260.63, $p < .01$, CFI = .85, RMSEA = .07,

SRMR = .06. The metric invariance model also fit well, χ^2 (389, $n = 628$) = 747.33,

$p < .01$, CFI = .83, RMSEA = .07, SRMR = .07. Moreover, this constrained model did

not significantly differ from the configural invariance model, $\Delta\chi^2$ ($\Delta df = 32$) = 28.61,

$p > .05$, and thus is deemed the better fitting model. All of the factor loadings for the

metric invariance model were equal across caregiver ethnicity, indicating that there is

factor equivalence for Caucasian, African American, and Latino foster parents. To summarize, factor loadings were equivalent across kinship status, ethnic match status, and foster parent ethnicity for the Time 1 PDR.

Multiple Indicator Multiple Cause (MIMIC; see Maruyama, 1998) models were used in order to determine if the single factor structure of the Time 1 PDR was invariant across continuous variables of interest. Specifically, the age of the foster parent and the number of days the child was in the home at the time of the initial baseline PDR assessment were examined in a single MIMIC model. The fully constrained model fit well, χ^2 (151, $n = 688$) $= 444.76$, $p < .01$, CFI $= .87$, RMSEA $=$ .05, SRMR $= .05$. Although the modification indices revealed no differences according to number of days in home, the indices indicated that the items "complain," "lie," and "irresponsible" differed with foster parent age.

A subsequent MIMIC model (i.e., partially constrained model) freeing the path coefficients identified by the modification indices was tested in order to determine whether the constrained or partially constrained model fit better and also to obtain the values of the path coefficients. The partially constrained model fit well, χ^2 (148, $n =$ 688) $= 428.82$, $p < .01$, CFI $= .88$, RMSEA $= .05$, SRMR $= .05$, and the chi-square difference test indicated that this model, which is depicted in Figure 3, fit better than the fully constrained model, $\Delta\chi^2$ ($\Delta df = 3$) $= 15.94$, $p < .05$. The path coefficients were 2.657, 2.314, and 1.796 for the items "complain," "lie," and "irresponsible," respectively. The positive coefficients indicate that as the age of the foster parents increased, the more likely they were to endorse the three items.
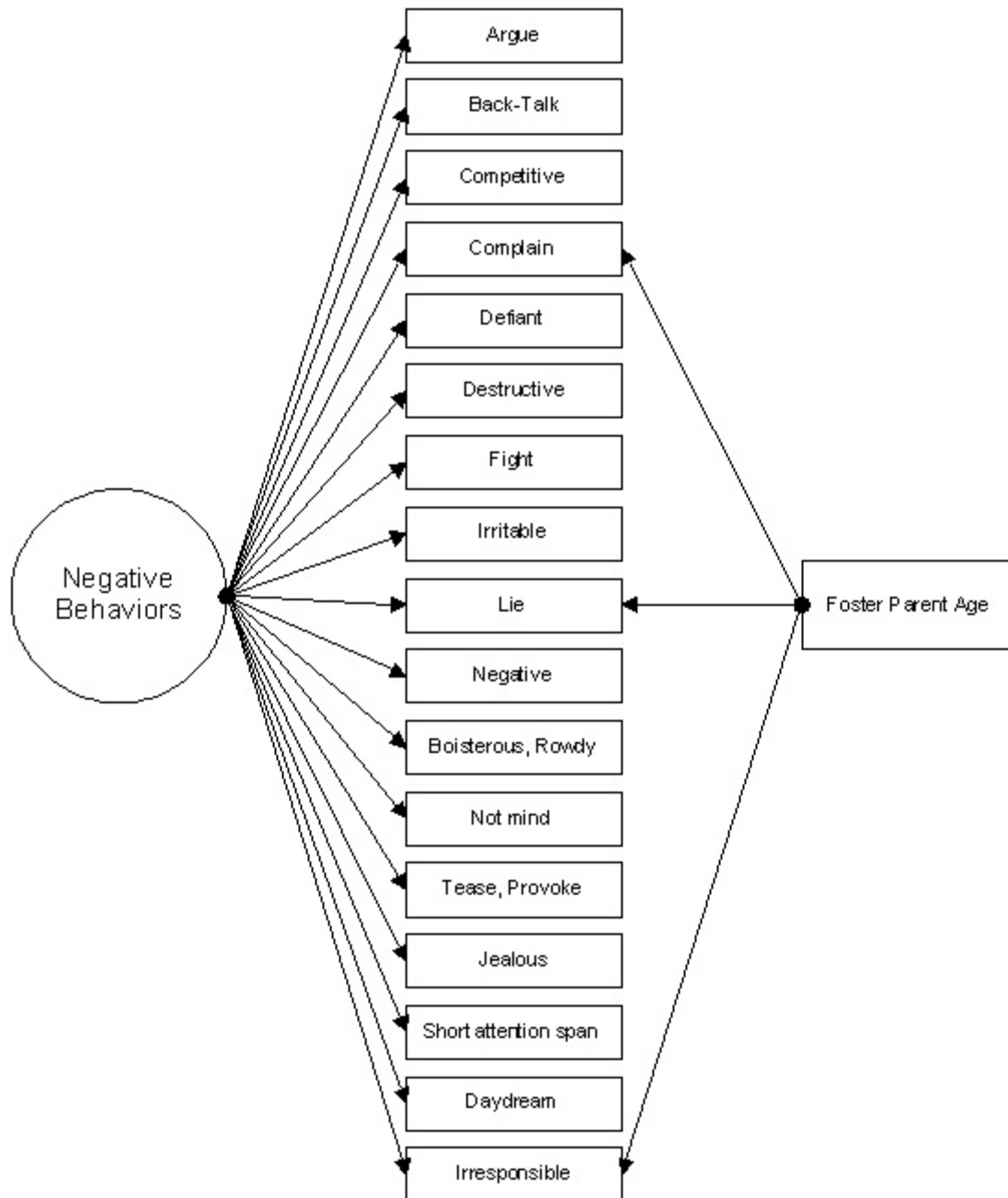
Figure 3: MIMIC Model for Time 1 PDR

Figure 3: MIMIC Model for Time 1 PDR

Latent Class Analysis (LCA). Exploratory LCA was used in order to investigate 1-, 2-, 3- and 4-class solutions for the Time 1 PDR. Classes were added iteratively until the model fit the data well from both a statistical and interpretive perspective. The Akaike Information Criterion (AIC) and sample size adjusted Bayesian Information Criterion (BIC) were used in order to determine the optimal class solution. In general, lower AIC and BIC values (i.e., closer to 0) indicated better model fit. The Lo-Mendell-Rubin Test (LMRT) was also used to compare models and determine if there was a significant difference between solutions that varied by one class.

Table 9 contains the AIC, BIC, and LMRT values for the latent class analyses conducted on the Time 1 PDR. The 2-class solution was better than the 1-class solution as seen by the significance test, LMRT = 1469.59, $p < .01$. The 3-class solution was deemed better than the 2-class solution due to the lower AIC and BIC values, and also the significant LMRT. The 4-class solution did not fit well, as seen by the higher BIC value and also the nonsignificant LMRT. Based on these data, the 3-class solution was deemed the best fitting model for the Time 1 PDR. Class 1 has 280 members (40.3%), Class 2 has 301 members (43.4%), and Class 3 has 113 members (16.3%).

Table 9: LCA Results for Time 1 PDR

Criterion	2-class solution	3-class solution	4-class solution
AIC	10655.25	10458.79	10435.91
BIC	10703.10	10531.25	10532.99
LMRT	1469.59, $p < .01$	230.39, $p < .01$	58.38, $p = .46$

The conditional response probabilities of the 17 items that compose the Time 1

PDR are listed in Table 10. Note that conditional response probabilities sum to 1, and

the probability of a "yes" response for each item is listed in Table 10. Qualitative

analysis of the response probabilities suggests that Class 1 reflects the least severe

group of children (i.e., lowest probability of "yes" responses across items). Class 3

reflects the most severe group of children, with the highest conditional response

probabilities across items (i.e., highest likelihood of "yes" responses with probabilities

close to 1). Finally, children in Class 2 exhibit a moderate level of behavioral

disturbance, as seen by the conditional response probabilities which fall between those

found for Class 1 and Class 3. Thus, classes appear to differ with regards to the level

or severity of behavioral disturbance. Class 1 is deemed the *Low* group, Class 2 is

identified as the *Moderate* group, and Class 3 is considered the *Severe* group.

Variables related to class membership. Multinomial logistic regression analyses

were conducted in order to examine child age, gender, and ethnicity (Caucasian,

African American, and Latino) as variables which may be related to the three classes

identified by LCA. Regression analyses were conducted separately because the

variables were not significantly correlated with one another. Gender was not

significantly associated with class membership, whereas child age and ethnicity was

significantly related to class membership. Specifically, as age increases, children are

1.1 times more likely to be in the Low group as compared to the Severe group, β =

-.133, $p < .01$, odds ratio (OR) = 1.143. The odds ratio can be conceptualized as a

method to compare whether the probability of a certain event is the same for two

Table 10: Three-Class Solution for Time 1 PDR[a]

Item	Class 1: Low (*n* = 280)	Class 2: Moderate (*n* = 301)	Class 3: Severe (*n* = 113)
Argue	.059	.549	.975
Back-talk	.037	.368	.852
Competitive	.270	.571	.700
Complain	.193	.573	.951
Defiant	.000	.239	.828
Destructive	.007	.048	.241
Fight	.037	.235	.543
Irritable	.090	.486	.895
Lie	.112	.359	.618
Negative	.044	.365	.810
Rowdy	.095	.417	.670
Not mind	.081	.319	.681
Tease	.011	.015	.020
Jealous	.004	.000	.000
Short attention span	.010	.057	.161
Daydream	.066	.426	.705
Irresponsible	.073	.371	.620

[a]Conditional response probabilities sum to 1, probability of "yes" response listed.

groups. In addition, as age increases, children are 1.1 times more likely to be in the Low group, as compared to the Moderate group, $\beta = -.111, p < .01, OR = 1.117$. To summarize, as children get older, they are more likely to be in the Low group than either the Severe or Moderate groups.

Ethnicity was significantly associated with class membership such that African American children are 2.3 times more likely to be in the Low group than the Severe group, as compared to Caucasian children, $\beta = -.833, p < .05, OR = 2.299$. A similar pattern emerged when comparing Latino and Caucasian children in that Latino children are 1.3 times more likely to be in the Low group than the Severe group, as compared to Caucasian children, $\beta = -.248, p < .05, OR = 1.282$. In addition, Latino children were found to be 1.3 times more likely to be in the Low group than the Moderate group, relative to Caucasian children, $\beta = -.264, p < .01, OR = 1.302$. When comparing African American and Latino children, ethnicity was not significantly associated with class membership. Taken together, African American children are more likely to be classified into the Low group than the Severe group, and Latino children are more likely to be in the Low group than either the Moderate or Severe groups, relative to Caucasian children.

Summary of Results for Time 1 PDR

The analyses conducted on the Time 1 PDR indicate that the single factor structure of the 17-item PDR is equivalent across gender, but not language. It should be noted, however, that the lack of configural invariance across language was based on only one item, "daydream," loading differently. Specifically, "daydream" loaded

significantly on the Spanish version but not the English version. Thus, the lack of equivalence was not a result of highly discrepant factor structures, but rather a single item loading differently across language. Subsequent multiple group analyses suggested that the single factor structure did not vary across kinship status, ethnic match status, or foster parent ethnicity. While the MIMIC model indicated that the factor structure of the Time 1 PDR did not vary with the number of days the child was in the home, it did suggest that there is a relation between caregiver age and three items of the Time 1 PDR: complain, lie, and irresponsible. Latent class analyses suggested the presence of three distinct classes of children, identified by the severity of their behavioral disturbance (i.e., Low, Moderate, Severe). Logistic regression analyses indicated that child age and ethnicity were significantly associated with class membership.

Time Composite Parent Daily Report (Average of Three Baseline Assessments)

     Exploratory factor analysis. In order to examine the Time Composite PDR, EFA was conducted in an iterative fashion in order to find the set of items that loaded consistently across time. Since a stable and pure measure of child behavior problems was desired, items with inconsistent loadings and significant cross-loadings were excluded. More specifically, items were dropped individually according to the following three criteria: (a) item did not load significantly on any factor, (b) item loaded differently across the three baseline administrations of the PDR, and (c) item loaded significantly on more than one factor (Comry & Lee, 1992; Tabachnik & Fidell, 2001). This iterative process eventually yielded a 10-item PDR (i.e., Time

Composite PDR) that was consistent across all three baseline administrations. The

factor loadings of these 10 items are listed in Table 11.

Table 11: Factor Loadings for Time Composite PDR

Item	Time 1		Time 2		Time 3	
	Factor 1	Factor 2	Factor 1	Factor 2	Factor 1	Factor 2
Argue	.602	.263	.519	.403	.628	.327
Back-talk	.771	.048	.750	.077	.852	-.020
Competitive	-.125	.639	-.111	.801	-.075	.825
Defiant	.908	-.038	.959	-.019	.982	-.083
Fight	.121	.514	.184	.593	.073	.652
Irritable	.571	.239	.516	.335	.623	.166
Negative	.500	.329	.564	.198	.588	.213
Not mind	.425	.214	.782	-.054	.711	.052
Tease, provoke	.018	.703	.032	.700	-.057	.849
Jealous	.015	.618	-.049	.642	.229	.536

EFA, again using a promax rotation, revealed a two-factor solution with both

factors yielding high internal consistency values (Cronbach's alpha for Factor 1 = .89;

Cronbach's alpha for Factor 2 = .81). Factor 1 (eigenvalue = 5.57, variance accounted

for = .56), entitled Oppositional Behaviors, contained the following six items: argue,

back-talk, defiant, irritable, negative, and not mind. Factor 2 (eigenvalue = 1.05,

variance accounted for = .11) was entitled Social Conflict, and included the following

four items: competitive, fight, tease/provoke, and jealous. This two-factor model of the

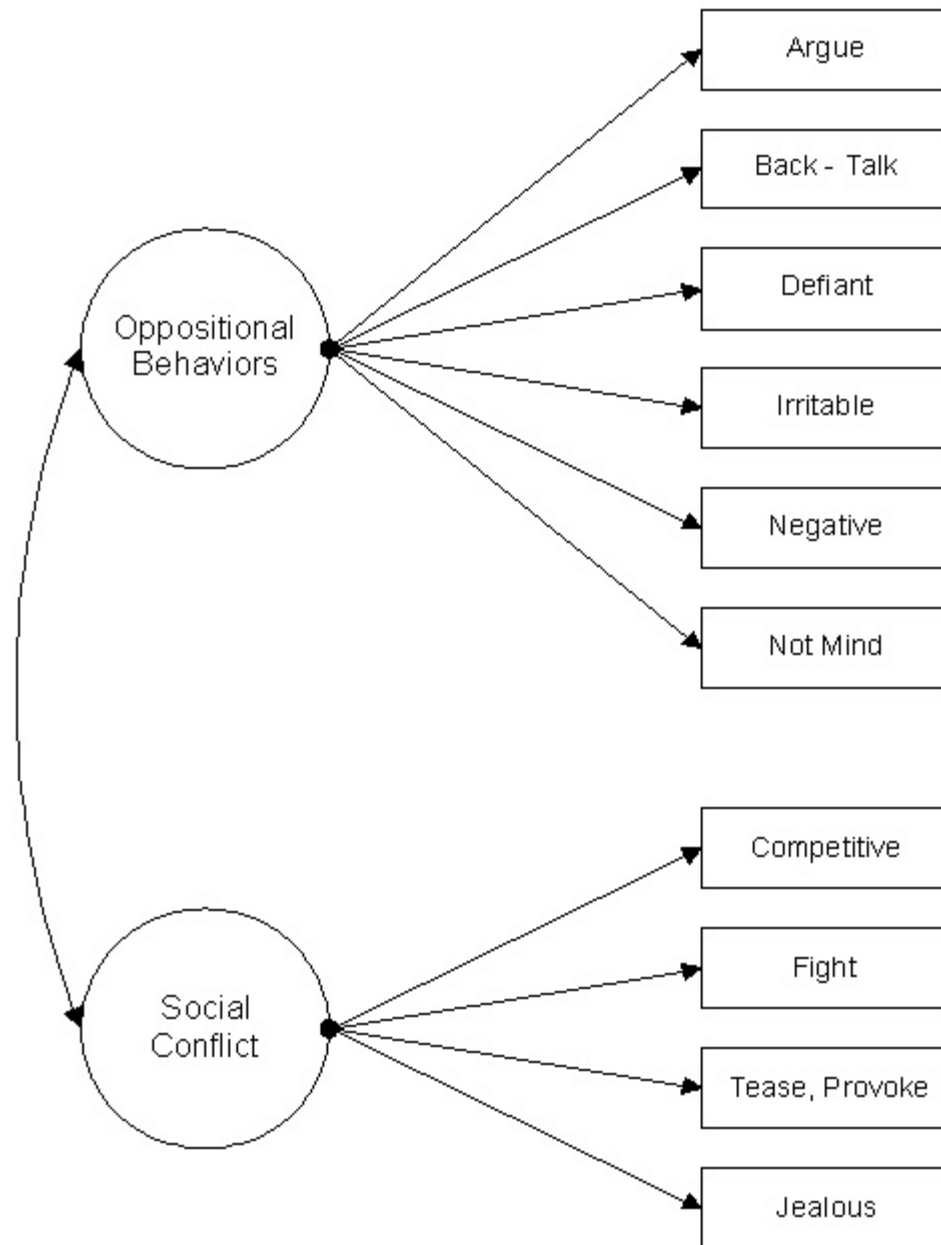Time Composite PDR is depicted in Figure 4.

Figure 4: Factor structure of time composite PDR

Multiple group analysis: Language. Using the same criteria described above with the Time 1 PDR, the two-factor model fit well for both the English version, χ^2 (34, *n* = 495) = 182.19, *p* < .01, CFI = .92, RMSEA = .09, SRMR = .05, and the Spanish version, χ^2 (34, *n* = 180) = 91.70, *p* < .01, CFI = .92, RMSEA = .10, SRMR = .05. All of the factor loadings for both language groups were large, positive, and significant (values ranging from .79 to 1.22 for the English version; from .75 to 1.00 for the Spanish version). The interfactor correlation was not significant for either language group (.043 for English and .059 for Spanish). Thus, configural invariance was established.

To test for metric invariance, the factor loadings of the two-factor model were constrained to be equal across language groups. This model also fit well, χ^2 (76, *n* = 675) = 299.70, *p* < .01, CFI = .92, RMSEA = .09, SRMR = .07. However, this model was significantly different from the configural invariance model, $\Delta\chi^2$ (Δdf = 8) = 25.81, *p* < .01, suggesting that the unconstrained model fits better for each language group. Model modification indices suggested that the items "argue," "defiant," and "irritable," may not load consistently on factor 1. The modification indices also indicated that the items "competitive," "fight," "tease," and "jealous" may not load consistently on factor 2.

A revised metric invariance model was then tested which did not constrain the loadings identified by the modification indices (i.e., factor loadings were freed). This analysis found that three items were statistically different across language. Specifically, "irritable" (.86 for English and 1.11 for Spanish), "fight" (.77 for English

and 1.01 for Spanish), and "tease" (1.15 for English and .95 for Spanish) had different

factor loadings for the English and Spanish versions of the Time Composite PDR.

These factor loadings suggest that the items "irritable" and "fight" are better indicators

for the Spanish language version, whereas the item "tease" is a better indicator for the

English version of the Time Composite PDR. The partially constrained model fit well,

χ^2 (73, n = 675) = 284.50, p < .01, CFI = .92, RMSEA = .09, SRMR = .06 and did

not significantly differ from the configural invariance model, $\Delta\chi^2$ (Δdf = 5) = 10.61,

p > .05. Thus, the partially constrained model was deemed the better fitting model,

suggesting that there is partial invariance across language for the Time Composite

PDR.

   Multiple group analysis: Gender. The two-factor model fit well for both males,

χ^2 (34, n = 321) = 108.46, p < .01, CFI = .94, RMSEA = .08, SRMR = .05, and

females, χ^2 (34, n = 354) = 159.76, p < .01, CFI = 90, RMSEA = .10, SRMR = .05.

All of the factor loadings for both gender groups were large, positive, and significant

(values ranging from .78 to 1.00 for males; from .74 to 1.00 for females). The inter-

factor correlation was not significant for either gender (.049 for males and .045 for

females), establishing configural invariance for the two-factor model.

   To test for metric invariance, the factor loadings of the two-factor model were

constrained to be equal across males and females. This model also fit well, χ^2 (76, n =

675) = 271.52, p < .01, CFI = .92, RMSEA = .09, SRMR = .05. Moreover, this model

did not significantly differ from the configural invariance model, $\Delta\chi^2$ (Δdf = 8) = 3.30,

p > .05, and thus was deemed the more parsimonious and better fitting model. All of

the factor loadings for the metric invariance model were equal across gender groups, suggesting that there is factor equivalence across gender for the Time Composite PDR.

Exploratory tests of invariance. MGA was again used in order to examine the additional categorical variables of interest (i.e., kinship status, ethnic match status, caregiver ethnicity) to determine if the factor loadings are group invariant. The two-factor model fit well for both kinship caregivers, χ^2 (34, $n = 229$) = 126.93, $p < .01$, CFI = .90, RMSEA = .10, SRMR = .06, and nonkinship caregivers, χ^2 (34, $n = 446$) = 153.42, $p < .01$, CFI = .93, RMSEA = .09, SRMR = .05. The metric invariance model, which constrained the factor loadings between kin and nonkin caregivers to be equal, also fit well, χ^2 (76, $n = 675$) = 288.46, $p < .01$, CFI = .92, RMSEA = .09, SRMR=.06. Moreover, this constrained model did not significantly differ from the configural invariance model, $\Delta\chi^2$ ($\Delta df = 8$) = 8.11, $p > .05$, and thus is deemed the more parsimonious and better fitting model. All of the factor loadings for the metric invariance model were equal across kinship status, indicating that there is factor equivalence.

The two-factor model fit well for both ethnically matched caregiver-child dyads, χ^2 (34, $n = 543$) = 190.29, $p < .01$, CFI = .93, RMSEA = .09, SRMR = .05, and ethnically mismatched dyads, χ^2 (34, $n = 128$) = 73.84, $p < .01$, CFI = .90, RMSEA = .10, SRMR = .06. The metric invariance model also fit well, χ^2 (76, $n = 671$) = 271.27, $p < .01$, CFI = .92, RMSEA = .09, SRMR = .05. Moreover, this constrained model did not significantly differ from the configural invariance model, $\Delta\chi^2$ ($\Delta df = 8$) = 7.14, $p > .05$, and thus is deemed the better fitting model. All of the factor loadings for the

metric invariance model were equal for ethnically matched and mismatched dyads, indicating that there is factor equivalence across ethnic match status for the Time 1 PDR.

The two-factor model fit well for Caucasian caregivers, χ^2 (34, $n = 185$) = 96.291, $p < .01$, CFI = .91, RMSEA = .10, SRMR = .06, African American caregivers, χ^2 (34, $n = 172$) = 89.98, $p < .01$, CFI = .92, RMSEA = .09, SRMR = .06, and Latino caregivers, χ^2 (34, $n = 252$) = 128.19, $p < .01$, CFI = .91, RMSEA = .10, SRMR = .05. The metric invariance model constraining the factor loadings to be equal across foster parent ethnicity also fit well, χ^2 (118, $n = 609$) = 334.25, $p < .01$, CFI = .91, RMSEA = .09, SRMR = .08. Moreover, this constrained model did not significantly differ from the configural invariance model, $\Delta\chi^2$ ($\Delta df = 16$) = 19.79, $p > .05$, and thus is deemed the better fitting model. All of the factor loadings for the metric invariance model were equal across caregiver ethnicity, indicating that there is factor equivalence for Caucasian, African American, and Latino foster parents. Taken together, these findings establish metric invariance across kinship status, ethnic match status, and foster parent ethnicity for the Time Composite PDR.

MIMIC models were used in order to determine if the two-factor structure of the Time Composite PDR was invariant across continuous variables of interest. Specifically, the age of the foster parent and the number of days the child was in the home at the time of the initial baseline PDR assessment were examined in a single MIMIC model. The constrained model fit well, χ^2 (50, $n = 659$) = 257.422, $p < .01$, CFI = .92, RMSEA = .08, SRMR = .04. However, the modification indices suggested

that the items "competitive" and "not mind" varied according to the number of days

the child was in the home, and the item "back-talk" differed with foster parent age.

A subsequent MIMIC model (i.e., partially constrained model) freeing the path

coefficients identified by the modification indices was tested in order to determine

whether the constrained or partially constrained model fit better and also to obtain the

values of the path coefficients. The partially constrained model fit well, χ^2 (47, $n =$

659) = 244.73, $p < .01$, CFI = .92, RMSEA = .08, SRMR = .04, and the chi-square

difference test indicated that this model, which is depicted in Figure 5, fit better than

the fully constrained model, $\Delta\chi^2$ ($\Delta df = 3$) = 12.69, $p < .05$. The path coefficients were

-2.107, -2.072, and 1.928 for the items "competitive," "not mind," and "back-talk,"

respectively. For those items with negative coefficients (i.e, competitive, not mind), as

the number of days the child was in the home increased, the less likely the foster

parents were to endorse the items. For the item "back-talk," which has a positive

coefficient, as the age of the foster parents increased, the more likely they were to

endorse the item.

Latent Profile Analysis (LPA). Since the Time Composite PDR represents an

average of PDR items across the three baseline time points and thus is not composed

of binary items, exploratory LPA rather than LCA was used in order to examine 1-, 2-,

3-, and 4-class solutions. The same criteria (i.e., AIC, BIC, LMRT) employed with the

LCA of the Time 1 PDR were utilized. Table 12 contains the AIC, BIC, and LMRT

values for the latent profile analyses. The 2-class solution was better than the 1-class

solution as seen by the significance test, LMRT = 1955.60, $p < .01$. The 3-class

Figure 5: MIMIC Model for Time Composite PDR
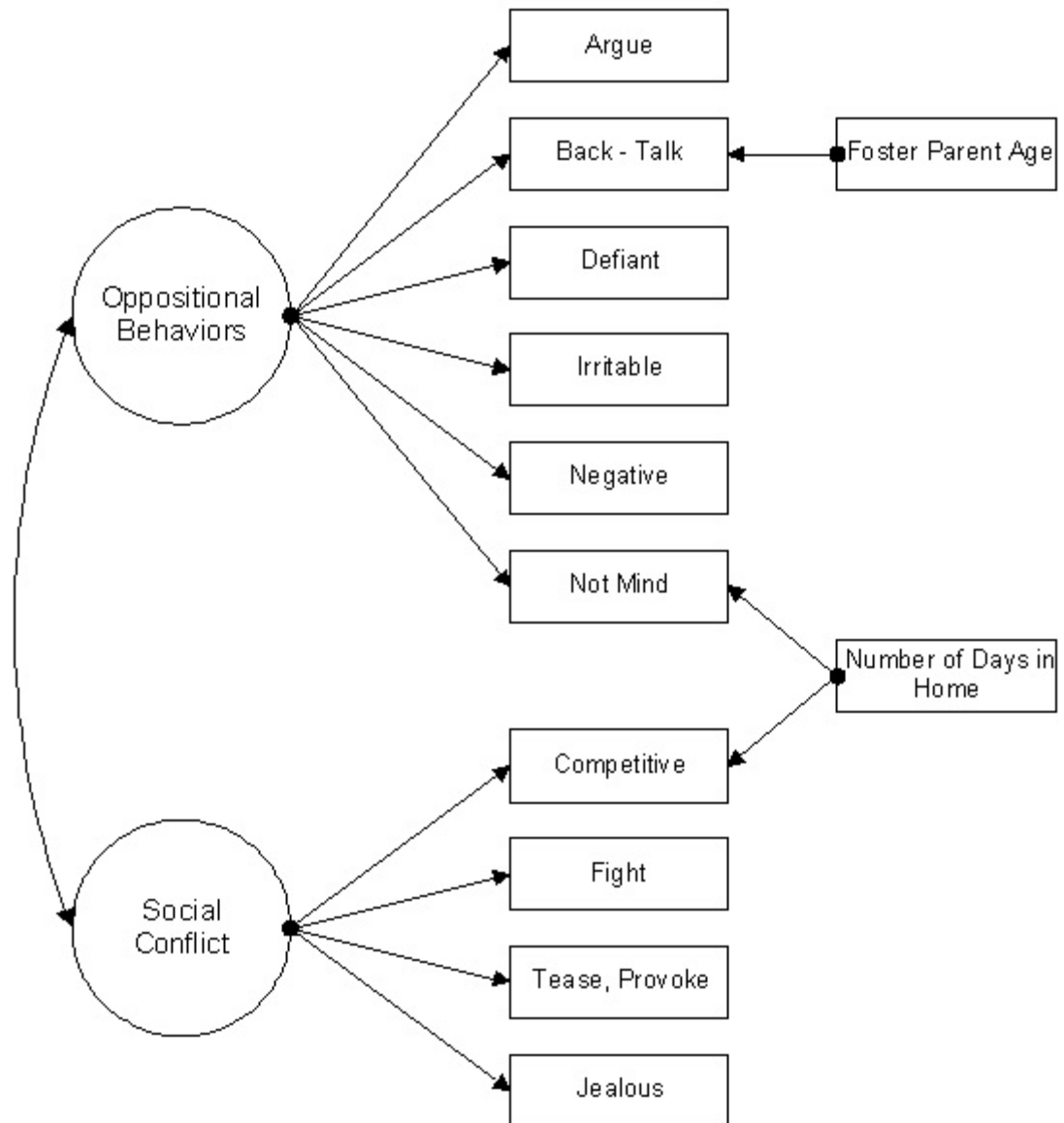
Figure 5: MIMIC Model for Time Composite PDR

solution was deemed better than the 2-class solution due to the lower AIC and BIC

values, and also the significant LMRT. The 4-class solution could not be executed

because the model did not converge, indicating that it poorly fit the data. Based on

these results, the 3-class solution was deemed the best fitting model for the Time

Composite PDR. Class 1 has 385 members (57.0%), Class 2 has 178 members

(26.4%), and Class 3 has 112 members (16.6%).

Table 12: Results for Time Composite PDR

Criterion	2-class solution	3-class solution	4-class solution[a]
AIC	2891.76	2456.17	n/a
BIC	2933.29	2512.53	n/a
LMRT	1955.60, $p < .01$	451.29, $p < .01$	n/a

[a]Solution did not converge.

The means of the 10 items that compose the Time Composite PDR are listed in

Table 13. Note that "yes" responses were coded as a "1" and "no" responses were

coded as a "2." Qualitative analysis of the means suggest that Class 1 reflects the least

severe group of children, with the means closest to "2" (i.e., items were not frequently

endorsed). Class 3 reflects the most severe group of children (i.e., items were endorsed

frequently as seen by the means being closest to "1"). Finally, the children in Class 2

exhibit a moderate level of behavioral disturbance, as seen by the means which fall

between those found for Class 1 and Class 3. Thus, as in the case of the Time 1 PDR,

classes appear to differ with regards to the level or severity of behavioral disturbance.

Class 1 is deemed the *Low* group, Class 2 is identified as the *Moderate* group, and

Class 3 is considered the *Severe* group.

Table 13: Three-Class Solution for Time Composite[a]

Item	Class 1: Low ($n = 385$)	Class 2: Moderate ($n = 178$)	Class 3: Severe ($n = 112$)
Argue	1.871	1.346	1.161
Back-talk	1.908	1.589	1.279
Competitive	1.713	1.415	1.315
Defiant	1.927	1.771	1.215
Fight	1.922	1.735	1.523
Irritable	1.814	1.500	1.217
Negative	1.873	1.616	1.300
Not mind	1.864	1.698	1.266
Tease	1.864	1.536	1.390
Jealous	1.875	1.676	1.453

[a]Means closer to "1" indicate higher probability of "yes" response.

Variables related to class membership. Multinomial logistic regression analyses

were conducted to examine the relations between child age, gender, and ethnicity and

class membership. Regression analyses were conducted separately since the variables

were not significantly correlated with one another. The regression analyses indicated

that child age, gender, and ethnicity were all significantly associated with class

membership. Specifically, as age increases, children are 1.1 times more likely to be in

the Moderate group, as compared to the Low group, $\beta = .103$, $p < .01$, $OR = 1.108$. In

addition, as age increases, the children are 1.2 times more likely to be in the Low

group as compared to the Severe group, $\beta = .115$, $p < .01$, $OR = 1.168$. In summary, as children get older, they are more likely to be in the Moderate or Low groups than the Severe group. Gender was also found to be significantly associated with class membership, $\beta = .486$, $p < .05$, $OR = 1.626$, such that females are 1.6 times more likely to be classified into the Low group rather than the Severe group, relative to males.

Ethnicity was significantly associated with class membership such that African American children are 2.1 times more likely to be in the Low group than the Severe group, as compared to Caucasian children, $\beta = .760$, $p < .05$, $OR = 2.137$. A similar pattern emerged when comparing Latino and Caucasian children. Specifically, Latino children are 1.3 times more likely to be in the Low group than the Severe group, as compared to Caucasian children, $\beta = .282$, $p < .01$, $OR = 1.325$. When comparing African American and Latino children, ethnicity was not significantly associated with class membership. Taken together, these findings suggest that African American and Latino children are more likely to be identified as Low than Severe, relative to Caucasian children.

Summary of Results for Time Composite PDR

The results from the Time Composite PDR indicate that the two-factor structure of the 10-item PDR is equivalent across gender, and partially equivalent across language. Although the items "irritable," "fight," and "tease" were statistically different across language, the largest difference in their factor loadings was .25. Thus, the Time Composite PDR is deemed to have *practical* equivalence across language,

despite the lack of *statistical* equivalence. As in the case of the Time 1 PDR,

subsequent multiple group analyses suggested that the factor structure did not vary

across kinship status, ethnic match status, or foster parent ethnicity. The MIMIC

model indicated that the factor loadings of the Time Composite PDR varied according

to both the number of days the child was in the home (for the items "competitive" and

"not mind") and foster parent age (for the item "back-talk"). Latent profile analyses

indicated the presence of three classes of children that, as in the case of the Time 1

PDR, appear to differ with regards to the level or severity of behavioral problems.

Logistic regression analyses indicated that child age, gender, and ethnicity were all

significantly associated with class membership.

CHAPTER 4

Discussion

The central goal of the present investigation was to examine the psychometric properties of the PDR using a diverse child welfare sample, with a particular emphasis on its factor structure. This goal was attained by examining the factor structure of the PDR (Aim 1) and also comparing it across language (Aim 2) and gender (Aim 3). The final study aim was to determine whether there were distinct classes or subtypes of youth based on the 31 items of the PDR (Aim 4). The analysis of both the Time 1 PDR (i.e., initial baseline administration of the PDR) and Time Composite PDR (i.e., average of three baseline administrations of PDR) was deemed necessary in order to examine how the factor structure may vary depending upon the administration schedule (i.e., when administered once versus administered repeatedly).

Time 1 PDR

The first three study aims were all designed to provide insight into the factor structure and the construct underlying the PDR. Hypothesis 1, in which it was predicted that the PDR would have two factors, was partially supported. The initial factor analysis did, indeed, indicate that there were two factors. However, the second factor was excluded from any subsequent analyses due to its extremely poor internal consistency. These results indicated that when the PDR was administered once, as in the case of the Time 1 PDR, there was a single broad construct represented by 17 items with strong internal consistency. These items reflect a wide range of negative, oppositional, and socially inappropriate behaviors.

In hypothesis 2 it was predicted that the factor structure of the PDR would be equivalent across the English and Spanish versions of the PDR. Although this hypothesis was not supported, the lack of equivalence was based on only one item, "daydream," which did not load consistently. Specifically, the item loaded significantly on the Spanish version of the PDR but not on the English version. A qualitative comparison of the concept of daydreaming in English versus Spanish may help provide insight into what may account for this differential item loading. It may also be useful to determine whether there is factor equivalence across language when this item is excluded from the measure. Hypothesis 3, in which it was predicted that there would be factor equivalence across gender, was supported for the Time 1 PDR. All of the factor loadings were equal for both males and females.

The final study aim was exploratory in that there was no specific hypothesis regarding the expected number of classes. The findings of the latent class analysis supported a 3-class solution, based upon the level or severity of behavioral disturbance. Specifically, there were three groups of children: those with low levels of behavior problems, those with moderate levels, and those with severe behavior problems.

Time Composite PDR

Hypothesis 1, in which it was predicted that the PDR would have two factors, was supported for the Time Composite PDR. The factor analysis revealed a stable two-factor structure composed of 10 items with strong internal consistency. The first factor was comprised of six items reflecting oppositional behaviors, whereas the second

factor reflected more interpersonal forms of conflict as seen by the following four items: competitive, fight, tease, and jealous.

In hypothesis 2 it was predicted that the factor structure of the PDR would be equivalent across the English and Spanish versions of the PDR. This hypothesis was generally supported, despite the lack of *statistical* equivalence. Although the items "irritable," "fight," and "tease" were statistically different across language, the largest difference in their factor loadings was .25, and thus the factor structure of the Time Composite PDR was deemed to have *practical* equivalence across the English and Spanish versions. Hypothesis 3, in which it was predicted that there would be factor equivalence across gender, was also supported for the Time Composite PDR. The latent profile analysis conducted for the final study aim supported a 3-class solution, again based upon the level or severity of behavioral disturbance (i.e., low, moderate, severe).

Comparing the Time 1 PDR and Time Composite PDR

The results regarding the factor structure of the Time 1 PDR and Time Composite PDR generally suggest that the measure has sound psychometric properties and a stable factor structure, regardless of administration schedule. The exploratory tests of invariance also provided support for the use of both the Time 1 PDR and Time Composite PDR in that the results of these analyses indicated that the factor structure of the measure was consistent across several caregiver characteristics, specifically kinship status (i.e., kin vs. nonkin caregiver), ethnic match status (i.e., caregiver-child ethnic match vs. ethnic mismatch), and caregiver ethnicity (Caucasian, African

American, and Latino). These findings are consistent with the assertion made by

Chamberlain et al. (2006) that the PDR may be less susceptible to reporter biases

given its checklist format.

While there is support for both versions of the PDR, and thus either a single

administration or repeated administration schedule, there are relative advantages and

disadvantages for each. A single administration of the PDR will result in a valid

estimate of a broad range of negative behaviors and can even provide information

regarding the general level of behavioral disturbance. If it is not possible to administer

the PDR again due to time constraints or the child relocating, for example, the results

remain useful. Thus, the greatest asset of a single administration, given the context of

child welfare settings, may be its ability to provide a quick estimate of general

behavior problems.

However, repeated administration of the PDR is generally recommended

because this will provide a more specific understanding of problematic child

behaviors. The factor structure of the Time Composite PDR suggests that as the

measure is administered repeatedly (i.e., at least three times), it offers a more nuanced

view of child behavior problems in that it provides information regarding oppositional

behaviors *and* social conflict. This more refined understanding of the types of behavior

problems may be necessary to make well-informed decisions regarding placement and

treatment planning, for example. Further, since the PDR is a checklist, it is possible

that a single administration may not reflect the child's prototypical behavior (e.g., the

assessment may have occurred after an especially good or bad 24-hour period). Thus,

if a more reliable estimate of problematic child behaviors is needed, and time and resources permit, it is recommended that the PDR be administered on at least three occasions, which is consistent with the recent suggestion of Chamberlain and colleagues (2006).

Although the purpose of the present investigation was not to determine which items should be deleted from the original 31-item PDR, the collection of findings regarding the Time Composite PDR suggests that the 10 items that comprise it represent the most stable and consistent indicators of behavior problems. Thus, the items "argue," "back-talk," "defiant," "irritable," "negative," "not mind," "competitive," "fight," "tease," and "jealous" could be considered core indicators of behavior problems, and may represent the most efficient and reliable method to assess the behavioral disturbance of children and adolescents in child welfare settings.

Given these results, it may be useful to create a short-form PDR which is composed of the 10 items, and then also have an extended-form PDR, which can vary somewhat depending upon the population of interest. The 10 items included in the Time Composite PDR could be conceptualized as the core items that comprise the short-form PDR. Using the KEEP project as an example, the short-form PDR could be used during the 16-week intervention as a monitoring tool that is more efficient and targeted to the intervention. Given the brevity and simplicity of these items, the short-form PDR could be administered to caregivers using a recorded telephone program that instructs them to press "1" if the behavior occurred and "0" if the behavior did not occur.

The extended-form PDR could include low-frequency items and those that are more individualized to the population of interest. This extended version could include items that are endorsed infrequently such as "run away" and "stay out late" because they still provide relevant and valuable information. Further, it is important to remember that when a measure only focuses on occurrence or nonoccurrence of events, as in the case of a checklist, low frequency items are not expected to load consistently across time (Turner & Wheaton, 1995). Indeed, one of the strengths of the checklist format of the PDR is that it allows for sampling of a range of behavior problems, some more common than others. Those items not frequently endorsed provide valuable qualitative information and may serve as useful indicators of more severe behavioral disturbance. In the case of KEEP, the extended-form PDR could be used for the initial baseline assessment and also the assessment conducted at the termination of the intervention in order to gather more comprehensive pre- and postintervention data.

Insights Into Class Membership

The findings from the latent class analyses and logistic regression analyses provided insight into several factors that are associated with class membership (i.e., Low, Moderate, or Severe). Consistent with the existing literature on the developmental patterns of aggression (e.g., Nagin & Tremblay, 1999), the results indicated that older children were more likely to exhibit lower levels of behavior problems. In the case of both the Time 1 PDR and the Time Composite PDR, as the children got older, they were more likely to be classified into the Low group than

either the Severe or Moderate groups. Also consistent with the existing body of evidence, females were more likely to be in the Low than the Severe group, relative to males (e.g., American Psychiatric Association, 1994; Maccoby, 2004), for the Time Composite PDR.

Another interesting finding is that both African American and Latino children were more likely to be classified into the Low group than the Severe group, as compared to Caucasian children. A number of different explanations may help to explain this pattern of results. For example, if the African American and Latino children were more likely to be placed in kinship care, this may help to account for the findings given that there is evidence that youth in kinship care have lower levels of behavioral disturbance than those in nonrelative foster care (Benedict, Zuravin, & Stallings, 1996; Keller et al., 2001). However, in the present sample, 62 (32%) of the Caucasian children, 55 (31%) of the African American children, and 99 (38%) of the Latino children were placed in kinship care, so it does not appear that this explanation can account for the pattern of findings.

An alternative explanation is that there may be a lower threshold for placement into out-of-home care for minority children than Caucasian children. Chand (2000) suggests that children from ethnic minorities may be more likely to be reported to child welfare and removed from the home. A recent study by Lau et al. (2003) examined racial/ethnic differences in youth-reported maltreatment and rates of placement into out-of-home care and found support for this explanation. These researchers found that African American children were significantly more likely to be

placed in care, even after controlling for variables such as maltreatment, age, gender, and income. In fact, African American children were over 12 times more likely to be placed in foster care than Caucasian children, holding maltreatment history constant. Thus, it is possible that the Caucasian youth placed in foster care in the present sample may exhibit more severe behavioral profiles due to a higher threshold for being placed into out-of-home care.

Related to the threshold hypothesis, the type, frequency, and severity of maltreatment may also have contributed to the pattern of lower levels of behavioral disturbance among the African American and Latino youth. It is possible that the Caucasian children in this sample had more severe maltreatment histories, perhaps resulting in their more severe behavioral profiles. It would be useful to examine whether Latinos and African Americans, as compared to Caucasian children, experienced neglect at higher rates, as opposed to physical abuse or witnessing domestic violence, for example. Since children placed in kinship care are more likely to be removed due to neglect rather than abuse (Iglehart, 1994; Landsverk et al., 1996), and kinship care is more common among ethnic minority groups (Ehrle & Geen, 2002; C. F. Smith & Devore, 2004), it would be interesting to investigate these possible explanations using the present sample. Examining the reasons for removal from the home may help shed light on why the behavior problems of the minority youth were more likely to be classified as Low relative to the Caucasian youth.

Potential for Use in Child Welfare Settings

The PDR is appealing for use in child welfare settings for several reasons. First, the checklist format of the PDR provides an efficient estimate of the number of problematic child behaviors that are occurring in a given 24-hour period. Further, this measure has been validated for use with several of the diverse groups found among child welfare settings. The results of this investigation indicated that there was factor equivalence across gender, kinship status, and ethnic match status. There was also equivalence across caregiver ethnicity, which indicates that the PDR can be used with Caucasian, Latino, and African American foster parents. The PDR also appears to be valid when administered in Spanish to Latino caregivers, which is crucial given that Latino children and families constitute the fastest growing ethnic group in the child welfare system (Rivera, 2002). However, since this was the first investigation of the Spanish PDR, and there was a differential factor loading with the item "daydream," it is recommended that additional investigations be conducted in order to either support or challenge these conclusions.

An additional advantage of the PDR is that the reporter only has to recall whether or not a specific behavior occurred, rather than having to estimate the frequency or severity of behaviors, as is the case with behavior rating scales. The format of the PDR is precise in that it identifies the specific behaviors and time frame of interest (e.g., "In the last 24 hours, did your child fight?"). This format generally yields more reliable estimates of behavior than measures which contain items describing more global behaviors (Sattler, 2002). Although behavior rating scales

similarly specify the behaviors of interest, they also require reporters to make a subjective judgment regarding the specific behavior (e.g., Does your child fight "often," "once in a while," or "not at all?"). This subjectivity is an obstacle that is inherent to all behavior rating scales because of the reliance on reporters' (e.g., parents, teachers) perceptions of a child's behavioral functioning. Thus, relative to checklists, the use of rating scales requires more careful attention to the credibility of the reporter.

Related to the issue of reporter credibility, there is evidence that the reports of biological parents involved with child welfare may be especially biased. Reid and colleagues (1987) examined reporter bias in a sample of maltreating parents involved with child welfare and control parents with no history of maltreatment. Using a number of measures of behavior problems including the CBCL, they found that abusive parents overreported the presence of externalizing problems. These findings suggest that it may be difficult to obtain a valid assessment of the behavior problems of youth in child welfare if behavior rating scales completed by parents are relied upon in isolation. Thus, it may be beneficial to use behavior checklists such as the PDR to obtain a less biased estimate of the level of child behavior problems.

The checklist format of the PDR is also appealing for use in child welfare settings because it does not require the extended contact and familiarity with the child that is necessary for behavior rating scales. This relative advantage of checklists is particularly important to consider given the transitory nature of child welfare settings. A number of individuals, including caseworkers, biological parents, foster parents, or

teachers, for example, may be asked to report on a child's behavior. These individuals may have limited familiarity with the child in question, making it difficult to provide a valid estimate of the frequency or intensity of behavior problems. The checklist format of the PDR addresses this challenge in that it only requires the reporter to have known the child for a single 24-hour period.

There are additional characteristics of child welfare settings which make the PDR a particularly appealing measure of child behavior problems. Cost and resource limitations are two especially salient obstacles that must be considered (B. D. Smith & Donovan, 2003). Cameron and Vanderwoerd (1997) emphasize that any programs or procedures implemented in child welfare settings must consider financial viability. Given this context, it is quite appealing that it only takes 5 to 10 minutes to administer the PDR. In addition, the measure can be administered reliably after only 1 hour of training (Chamberlain & Reid, 1987). Since behavior rating scales are relatively more time consuming and also require more extensive training for interpretation (Sattler, 2002), the PDR would likely be a more cost effective method of obtaining an estimate of the level of child behavior problems.

Despite the simplicity of the PDR, it includes items that are conceptually similar to those found on more commonly used measures of behavior problems, especially in the case of the 10 items included in the Time Composite PDR. More specifically, the PDR items are similar to the items that compose the externalizing scales of measures such as the CBCL and BASC. The externalizing scale of the BASC, for example, contains items such as arguing, teasing, and stealing, all of which

are also included on the PDR. Similarly, the CBCL contains items that parallel those found on the PDR, such as disobeying, arguing, fighting, and teasing. In fact, a qualitative analysis of these measures reveals that the vast majority of the 31 items of the PDR are subsumed in the more comprehensive lists of behaviors found on the BASC and CBCL. Although this clearly does not establish the convergent validity of the PDR, it does suggest that the measure may be able to provide insight into similar types of behavior problems.

Given the relative efficiency and cost effectiveness of the PDR, the measure may be especially useful as a screening instrument in child welfare settings. Keil and Price (2006a) suggest that the PDR may serve as a useful initial assessment, which could then help guide subsequent decisions regarding assessment, appropriate placement, and treatment planning. The study conducted by Chamberlain and colleagues (2006) provided evidence that the PDR can predict placement disruptions, which are linked to further escalation of behavior problems (Newton et al., 2000). Thus, the level of behavior problems revealed by the PDR could help determine the level and type of care that is needed for youth exhibiting externalizing behaviors. For example, the needs of those youth exhibiting high levels of behavior problems may be more effectively met by placing them in treatment foster care with caregivers specially trained to manage youth with difficult behavioral profiles. Such a placement decision may help prevent subsequent disruptions and escalation of behavior problems.

The collection of findings from the present investigation corroborate the findings of Chamberlain and colleagues (2006) and are consistent with the suggestions

of Keil and Price (2006a) in that it provides additional support for the utility of the

PDR in child welfare settings. The results of the latent class analyses indicate that the

PDR can help identify subgroups of children with varying degrees of behavioral

disturbance, from Low to Moderate to Severe profiles. This has significant

implications for case management and treatment planning. For example, a child

that falls into the Low category may not be at high risk for placement disruptions and

may be able to be effectively maintained in the home. In contrast, youth with Severe

profiles and their families may need additional resources such as intensive wraparound

services or perhaps treatment foster care. This more individualized and targeted

approach to case management may result in more successful outcomes at the child,

family, and system levels.

CHAPTER 5

Limitations, Future Directions, and Conclusions

Limitations

There are several limitations of the present investigation that must be addressed. First, the sample used in this study was composed of foster parents that agreed to participate in a research study examining the effectiveness of the KEEP intervention. Although a majority of the foster parents that were contacted chose to participate (62%), there was a sizeable group of foster parents who declined to participate (38%), which may limit the generalizability of the findings. Given that the present sample was self-selected, it would be useful to examine the sample that declined to participate in the research study and compare it to the participating foster parents in order to determine whether there are meaningful differences between the two populations. For example, the number of children in the home, the number of years of experience as a foster parent, and socioeconomic status are just a few variables which may be important to consider.

An additional challenge related to the sample is that there was a significant multiracial category of foster children that could not be examined due to limitations in sample size. The ethnic comparisons in the present investigation only examined the three largest ethnic groups that were represented in the sample: Latino (37.4% of caregivers, 32.6% of children), Caucasian (27.4% of caregivers, 22.4% of children), and African American (25.7% of caregivers, 21.1% of children). Twenty-two percent of the children were classified as multiracial and 7% of the parents identified

themselves as multiracial. This methodological challenge of understanding and studying multiracial participants is not unique to the present study and is in fact appropriate given the diversity of the child welfare population. When considering the ethnic comparisons presented in this study, it is also important to remember that the foster parents reported the ethnicity of themselves and also their foster children. Thus, the child's ethnicity used in the study analyses is more accurately described as the caregiver's perception of child ethnicity. To determine the accuracy of caregivers' perceptions of child ethnicity, it may be useful to compare their report to that contained in the child welfare database.

Third, although both the English and Spanish versions of the PDR were examined in the present investigation, there are potentially confounding variables that must be considered. For example, among the Latino caregivers, some chose to participate in the Spanish version of the intervention, while others chose the English version. Of the 262 Latino caregivers, 70.5% chose to receive the Spanish version of the PDR, whereas the remaining 29.5% chose the English version. It would be worthwhile to examine factors such as acculturation and immigration history (e.g., number of years in the United States) which may have impacted this decision. It is possible that Latino caregivers that chose to participate in the study using Spanish are qualitatively different from those that chose to use English. For example, perhaps the Spanish-speaking Latino caregivers were significantly less acculturated than their English-speaking counterparts, which could then impact other important study variables such as attrition, caregiver satisfaction, and treatment outcome. Although the

data necessary to investigate these possibilities were not collected as part of Project KEEP, it would be beneficial for future investigations to collect such information in order to help tease apart complex and interrelated cultural variables such as language, ethnicity, and acculturation.

A final important limitation to consider is that the present investigation focused on the 31-item child welfare version of the PDR that was recently used in the effectiveness study of the KEEP intervention. Although it may be appealing to assume that the current findings generalize to all studies that have utilized the PDR, this would not be appropriate given that a small number of items are typically deleted or added to the measure depending upon the setting and population of interest. While the current study can certainly be used to inform research on different versions of the PDR, it is recommended that researchers utilizing altered versions of the measure conduct their own evaluations of its psychometric properties.

Future Directions

Despite these limitations, the present investigation has contributed significantly to the existing literature on the PDR. The generally favorable findings regarding the psychometric properties of the PDR can be used to inform and guide additional areas of research which would benefit from further examination. Suggestions for future research are grouped into three broad areas: (a) additional research focusing on the psychometric properties of the PDR, (b) applications of the present findings to Project KEEP and other intervention studies, and (c) systems-level research.

First, it would be beneficial to further examine the reliability and validity of the PDR. Although this study has begun to examine variables which may be significantly related to the factor structure of the PDR, such as kinship status and the number of days the child was in the home, there are numerous others to consider. Additional variables such as the duration of the phone call and time of day may be important factors to consider. For example, it is possible that as the duration of the phone call increases, foster parents may be more or less likely to endorse certain items. Or perhaps caregivers are more or less likely to endorse specific items in the evening relative to the morning hours. Studies examining these possibilities would help provide additional insight into the psychometric properties of the PDR.

Although caregiver age and ethnicity were examined in the present study, there are additional caregiver characteristics which may impact their report of problematic child behaviors. Specifically, there is a growing body of evidence which suggests that parental psychopathology can result in biased reports of behavior problems. For example, Webster-Stratton and Hammond (1990) found that higher levels of parental depression significantly predicted more negative perceptions of child adjustment for both mothers and fathers. In addition, level of maternal stress has been found to be significantly positively related to the report of child behavior problems (Eyberg, Boggs, & Rodriguez, 1992; Seiffge-Krenke & Kollmar, 1998). Since data regarding caregiver stress were collected as part of Project KEEP, it would be useful to examine the relation between caregiver stress and the number and type of child behavior problems.

Foster parent age and the number of days the child was in the home were two variables that were found to be significantly associated with particular items of both the Time 1 PDR and Time Composite PDR. For the Time 1 PDR, there was a relation between caregiver age and the following items: complain, lie, and irresponsible. In the case of the Time Composite PDR, factor loadings varied according to both the number of days the child was in the home (for the items "competitive" and "not mind") and foster parent age (for the item "back-talk"). It is unclear at the present time what could account for these differential relations. One possibility is that as caregivers get older, they become less tolerant of behavior problems or perhaps they have a lower threshold for stress. For the items "competitive" and "not mind," which were less likely to be endorsed as the child was in the home longer, it is possible that as the child becomes more acquainted to their new surroundings, they are less likely to disobey house rules and engage in competitive behavior with others who live in the home, such as other foster children. These are just a couple of plausible explanations which would benefit from further investigation.

With regards to the psychometric properties of the PDR, a final suggestion is to evaluate the convergent validity of the measure, which could be done relatively easily given that Project KEEP included other more commonly used measures of behavior problems such as the ECBI. Qualitative analysis of the items of the PDR and ECBI reveal overlapping item content. However, statistical analyses are needed in order to provide empirical evidence of convergent validity. These analyses would provide

insight into the relation between the PDR and ECBI and, more broadly, how the PDR

compares to measures that are believed to have similar underlying constructs.

The findings from the present investigation also have implications for the larger

KEEP effectiveness study. For example, the three classes of children (i.e., Low,

Moderate, Severe) identified by the latent class analyses can now be used in order to

determine if the KEEP intervention is more or less effective for a particular behavioral

profile. In other words, class membership may be a significant predictor of treatment

outcome. If this is indeed the case, the PDR could be used to identify those youth in

child welfare that are likely to receive the most benefit from the KEEP intervention.

The PDR could also be used to identify youth that require more intensive

interventions. This targeted approach to intervention may represent a more affordable

and effective way of meeting the needs of youth with externalizing behavior problems

in child welfare settings.

The findings from this study are also relevant to research beyond the KEEP

effectiveness study. There are currently several research studies examining the

effectiveness of evidence-based treatments in child welfare settings (Keil & Price,

2006b). Specifically, parent-child interaction therapy, parent management training, and

multisystemic therapy are among the approaches being applied to child welfare

populations. The PDR could serve as an efficient measure of the level of child

behavior problems pre- and posttreatment. Assuming that these interventions are

indeed found to be effective in child welfare settings, the PDR may also be an

extremely helpful tool for researchers in that it can be used to determine which

children are likely to benefit the most from these interventions. Essentially the PDR

could be used to help triage the limited resources found among child welfare settings.

Given that it is recommended that the PDR be used within child welfare

settings, systems-level research is needed, particularly because this line of research has

yet to be explored. Obstacles at the systems level must be considered and evaluated

because successful dissemination into community settings requires the feedback and

cooperation of service providers. It is recommended that future investigations examine

the use of the PDR by caseworkers in routine practice. For example, researchers could

compare the experiences of caseworkers using "practice as usual" versus those that

are trained to use the PDR as part of their routine practice. Qualitative studies of

caseworkers' experiences with the PDR, in addition to more quantitative measures of

caseworker stress and satisfaction, would help provide insight into the acceptability of

the measure.

Another important research step would be to examine the utility of the PDR

when used by caseworkers. This line of research shifts the focus from the use of the

PDR as a tool for researchers, to a tool used by caseworkers to facilitate effective case

management. For example, the PDR may help caseworkers choose a more appropriate

out-of-home placement for children with severe behavioral profiles because it is a

relatively simple method of assessing behavioral functioning. Thus, the PDR may help

caseworkers find the optimal first placement for youth in child welfare settings. Once a

child has been placed, the PDR could be used to help caseworkers identify youth that

need to be referred for additional services such as outpatient psychotherapy, which

could therefore improve caseworkers' responsiveness to the needs of foster families.

The number of referrals for intervention and foster parents' satisfaction with their

caseworkers may be interesting outcome measures to consider. If reunification occurs,

caseworkers could use the PDR as a way to monitor any changes in a child's

behavioral functioning. In summary, the PDR may have great potential for use by

caseworkers throughout the removal and reunification process.

Finally, formal cost effectiveness and feasibility studies are needed in order

to provide support for the implementation and dissemination of the PDR into child

welfare settings. Although the brevity of the PDR certainly suggests that it would be a

cost effective method of assessing problematic child behaviors, there are currently no

studies that focus on the economic costs of implementing the measure. Taken together,

these suggestions indicate that systems-level research is an imperative next step that

will provide the necessary insight into the acceptability and performance of the PDR

when used in child welfare settings.

Conclusions

The primary goal of the present investigation was to examine the factor

structure of the Parent Daily Report and its potential for use in child welfare settings.

The collection of findings generally suggests that the measure has sound psychometric

properties and a stable factor structure. The factor structure of the PDR is equivalent

across several important groups found among diverse child welfare samples, such as

language (when administered repeatedly), gender, kinship status, and foster parent

ethnicity. Although the factor structure of the PDR is somewhat different depending

upon whether it is administered once or on several occasions, the measure provides a

useful estimate of the level of behavioral disturbance, regardless of the number of

administrations. In addition, the PDR appears to have utility in child welfare settings

in that it can be used to predict placement disruptions. Given these strengths, the PDR

has great promise for more widespread use and dissemination into child welfare

settings and future investigations should focus on examining the feasibility and

acceptability of the measure.

REFERENCES

Achenbach, T. M. (1991a). *Manual for the Child Behavior Checklist and 1991 Profile.* Burlington, VT: University of Vermont Department of Psychiatry.

Achenbach, T. M. (1991b). *Manual for the Teacher's Report Form and 1991 Profile.* Burlington, VT: University of Vermont Department of Psychiatry.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

American Psychological Association. (2002). Ethical principles of psychologists and code of conduct. *American Psychologist, 57*, 1060-1073.

Ayon, C., & Lee, C. D. (2005). A comparative analysis of child welfare services through the eyes of African American, Caucasian, and Latino parents. *Research on Social Work Practice, 15*, 257-266.

Becker, W. C., Madsen, C. H., Arnold, C. R., & Thomas, D. R. (1967). The contingent use of teacher attention and praise in reducing classroom behavior problems. *Journal of Special Education, 1*, 287-307.

Benedict, M. I., Zuravin, S., & Stallings, R. Y. (1996). Adult functioning of children who lived in kin vs. non-relative family foster homes. *Child Welfare, 75*, 529-549.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*, 238-246.

Boggs, S. R., Eyberg, S. M., & Reynolds, L. A. (1990). Concurrent validity of the Eyberg Child Behavior Inventory. *Journal of Clinical Child Psychology, 19*, 75-78.

Bower, G. H. (1981). Mood and memory. *American Psychologist, 36*, 129-148.

Cameron, G., & Vanderwoerd, J. (1997). *Protecting children and supporting families: Promising programs and organizational realities.* New York: Aldine de Gruyter.

Chamberlain, P., Moreland, S., & Reid, K. (1992). Enhanced services and stipends for foster parents: Effects on retention rates and outcomes for children. *Child Welfare, 5*, 387-401.

Chamberlain, P., Price, J. M., Reid, J. B., Landsverk, J., Stoolmiller, M., & Fisher, P. A. (2006). Who disrupts from placement in foster and kinship care? *Child Abuse & Neglect, 30*, 409-424.

Chamberlain, P., & Reid, J. B. (1987). Parent observation and report of child symptoms. *Behavioral Assessment, 9*, 97-109.

Chamberlain, P., & Reid, J. B. (1991). Using a specialized foster care community treatment model for children and adolescents leaving the state mental hospital. *Journal of Community Psychology, 19*, 266-276.

Chand, A. (2000). The over-representation of Black children in the child protection system: Possible causes, consequences and solutions. *Child and Family Social Work, 5*, 67-77.

Clausen, J. M., Landsverk, J., Ganger, W., Chadwick, D., & Litrownik, A. (1998). Mental health problems of children in foster care. *Journal of Child and Family Studies, 7*, 283-296.

Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Conners, C. K. (1969). A teacher rating scale for use in drug studies with children. *American Journal of Psychiatry, 126*, 884-888.

Conners, C. K. (1990). *Conners' rating scales manuals*. Toronto: Multi-Health Systems.

Courtney, M. E. (1999). Foster care and the costs of welfare reform. In P. A. Curtis, G. Dale, & J. C. Kendall (Eds.), *The foster care crisis* (pp. 129-151). Lincoln: University of Nebraska Press.

Crick, N. R., Grotpeter, J. K., & Bigbee, M. A. (2002). Relationally and physically aggressive children's intent attributions and feelings of distress for relational and instrumental peer provocations. *Child Development, 73*, 1134-1142.

Curtis, P. A. (1999). The chronic nature of the foster care crisis. In P. A. Curtis, G. Dale, & J. C. Kendall (Eds.), *The foster care crisis* (pp. 1-14). Lincoln: University of Nebraska Press.

De Los Reyes, A., & Kazdin, A. E. (2005). Informant discrepancies in the assessment of childhood psychopathology: A critical review, theoretical framework, and recommendations for further study. *Psychological Bulletin, 131*, 483-509.

Ehrle, J., & Geen, R. (2002). Kin and non-kin foster care – Findings from a national survey. *Children and Youth Services Review, 24*, 15-35.

Embry, R. A., Buddenhagen, P., & Bolles, S. (2000). Managed care and child welfare: Challenges to implementation. *Children and Youth Services Review, 22*, 93-116.

Eyberg, S. M., Boggs, S. R., & Rodriguez, C. M. (1992). Relationships between maternal parenting stress and child disruptive behavior. *Child and Family Behavior Therapy, 14*, 1-9.

Eyberg, S. M., & Robinson, E. A. (1983). Conduct problem behavior: Standardization of a behavioral rating scale with adolescents. *Journal of Clinical Child Psychology, 12*, 347-354.

Federal Guidelines. (2001). Child and Family Services Reviews, administered by the Children's Bureau, Administration for Children and Families, U.S. Department of Human Services.

Fisher, P. A., Burraston, B., & Pears, K. (2005). The Early Intervention Foster Care Program: Permanent placement outcomes from a randomized trial. *Child Maltreatment, 10*, 61-71.

Fisher, P. A., Gunnar, M. R., Chamberlain, P., & Reid, J. B. (2000). Preventive intervention for maltreated preschool children: Impact on children's behavior, neuroendocrine activity, and foster parent functioning. *Journal of the American Academy of Child and Adolescent Psychiatry, 39*, 1356-1364.

Forgatch, M. A., & Toobert, D. J. (1979). A cost-effective parent training program for use with normal preschool children. *Journal of Pediatric Psychology, 4*, 129-145.

Gorman-Smith, D., & Loeber, R. (2005). Are developmental pathways in disruptive behaviors the same for girls and boys? *Journal of Child and Family Studies, 14*, 15-27.

Gould, S. J. (1996). *The mismeasure of man.* New York: Norton.

Groth-Marnat, G. (2003). *Handbook of psychological assessment* (4th ed.). New York: Wiley.

Helfinger, C. A., Simpkins, C. G., & Combs-Orme, T. (2000). Using the CBCL to determine the clinical status of children in state custody. *Children and Youth Services Review, 22*, 55-73.

Hines, A. M., Lemon, K., Wyatt, P., & Merdinger, J. (2004). Factors related to the disproportionate involvement of children of color in the child welfare system: A review and emerging themes. *Children and Youth Services Review, 26*, 507-527.

Hoyle, R. H. (2000). Confirmatory factor analysis. In H. E. A. Tinsley & S. D. Brown (Eds.), *Handbook of applied multivariate statistics and mathematical modeling* (pp. 465-497). San Diego, CA: Academic Press.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1-55.

Hulsey, T. C., & White, R. (1989). Family characteristics and measures of behavior in foster and nonfoster children. *American Journal of Orthopsychiatry*, 59, 502-509.

Iglehart, A. (1994). Kinship foster care: Placement, service, and outcome issues. *Children and Youth Services Review, 16*, 107-122.

Jones, R. R. (1974). Observation by telephone. *Oregon Research Institute Technical Report, 14*, 1.

Keil, V., & Price, J. M. (2006a). Externalizing behavior disorders in child welfare settings: Definition, prevalence, and implications for assessment and treatment. *Children and Youth Services Review, 28*, 761-779.

Keil, V., & Price, J. M. (2006b). *Treatment of externalizing behavior disorders in child welfare settings: Integrating evidence-based practice.* Manuscript submitted for publication.

Keller, T. E., Wetherbee, K., Le Prohn, N. S., Payne, V., Sim, K., & Lamont, E. R. (2001). Competencies and problems behaviour of children in family foster care variations by kinship placement status and race. *Children and Youth Services Review, 23,* 915-940.

Landsverk, J., Davis, I., Ganger, W., Newton, R., & Johnson, I. (1996). Impact of child psychosocial functioning on reunification from out-of-home care. *Children and Youth Services Review, 18*, 447-462.

Lau, A. S., McCabe, K. M., Yeh, M., Garland, A. F., Hough, R. L., & Landsverk, J. (2003). Race/ethnicity and rates of self-reported maltreatment among high-risk youth in public sectors of care. *Child Maltreatment, 8*, 183-194.

Lett, N. J., & Kamphaus, R. W. (1997). Differential validity of the BASC Student Observation System and the BASC Teacher Rating Scale. *Canadian Journal of School Psychology, 13*, 1-14.

Maccoby, E. E. (2004). Aggression in the context of gender development. In M. Putallaz & K. L. Bierman (Eds.), *Aggression, antisocial behavior, and violence among girls: A developmental perspective* (pp. 3-22). New York: Guilford Press.

Margalit, M. (1983). Diagnostic application of the Connors Abbreviated Symptom Questionnaire. *Journal of Clinical Child Psychology, 12*, 355-357.

Maruyama, G. M. (1998). *Basics of structural equation modeling.* Thousand Oaks, CA: Sage Publications.

McIntyre, A., & Keesler, T. Y. (1986). Psychological disorders among foster children. *Journal of Clinical Child Psychology*, 15, 297-303.

Nagin, D., & Tremblay, R. E. (1999). Trajectories of boys' physical aggression, opposition, and hyperactivity on the path to physically violent and nonviolent juvenile delinquency. *Child Development, 70*, 1181-1196.

Needell, B., Webster, D., Cuccaro-Alamin, S., Armijo, M., Lee, S., Brookhart, A., et al. (2001). *Child Welfare Service reports for California.* Retrieved July 11, 2005, from http://www.cssr.berkeley.edu/CWSCMSreports/

Newton, R. R., Litrownik, A. J., & Landsverk, J. A. (2000). Children and youth in foster care: Disentangling the relationship between problem behaviors and number of placements. *Child Abuse and Neglect, 24*, 1363-1374.

Patterson, G. R. (1964). An empirical approach to the classification of disturbed children. *Journal of Clinical Psychology, 20*, 326-337.

Patterson, G. R. (1976). Interventions for boys with conduct problems: Multiple settings, treatments, and criteria. In C. M. Franks & G. T. Wilson (Eds.), *Annual review of behavior therapy: Theory and practice.* New York: Brunner/Maze.

Patterson, G. R., Chamberlain, P., & Reid, J. B. (1982). A comparative evaluation of parent training procedures. *Behavior Therapy, 13*, 638-650.

Pilowsky, D. (1995). Psychopathology among children placed in family foster care. *Psychiatric Services*, 46, 906-910.

Price, J. M., Chamberlain, P., Reid, J. B., & Landsverk, J. (2006). *Training and support for foster parents: Outcomes for child problem behavior and placement stability.* Manuscript submitted for publication.

Quay, H. C., & Peterson, D. R. (1987). *Manual for the Revised Behavior Problem Checklist.* Coral Gables, FL: Author.

Quay, H. C., & Peterson, D. R. (1996). *Manual for the Revised Behavior Problem Checklist – PAR version.* Odessa, FL: Psychological Assessment Resources.

Reid, J. B., Kavanagh, K., & Baldwin, D. V. (1987). Abusive parents' perceptions of child problem behaviors: An example of parental bias. *Journal of Abnormal Child Psychology, 15*, 457-466.

Reid, J. B., & Patterson, G. R. (1976). The modification of aggression and stealing behavior of boys in the home setting. In E. Ribes-Inesta & A. Bandura (Eds.), *Behavior modification: Experimental analysis of aggression and delinquency.* Hillsdale, NJ: Lawrence Erlbaum.

Reynolds, C. R., & Kamphaus, R. W. (1992). *Behavior Assessment System for Children*. Circle Pines, MN: American Guidance.

Rio, A. T., Quay, H. C., Santisteban, D. A., & Szapocznik, J. (1989). Factor-analytic study of a Spanish translation of the Revised Behavior Problem Checklist. *Journal of Clinical Child Psychology, 18*, 343-350.

Rivera, H. P. (2002). Developing collaborations between child welfare agencies and Latino communities. *Child Welfare, 81*, 371-384.

Robinson, E. A., & Eyberg, S. M. (1981). The dyadic parent-child interaction coding system: Standardization and validation. *Journal of Consulting and Clinical Psychology, 49*, 245-250.

Robinson, E. A., Eyberg, S. M., & Ross, A. W. (1980). The standardization of an inventory of child conduct problem behaviors. *Journal of Clinical Child Psychology, 9*, 22-28.

Sattler, J. M. (2002). *Assessment of children: Behavioral and clinical applications*. San Diego, CA: Author.

Seiffge-Krenke, I., & Kollmar, F. (1998). Discrepancies between mothers' and fathers' perceptions of sons' and daughters' problem behaviour: A longitudinal analysis of parent-adolescent agreement on internalising and externalising problem behaviour. *Journal of Child Psychology and Psychiatry, 39*, 687-697.

Silver, S. E., Duchnowksi, A. J., Kutash, K., Friedman, R. M., Eisen, M., Prange, M. E., et al. (1992). A comparison of children with serious emotional disturbance served in residential and school settings. *Journal of Child and Family Studies, 1*, 43-59.

Smith, B. D., & Donovan, S. E. F. (2003). Child welfare practice in organizational and institutional context. *The Social Service Review, 77*, 541-563.

Smith, C. F., & Devore, W. (2004). African American children in the child welfare and kinship system: From exclusion to over inclusion. *Children and Youth Services Review, 26*, 427-446.

Snowden, L. R. (2003). Bias in mental health and intervention. *American Journal of Public Health, 93*, 239-243.

Steiger, J. S. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research, 25*, 173-180.

Stone, A. A., Broderick, J. E., Kaell, A. T., DelesPaul, P. A. E. G., & Porter, L. E. (2000). Does the peak-end phenomenon observed in laboratory pain studies apply to real-world pain in rheumatoid arthritics? *Journal of Pain, 1*, 212-217.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Boston, MA: Allyn & Bacon.

Thompson, A. H., & Fuhr, D. (1992). Emotional disturbance in fifty children in the care of a child welfare system. *Journal of Social Service Research, 15*, 95-112.

Tourangeau, R. (2000). Remembering what happened: Memory errors and survey reports. In A. A. Stone & J. S. Turkkan (Eds.), *The science of self-report: Implications for research and practice* (pp. 29-47). Mahwah, NJ: Lawrence Erlbaum Associates.

Turner, R. J., & Wheaton, B. (1995). Checklist measurement of stressful life events. In S. Cohen, R. C. Kessler, & L. U. Gordon (Eds.), *Measuring stress: A guide for health and social scientists* (pp. 29-58). New York: Oxford University Press.

Urquiza, A. J., Wirtz, S. J., Peterson, M. S., & Singer, V. A. (1994). Screening and evaluating abused and neglected children entering protective custody. *Child Welfare, 123*, 155-171.

U.S. Census Bureau. (2000). *State and County quick facts: California.* Retrieved July 11, 2005, from http://www.quickfacts.census.gov/qfd/states/06000.html

U.S. Department of Health and Human Services. (2001a). *Foster care national statistics.* Retrieved July 13, 2005, from http://www.calib.com/nccanch/ pubs/factsheets/foster.cfm

U.S. Department of Health and Human Services. (2001b). *Mental health: Culture, race, and ethnicity – A supplement to mental health: A report of the Surgeon General.* Rockville, MD: Author.

Webster-Stratton, C., & Hammond, M. (1990). Predictors of treatment outcome in parent training for families with conduct problem children. *Behavior Therapy, 21*, 319-337.

Webster-Stratton, C., & Hammond, M. (1997). Treating children with early-onset conduct problems: A comparison of child and parent training interventions. *Journal of Consulting and Clinical Psychology, 65*, 93-109.

Webster-Stratton, C., Hollinsworth, T., & Kolpacoff, M. (1989). The long-term effectiveness and clinical significance of three cost-effective training programs for families with conduct-problem children. *Journal of Consulting and Clinical Psychology, 57*, 550-553.

Webster-Stratton, C., Kolpacoff, M., & Hollinsworth, T. (1988). Self-administered videotape therapy for families with conduct-problem children: Comparison with two cost-effective treatments and a control group. *Journal of Consulting and Clinical Psychology, 56*, 558-566.

Weinrott, M. R., Bauske, B., & Patterson, G. R. (1979). Systematic replication of a social learning approach. In P. O. Sjoden, S. Bates, & W. S. Dockens (Eds.), *Trends in behavior therapy* (pp. 331-352). New York: Academic Press.