The Pennsylvania State University

The Graduate School

College of Education

RELIABILITY AND VALIDITY EVIDENCE FOR THE CONFUSION, HUBBUB, AND ORDER SCALE (CHAOS) WHEN USED IN RURAL HOMES

A Dissertation in

School Psychology

by

Sarah A. Wollersheim Shervey

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

August 2013

The dissertation of Sarah A. Wollersheim Shervey was reviewed and approved* by the following:

James C. DiPerna
Associate Professor of Education
Professor-in-Charge, School Psychology Program
Dissertation Adviser, Chair of Committee

Thomas W. Farmer
Associate Professor of Education

Mark T. Greenberg
Professor of Human Development and Psychology

Shirley A, Woika
Assistant Professor of Education

*Signatures are on file in the Graduate School.

Abstract

The purpose of the current study was to examine reliability and validity of scores from the Confusion, Hubbub, and Order Scale (CHAOS; Matheny, Wachs, Ludwig, & Phillips, 1995). Participants included 1,292 rural families drawn from a larger longitudinal study, the Family Life Project (Vernon-Feagans, Cox, & the Family Life Project Key Investigators, in press). Confirmatory Factor Analyses (CFAs) were conducted to examine the structural validity of the CHAOS. Results of these analyses indicated that there was no clear best-fitting model for the data, and not all items were equivalent across income groups. In addition, a repeated measures 2 x 2 ANOVA indicated that CHAOS scores were higher at a follow-up home visit. Regression analysis indicated that the primary caregiver's partner living in the household, number of recent job changes, and number of hours of TV per week were significantly related to CHAOS scores for both the low and middle/high income group, and the overall model accounted for 33% of the variance in CHAOS scores across scores. Because psychometric evidence for scores from the CHAOS was mixed, use of the CHAOS should be considered with caution.

Table of Contents

List of Tables

List of Figures

CHAPTER 1: INTRODUCTION AND LITERATURE REVIEW

Environmental factors have a significant impact on the development of children (e.g., Vernon-Feagans, Garrett-Peters, DeMarco, & Bratsch, 2012; Wachs & Evans, 2010). Bronfenbrenner (1999) described four spheres of context in which a person operates: the microsystem, mesosystem, exosystem, and macrosystem. While chaos at any of these levels can have an impact on childhood outcomes, the current study focused on the measurement of chaos at the microsystem level. Matheny, Wachs, Ludwig, and Phillips (1995) designed the Confusion Hubbub and Order Scale (CHAOS) to capture the "environmental confusion" that occurs in one aspect of Bronfenbrenner's microenvironment, referred to as "physical microenvironment" (p. 430).

Based on Bronfenbrenner's model, where all systems act on each other, and all aspects of each system act on each other, it is difficult to study one aspect of the microsystem in isolation. However, in order to better study aspects of a child's multilevel-multidimensional environment, it is important to quantify what children and their families are experiencing. The CHAOS attempts to capture environmental confusion in the physical microenvironment, which Matheny et al. described as "potentially stressful, nonspecific background factors such as noise, crowding, and situaltional traffic patterns" (p. 430). By identifying ways to measure environmental confusion, researchers can begin to form a more complete picture of this small piece of the multilevel, multidimensional puzzle. The purpose of the current study was to examine reliability and validity of scores from the CHAOS.

Impact of Household Chaos

Previously, it was thought that significant stimulation, especially noise, aided in language development (Wachs & Evans, 2010). Results of recent research on environmental chaos, however, suggested that high levels of chaos can have adverse effects on development. Specifically, high levels of chaos can impact the child-parent relationship (e.g., Dumas, Nissley, Nordstrom, Smith, Prinz, & Levine, 1995; McLoyd 1990; Wachs & Evans, 2010). Parents who reside in highly chaotic environments are less responsive, exhibit lower levels of involvement, and provide less stimulation relative to parents in a household with lower chaos levels. These byproducts of chaos, in turn, lead to developmental impairment. Because of these new research findings, Wachs and Evans (2010) conceptualized a curvilinear relationship between level of stimulation and development, and they concluded there should be a medium level of stimulation for optimal development. Conversely, high and low levels of stimulation can lead to impaired development.

Lichter and Wellington (2010) examined factors that influence chaos from a historical context. Demographic features such as the age of the youngest child in the home, how many parents reside in the home, how many geographic moves a child experiences, and how many adults are employed outside of the home all influenced levels of chaos. Additionally, they asserted that, “poverty and chaos are inextricably linked in children’s lives” (p. 17). However, chaos also has been shown to contribute independently to variability in child outcomes (Vernon-Feagans, Garett-Peters, Willoughby, Mills-Koonce, & the Family Life Project Key Investigators [FLPKI], 2012). Levels of environmental chaos have been shown to relate to the physiological, socioemotional, behavior, cognitive, and academic, outcomes of children (e.g., Evans, 2006, Matheny et al., 1995).

Environmental or household chaos has been defined in several different ways. Bronfrenbrenner and Evans (2000) characterized household chaos as “frenetic activity, lack of structure, unpredictability in everyday activities, and high levels of ambient stimulation” (p.

121). Further, Evans and Wachs (2010) described household chaos as “environments that are characterized by high levels of noise, crowding, and instability, as well as a lack of temporal and physical structuring (few regularities, routines, or rituals); nothing has its time or place” (p. 448; as cited in Vernon-Feagans, Garrett-Peters, DeMarco, et al., 2012). Definitions of household and environmental chaos, as specified in key studies of the relationship between chaos and child outcomes, are reported in Table 1.

Negative Outcomes Associated With Household Chaos

Regardless of how household chaos is defined, it is important to examine the measurement of environmental chaos because it can affect children’s development in a variety of ways. Household chaos and other chaos-related variables such as noise and crowding have been related to negative childhood outcomes. Negative childhood outcomes have been demonstrated in physiological, social, emotional, behavioral, cognitive, and academic domains (e.g., Adams, 2004; Evans, 2006; Vernon-Feagans, Garrett-Peters, DeMarco, et al., in press). Table 1 displays information about the sample, chaos measure, outcome measures, and magnitude of the relationship between chaos and outcome variables for studies that demonstrate the impact of household chaos.

Chaos and physiological problems. Environmental or household chaos variables have been linked to many negative health and physiological outcomes (Evans, 2006; Evans, Hygge, & Bullinger, 1995; Evans, Kliewer, & Martin, 1991; Evans, Lepore, Shejwal, & Palsane, 1998). A higher level of noise, for example, was related to higher blood pressure (Evans, 2006; Evans et al., 1991). Further, Evans, Hygge, et al. found that chronic noise was associated with elevated neuroendocrine measures and lower blood pressure reactivity.

Table 1 *Studies Examining Household Chaos and Related Constructs*

Study	Conceptualizing chaos	Sample	Chaos measure used	Outcome measures	Relationship with Household Chaos
Adam & Chase-Lansdale (2002)	"Family instability… parental separations and family mobility" (p. 792)	267 adolescent African American females	1. Family Distribution Variables: number of moves to new homes, separations from parents, quality of parental relationships, family support, and peer relationship quality 2. Perceived Quality of Neighborhood Scale	1. report card grades 2. The Center for Epidemiological Studies-Depression Scale 3. The National Longitudinal Survey of Youth, The Youth Deviance Scale 4. whether or not the participants were sexually active	adjustment problems 1. number of moves (R^2 = .10) 2. number of separations (R^2 = .05)
Bada et al. (2008)	"unpredictability in the home environment" (p. 173)	1388, age 3, urban environment, 78% African American	Disruptions in caregivers (Records of interventions of child protective services)	1. Child Behavior Checklist 2. Vineland Adaptive Behavior Scale	1. communication (decrease of 1/6 SD) 2. daily living (decrease of 1/6 SD)
Billows, Gradisar, Dohnt,	"Disorganized and unstructured home	217, Australian, ages 13-18	CHAOS	1. Sleep Hygiene Index 2. sleep onset latency 3. total sleep time	1. 1. total sleep time, sleep onset latency,

Johnston, McCappin, & Hudson (2009)	environment" (p.745)				and daytime sleepiness (R^2 = .11)
Chen, Cohen, & Miller (2010)	"crowding, greater noise exposure, and less predictable routines" (p. 32)	54 children, ages 9-18	CHAOS	1. Cortisol	1. R^2 = .25
Coldwell, Pike, & Dunn (2006)	"an environment that is high in noise and crowding and low in regularity and routines" (p. 1116)	118 English families, ages 4-8	1. Short-form CHAOS	1. Berkeley Puppet Interview 2. Parent-child Relationship Scale 3. Strengths and Difficulties	1. R^2 = .11-.31
Corapci & Wachs (2002)	"high levels of crowding, home traffic (number of people coming and going in the home), and ambient background noise, as well as temporal and structural regularity" (p. 183)	57, parent infant dyads	1. Purdue Home Stimulation Inventory (PHSI) 2. CHAOS	1. Parenting Sense of Competency Scale 2. Maternal Self-Efficacy Scale 3. Profiles of Mood State	Confusion-Crowding chaos factor score: 1. Lower levels of developmentally facilitative caregiver-child transactions (R^2 = .32) 2. higher levels of nonfaciliative caregiver child transactions (R^2 = .11) Noise-chaos factor

					score: 1. Lower levels of facilitative caregiver-child transactions (R^2 = .13) 2. Higher levels of nonfaciliative caregiver child transactions (R^2 = .08)
Deater-Deckard, Mulluineaux, Beckman, Petrill, Schatschneider, & Thompson (2009)	“noise levels, crowding and ‘traffic’ (people coming and going all the time), lack of predictability and family” routines” (p. 1301)	302 families, varying SES	CHAOS	1. Stanford-Binet 2. Disruptive Behavior Rating Scale	1. R^2 = .05
Dumas, Nissley, Nordstrom, Smith, Prinz, & Levine (2005)	“high levels of confusion and agitation in the home, as well as a sense of rush, disorganization, and pressure of time in daily	Study 1 106 mothers and pre school aged children, primarily Caucasian	Study 1 1. CHAOS	Study 1 1. Parenting Stress Index, Short Form 2. Parenting Scale 3. Modified Block Child Rearing Practices Report 4. Social Competence and	1. parenting quality and child adjustment (*r* = .27-.48) 2. child performance (*r* = .08-.29)

	routines" (p. 93)			Behavior Evaluation, Short Form 5. Children's Behavior Questionnaire 6. Parent-Child Toy-Sorting Task 7. Wechsler Preschool and Primary Scale of Intelligence 8. Emotion Identification task	
		Study 2 56 children, primarily African American	Study 2 1. CHAOS 2. Neighborhood Questionnaire	Study 2 1. Child Behavior Checklist, Parent and Teacher Forms	
Evans, Gonnella, Marcynyszyn, Gentile, & Salpekar (2005)	"Frenetic activity, lack of structure, and unpredictability, in conjunction with intense background stimulation" (p. 560)	223, upstate New York, oversample of children in poverty	1. CHAOS 2. Family Rituals Questionnaire 3. Family Routines Inventory	1. unsolvable puzzle task 2. Rutter Child Behavior Questionnaire 3. Children's Self-Control Scale	1. learned helplessness (ΔR^2 = .03) 2. psychological distress (ΔR^2 = .05) 3. self-regulatory behavior (ΔR^2 = .04)
Evans, Hygge, & Bullinger	"chronic noise" (p. 333)	135, 3rd and 4th graders	High or low noise area	1. blood pressure 2. urinary epinephrine and	1. practical significance

(1995)				norepinephrine 3. cortisol 4. attention 5. memory 6. Biglmaier Reading Test 7. Motivation 8. estimate of annoyance from 1 to 100 9. quality of life.	not reported
Evans, Lepore, Shejwal, & Palsane (1998)	"chronic residential crowding" (p. 1514)	281, India, 10-12 year old children	Living in crowded urban city	1. adjustment problems 2. blood pressure 3. social support 4. helplessness	1. academic achievement (R^2 = -.19) 2. adjustment problems (R^2 = .07) 3. Systolic BP (R^2 = .19) 4. Diastolic BP (R^2 = .12) 5. Parent support (R^2 = -.18) 6. parent conflict (R^2 = .37) 7. learned helplessness (R^2 = .0)1
Evans, Maxwell, & Hart (1999)	"high density living conditions"	42 Midwestern families, all SES represented	Person-to-room ratio room-"the	1. Language diversity - "the sum of different nouns, adjectives, and	1. 1. ΔR^2=.07

	(p.1020)		presence of a floor-to-ceiling wall with a defined entry separating the area from a hall or another room" (p. 1021).	adverbs addressed to the child per hour" (p. 1021) 2. Parental verbal responsiveness - "the proportion of all single parental verbal utterances to the child within 5 s of the target child's behavior not preceded by parental verbal initiation" (p. 1021).	
Fontain, McCrory, Boivin, Moffitt, & Viding (2011)	"disorganized, noisy household" (p. 732)	9,578 (Twins Early Development Study), ages 7 to 12, 94.4 % Caucasian	CHAOS	1. Strengths and Difficulties Questionnaire 2. Antisocial Process Screening Device 3. Parent Feelings Questionnaire 4. Parent Discipline (using semi-structured interview)	1. Odds ratios of 1.6-2.4 predicted membership in conduct problems and callous-unemotional groups
Forssman, Eninger, Tillman, Rodriguez, & Bohlin (2010)	"levels of commotion, routines and noise" (p. 288)	120 adolescents, ages 12-16	CHAOS	1. Children's Size-Ordering Task 2. 4-point scale of Diagnostic and Statistical Manual of Mental Disorders	1. CHAOS does not contribute its own separate variance to ADHD and ODD

				symptoms of ADHD and ODD 3. Response Tasks	
Gregory, Eley, O'Connor, Riksdijk, & Plomin (2005)	"family disorganization" (p. 1388)	6000 twin pairs, England and Wales	CHAOS, short-form	1. Edinburgh Postnatal Depression Scale 2. stressful life events 3. SES 4. sleep problems 5. Revised Rutter Parent Scale for Preschool Children	1. sleep (R^2 = .24 [males], R^2 = .47 [females]) 2. anxiety (R^2 = .41)
Hanscombe, Haworth, Davis, Jaffee, & Plomin (2011)	"environmental confusion and unpredictability of high levels of family chaos (i.e., noise, disorder and human traffic)" (p. 1212)	2,300 twin pairs, ages 9-12 (longitudinal), England and Wales	CHAOS, short-form	1. UK National Curriculum Criteria	1. R^2 = .63
Hart, Petrill, Deater-Deckard, & Thompson (2007)	"environmental confusion (i.e., noise, crowding, and traffic within the home" (p. 234)	272 twins, elementary age, western Pennsylvania and eastern Ohio	CHAOS	1. Stanford-Binet	R^2 = .02-.08
Kretschmer & Pike (2009)	"an environment that lacks routines and	118 families with children ages 4-8,	CHAOS, short form	1. Parental Discipline Interview 2. SES	1. sibling relationship quality (R^2 = -

regularity, and is disorganized and high in confusion” (p. 582)	Southern England	.23) 2. mother warmth with older sibling (R^2 = -.02) 3. mother warmth with younger sibling (R^2 = -.02) 4. mother discipline older sibling (R^2 = .01) 5. mother discipline younger sibling (R^2 = .001) 6. father warmth with older sibling (R^2 = -.04) 7. father warmth with younger sibling (R^2 = -.05) 8. father discipline older sibling (R^2 = .03) 9. father discipline

					younger sibling (R^2 = .01) 10. older sibling relationship (R^2 = -.01) 11. younger sibling relationship (R^2 = -.004)
Marcynyszyn, Evans, & Eckenrode (2008)	"family instability" (p. 380)	Study 1 1141 families, low SES Study 2 2339 families, low SES	1. intimate partner changes 2. work hour changes 3. residence changes 4. school transfer	1. Child Behavior Checklist 2. Teacher-Child Rating Scale 3. Academic Achievement 4. Criminal Justice Contacts 5. Substance Use 6. School Suspensions	1. externalizing problems (R^2 = .11) 2. internalizing problems (R^2 = .13) 3. academic grades (R^2 = .02-.05) 4. social skills, frustration tolerance, and task orientation (R^2 = .03-.08)
Martin, Razza, & Brooks-Gunn (2011)	"characterized by noise, crowding, and lack of routine and order " (p. 1)	882 preschoolers, representative of all Chicago neighborhoods	1. Family Instability 2. Lack of Routine 3. Television Generally On 4. Crowding 5. Noise	1. Peabody Picture Vocabulary Test 2. Child Behavior Checklist 3. Self-Regulation Inventory 4. HOME Inventory	1. receptive vocabulary (*d* = .07) 2. delayed gratification (*d* = .09)

Maxwell (2003)	"spatial density...amount of space per person" (p. 568)	73, 90% Black or African American, urban, elementary age	1. Classroom crowding 2. Ratio of rooms to people	1. Woodcock-Johnson Word Identification subtest 2. Lewis Feel Bad Scale	1. read five more words 2. scored 4 points higher on behavior measures
Milan, Pinderhughes, & The Conduct Problems Prevention Research Group (2006)	"family instability" (p. 44)	369, 60% of low SES	a rating of 0-7 based on: 1. the number of residential moves 2. deaths of family members 3. divorces or separations in the family 4. remarriage/re-entry to the home of a partner 5. parent-child separation on a temporary basis 6. addition of a new child into the family 7. parental job	1. Child Behavior Checklist 2. Diagnostic Interview Schedule for Children	1. externalizing problems (ΔR^2 = .02)

			changes experienced in the last year		
Mokrova, O'Brien, Calkins, & Keane (2010)	"an environment that is high in background noise and crowding and that is low in structural and temporal routine" (p. 120)	319 families, ages 7 and 10, 67% European American, 33% ethnic minority	CHAOS	1. Adult ADHD rating scale 2. Alabama Parenting Questionnaire 3. Coping with Children's Negative Emotions Scale 4. ADHD Rating Scale-IV: School Version	1. inconsistent discipline (ΔR^2 = .03) 2. Nonsupportive response to child negative emotions (ΔR^2 = 04)
Petrill, Pike, Prince, & Plomin (2004)	"the degree of organization and calm in the household versus chaos" (p. 446)	3,915 MZ twins, 3866, DZ twins, England and Wales	CHAOS	1. McArthur Communication Development Inventory 2. Parent Report of Children's Abilities	CHAOS score age 3 1. communication skills A. age 3 (R^2 =-.04) B. age 4 (R^2 = -.03) 2. cognitive ability A. age 3 (R^2 = -.05) B. age 4 (R^2 = -.04) CHAOS score age 4 1. communication skills A. age 3 (R^2 =

					-.04) B. age 4 (R^2 = -.03) *2.* cognitive ability A. age 3 (R^2 = -.04) *B.* age 4 (R^2 = -.05)
Raviv, Kessenich, & Morrison (2004)	"proximal environmental factors" (p.529)	1016, 3-year-old, 16% ethnic minorities	Home Observation for Measurement for the Environment (HOME)	1. parent-child interaction styles 2. Reynell Development Language Scales 3. Bracken Basic Concept Scale	1. Bracken Basic Concept Scale *(R^2* = .19) 2. Reynell Developmental Language Scale *(R^2* = .11)
Shamama-tus-Sabah, Gilani, & Wachs (2011)	"microsystem contexts such as the home, day care center, or school, which are characterized by high noise levels, high levels of density, or crowding, high context traffic patterns (many people coming and	203 children, primary grades, Pakistan	CHAOS	1. Behavior Assessment System for Children 2. 2. Raven's Standard Progressive Matrices	1. conduct problems (ΔR^2 = .09) 2. social skills (ΔR^2 = .03)

	going), and a lack of physical and temporal structure; few regularities or routines in the environment, little is scheduled nothing has its place" (p. 202)				
Smith, Prinz, Dumas, & Laughlin, (2001)	Organization (p. 968)	492, African American, longitudinal	Family Beliefs Inventory	1. Scale of Social Competence and School Adjustment 2. Academic competence measures	reading achievement (R^2 = .13) Parent Rating 1. social competence (R^2 = .15) 2. achievement competence (R^2 = .10) 3. communication effectiveness (R^2 = .08) 4. problem behavior (R^2 = .21) Teacher Rating 1. social competence (R^2 = .04)

					2. achievement competence (R^2 = .07) 3. communication effectiveness (R^2 = .06) 4. problem behavior (R^2 = .05)
Vernon-Feagans, Garrett Peters, Willoughby, Mills-Koonce, & the Family Life Project Key Investigators (2012)	"a system of frenetic activity, lack of structure, unpredictability in everyday activities, and high levels of ambient stimulation" (p. 339)	1,123 families from rural PA and NC, preschool	Chaos Disorganization and Instability	1. Wechsler Primary Preschool Inventory, Receptive Vocabulary 2. Preschool Language Skills	1. expressive language (R^2 = .25) 2. receptive language (R^2 = .23)
Wachs (1993)	"noisy homes, crowded homes, or homes with a high traffic pattern (people coming and going in the home)" (p.82)	56 toddlers , Caucasian, parents from varying education levels	PHSI	1. Observation of caregiver attentiveness and responsibility	2. investment (R^2 = .07) 3. verbal responsiveness (R^2 = .03)

Wachs & Camli (1991)	"physical environment…high levels of ambient background noise or home crowding" (p. 251)	35 infants, Indiana, 83% of primary caregivers had completed college	PHSI	1. Parenting questionnaires	Regular Mealtime 1. marital satisfaction (R^2 = .13) 2. Parental support (R^2 = .13) 3. child care satisfaction (R^2 = .11) Regular Naptime 1. child care satisfaction (R^2 = .12)

In addition to noise level, overcrowding also has an impact on physiological outcomes. For example, participants who experienced household crowding had higher skin conductance and higher rates of illness (Evans et al., 1991). Chen, Cohen, and Miller (2010) observed higher levels of cortisol for individuals whose parents indicated higher scores on the Confusion, Hubbub, and Order Scale (CHAOS; Matheny et al., 1995). Further, CHAOS scores also have demonstrated relationships with total sleep time, sleep hygiene, sleep problems, and sleep onset latency in both pre-schoolers and adolescents (Billows, Gradisar, Dohnt, Johnston, McCapin, & Hudson, 2009; Gregory, Eley, O'Connor, Rijsdijk, & Plomin, 2005). Additionally, males exposed to chronic overcrowding experienced higher blood pressure than their counterparts who were not exposed to chronic overcrowding (Evans et al., 1998). This evidence supports the theory that there are physiological changes in individuals when exposed to household chaos, overcrowding and chronic noise levels.

Chaos and socioemotional problems. In addition to negative physiological outcomes, household chaos relates to adjustment problems, social competence difficulties, and other internalizing problems (e.g., Adams, 2004; Evans, 2006, Evans, Gonnella, Marcynyszyn, Gentile, & Salpekar, 2005; McLoyd, 1990). The number of moves children or adolescents experienced, which reflects instability of the home, positively correlated with the number of adjustment problems they had (Adams, 2004; Evans, 2006). In addition, higher scores on the CHAOS related to lower levels of social competence (Dumas et al., 2005; Evans, 2006). For example, Dumas et al. found that mother and child pairs with higher CHAOS scores had more difficulty completing a toy-sorting task in which they were required to work together to solve problems. Higher CHAOS scores also related to poorer quality of sibling relationships (Kretschmer & Pike, 2009).

Beside social competency, household chaos variables also have related to internalizing problems. For example, a higher level of instability, as characterized by the number of moves experienced throughout childhood, was related to greater internalizing problems (Adam & Chase-Lansdale, 2002). A lack of routine was associated with lower scores on a delayed gratification task (Martin, Razza, & Brooks-Gunn, 2011). Also, on a household measure of home density, students from homes with a higher level of home density were more likely to "feel badly" on the Lewis Feel Bad Scale (Maxwell, 2003). Further, Evans, Lepore, et al. (1998) found that chronic overcrowding was positively associated with learned helplessness in girls. Higher rates of learned helplessness were observed across both genders upon exposure to chronic noise (Evans, 2006; Evans et al., 2005; Evans et al., 1991).

Household chaos variables were also associated with psychological variables. Evans (2006) found that crowding conditions related to withdrawal and elevated levels of neuroticism. Children who experienced higher levels of home chaos also exhibited higher levels of psychological distress and lower teacher-rated self-regulatory behavior (Evans et al., 2005). Further, participants who experienced more household crowding were more likely to demonstrate an external locus for control (Evans et al., 1991). Finally, based on her review of the empirical literature, Adams (2004) discussed the link between adolescent suicide and the number of living locations and separations from parent figures.

The household chaos variable of family disruption also related to child internalizing behaviors, externalizing behaviors, and psychological diagnoses (Milan, Pinderhughes, & The Conduct Problems Prevention Research Group [CPPRG], 2006). They found most children only experienced one family disruption event per year. However, for those children who did experience family disruption, the more instability experienced, the greater the negative effects were predicted. Growth curve modeling indicated that family instability predicted externalizing behaviors more strongly than internalizing behavior, and participants who experienced co-morbid disorders had higher levels of family instability. These results demonstrate the link not only between family disruption and internalizing problems but also externalizing problems.

Chaos and behavior. In addition to socioemotional problems, household chaos also has related to behavior difficulties. Bada et al. (2008) established a link between early living arrangements and behavior outcomes in that both adaptive functioning and problem behaviors are positively associated with the number of early disruptions. Deater-Deckard, Mulluineaux, Beckman, Petrill, Schatschneider, and Thompson (2009) found that higher CHAOS scores were associated with higher levels of disruptive behavior. In a sample of African American elementary students, higher levels of child behavior problems were reported, based on ratings by both parents and teachers, when their household CHAOS score was higher (Dumas et al., 2005). Household chaos predicted negative behavior outcomes above parenting style (Adams, 2004; Coldwell, Pike, & Dunn, 2006).

In addition to general externalizing problems, household chaos has been linked to specific problem behaviors. In a sample of adolescents, family instability was related to higher number of cigarettes smoked, lower frustration tolerance, higher risk for school suspension, more binge drinking, more classroom disruptive behavior, and more marijuana use (Marcynyszyn, Evans, & Eckenrode, 2008). Further, the number of times participants were separated from their parental figures was related to externalizing problems and greater teen sexual activity (Adam & Chase-Lansdale, 2002). Higher noise levels, chronic crowding, greater TV-use, and higher CHAOS scores related to higher levels of aggression (Dumas et al., 2005; Evans, 2006; Martin et al., 2011). CHAOS scores also predicted children exhibiting callous-unemotional traits and conduct problems (Fontaine, McCrory, Boivin, Moffitt, & Viding, 2011). Higher CHAOS scores and greater TV-use also were linked to attention problems (Dumas et al., Forssman, Eninger, Tillman, Rodriguez, & Bohlin, 2010; Martin et al., 2011). Further, higher scores on the CHAOS related to a greater number of ADHD symptoms in children ages 7 and 10 (Mokrova, O'Brien, Calkins, & Keane, 2010). Finally, household crowding, housing, and neighborhood quality contributed to higher rates of juvenile delinquency (Evans, 2006; Evans et al., 1991). The impact of household chaos also has been demonstrated in international samples. For example, higher CHAOS scores for a sample of children from Pakistan were related to higher internalizing, externalizing, and adaptive behaviors (Shamama-tus-Sabah, Gilani, & Wachs, 2011).

Chaos and cognitive, academic, and language skills. Household chaos was also related to cognitive, academic, and language difficulties (Deater-Deckard et al., 2009; Raviv, Kessenich, & Morrison, 2004). Exposure to toxins such as lead, mercury, and PBCs has been associated with poor cognitive outcomes (e.g., Evans, 2006). Further, Deater-Deckard et al. (2009) found that household chaos was inversely correlated with IQ scores. Additionally, Hart, Petrill, Deater-Deckard, and Thompson (2007) examined whether or not SES and household chaos mediated cognitive development. Results indicated that CHAOS scores were a significant predictor of cognitive ability above and beyond SES. Evans et al. (1995) found differences between children exposed to chronic noise and those that were not exposed to chronic noise such that children who lived in the high noise area made more errors than their counterparts in quiet neighborhoods. However, chronically higher noise levels led to modifications in cognitive strategies to accommodate for the noise (Evans et al., 1991).

There are further links between household chaos and cognitive outcomes. For example, Petrill, Pike, Prince, and Plomin (2004) found that SES and CHAOS scores mediated the relationship between verbal ability and parent perceived ability. The participants in this study were taken from the Twins Early Development Study (TEDS; Trouton, Spinath, & Plomin, 2002). Parents were administered questionnaires about their child's perceived verbal and nonverbal cognitive abilities. Results indicated that household chaos was a significant mediator of both verbal and nonverbal scores such that greater chaos related to lower scores. While this link is important, it is possible that the use of parent report of cognitive ability could have resulted in data that were not accurate.

In addition to cognitive differences, household chaos variables, such as higher levels of disorganization, are associated with poorer academic outcomes (Evans, 2006). Specifically, Marcynyszyn et al. (2008) found that a higher level of family instability was related to lower math and reading achievement scores. Higher scores on the short form of the CHAOS also were correlated with poorer performance on UK National Curriculum criteria (Hanscombe, Haworth, Davis, Jaffee, & Plomin, 2011). In addition, the number of times that children were separated from their parents was related to lower grades (Adam & Chase-Lansdale, 2002), and higher noise levels were associated with reading difficulties and long-term memory deficits (Evans, 2006). Finally, housing and neighborhood quality have been related to lower academic competency and lower standardized test scores (Evans, 2006).

Household chaos also has been associated with language development difficulties. Raviv et al. (2004) found a relationship between the physical environment and early language skills as well as comprehension skills. Household crowding was a statistically significant predictor of Woodcock-Johnson Word Identification subtest scores (Maxwell, 2003), and lack of routine was correlated with lower receptive vocabulary (Martin et al., 2011). Additionally, Vernon-Feagans, Garrett Peters, Willoughby, et al. (2012) found that household chaos, as measured by the Chaos, Disorganization, and Instability scale, accounted for variance in early language development.

Higher levels of household chaos can have a negative impact on the physical, psychological, behavior, cognitive, academic, and language outcomes of children. For example, differences in CHAOS scores accounted for small to medium amounts of variance in IQ, academic, and behavioral scores. The effects tended to be larger for externalizing and internalizing behaviors than for cognitive outcomes. Some effects were small while others were moderate for language outcomes. Because there were negative outcomes associated with household chaos, it was important to develop an effective measure of the construct that is useful for a variety of populations. However, most studies to date have focused on one aspect of chaos (e.g., Bada et al., 2008; Maxwell, 2003) or chaos at the classroom or neighborhood levels rather than household level (e.g., Evans et al., 1998; Maxwell & Evans, 2000). In addition, these studies have focused primarily on urban populations. In response to these limitations, the present study investigated psychometric properties of the CHAOS, the measure used most frequently in the empirical literature to date. The present study used a large (N = 1,292) rural sample and focused on chaos at the household level.

How Chaos Potentially Impacts Child Development

Parenting and language development. Higher levels of crowding and traffic patterns have been related to impaired parent-child interactions, including lower levels of parent involvement, verbal stimulation, and responsivity (Evans, Maxwell, & Hart, 1999; Wachs & Camli, 1991). Specifically, parents of children in crowded homes were not as responsive to their children as parents in homes with fewer people. Additionally, less diverse language was used in more crowded homes. Household chaos appeared to impact language acquisition and development because it affected the way that parents interacted with their children (Evans et al., 1999). Furthermore, higher levels of these components of chaos were associated with parents displaying lower levels of giving or demonstrating objects to their children (Wachs & Camili, 1991). Fewer interactions between parents and children impacted the ability of children to self-regulate (Evans, 2006). In summary, it appeared that elevated home chaos does not allow parents to be as attentive and offer as much verbal stimulation to their children, which may contribute to poorer language development**.** Because the home environment affects the quality of parent-child interactions, it is important to study characteristics of the home environment, such as household chaos (Raviv et al., 2004).

Parenting behaviors. Chaos also was related to parenting behaviors (Coldwell et al., 2006; Corapci & Wachs, 2002; Wachs, 1993). Mokrova et al. (2010), for example, established a link between parental ADHD and the amount of household chaos families experience. Additionally, Corapci and Wachs established the link between household chaos, parenting efficacy, and parent mood state. Their data supported a three-factor structure of chaos and efficacy (crowding, noise-chaos, and efficacy). Higher household chaos related to lower quality of parenting behavior. Further, higher scores on both noise and crowding were correlated with less responsive and stimulating parenting. High levels of noise, crowding, and traffic patterns were associated with lower caregiver attentiveness and responsibility (Wachs, 1993).

Coldwell et al. (2006) demonstrated that families who reported lower levels of chaos indicated lower levels of parent-child negativity, higher levels of parent-child positivity, and higher levels warmth/enjoyment with both parents. Conversely, higher household chaos scores were related to higher levels of anger/hostility toward parents. They also found the presence of high chaos moderated the relationships between dad-child positivity and negativity. Dumas et al. (2005) found that higher scores on the CHAOS led to higher scores of dysfunctional discipline practices.

There was a link between parental stress and the support parents provided for their children. Wahler and Dumas (1989) reviewed the relationship between parent-child interactions and environmental stressors. As a result of their review, they hypothesized that mothers who have a lot of environmental stressors are less able to effectively respond to their children. Additionally, they identified all of the systems acting on the parent that may influence their interactions with their children such as their spouse, friends, and boss. Chaos also interacted with family stress and affected the level of supportive parenting provided to children (Nelson, O'Brien, Blankson, Calkins, Keane, 2009). From a more positive perspective, Smith, Prinz, Dumas, and Laughlin (2001) demonstrated that family support and organization were related to parent- and teacher-reported behavior competence and positive reading outcomes.

Poverty and parenting. While some studies (e.g., Dumas et al., 2005; Vernon-Feagans, Garrett-Peters, DeMarco, et al., in press) have demonstrated that environmental chaos contributes to variance in parenting styles and child outcomes beyond poverty, Evans et al. (2005) found children in poverty experienced more home chaos than their middle- or high-income counterparts. Given these conflicting findings, it is important to consider the effects poverty has on parenting.

McLoyd (1990) conducted a literature review examining characteristics of African American families living in poverty and how poverty interacted with a number of other factors to contribute to psychological distress. Their model proposed that poverty and economic loss can lead to psychological distress, marital conflict, and parenting difficulties, which in turn can lead to child socioemotional problems. Protective factors for parent psychological distress and parenting styles included parent appraisal, personality, financial resources, and community/social support. These resources may be useful in informing potential interventions. The authors concluded that psychological distress mediated the relationship between poverty and quality of parenting.

Neighborhood. Neighborhood characteristics, such as overcrowding and chronic noise, also have been associated with negative childhood outcomes. Dumas et al. (2005) found there were lower positive and higher negative neighborhood attributes when there was higher home disorganization. Because neighborhood attributes were related to CHAOS scores and childhood outcomes, it was important to examine how sociocultural factors contribute to neighborhood characteristics. Sampson, Morenoff, and Earls (1999) used survey data from 8,782 Chicago residents to examine factors that result in efficacy for children. One such mechanism was social control. Results included bivariate correlation coefficients of characteristics across three domains: intergenerational closure, reciprocated exchange, and child-centered social control. They found that these factors were related to factors of overcrowding such as population density as well as poverty. Results also indicated that affluence and poverty were concentrated to certain neighborhoods. Neighborhoods with people of color tended to have higher population densities than neighborhoods primarily with people who are Caucasian. These results indicated that there may be racial/ethnic differences in the amount of environmental chaos resulting from crowding.

Interacting Mechanisms. In addition to noise and overcrowding, other environmental factors can influence child outcomes. For example, Adam (2004) found that family instability contributed to a loss of trust, feelings of abandonment, and lack of social support. Further, family instability led to a disruption in activities, routines, environment, and parenting quality.

Additionally, Smith et al. (1997) used a sample of 201 fourth grade students and their families to examine ecological factors such as parent involvement at school, parent involvement at home, importance of education, attitude toward involvement, teacher provision of involvement opportunities, barriers, school climate, neighborhood climate, and family demographic variables. They developed a model that described how these factors were related to each other within the sample. The key relationship was a path that traveled through parent education level, family structure, income, neighborhood climate, school climate, teacher provision of involvement, and parent involvement at both home and school. School climate also related to attitudes towards involvement and barriers. These attitudes can impact parent involvement at home and at school. While the study expanded on how environmental characteristics related to parent involvement, more direct measurement techniques are needed, as these data were self-report.

Conclusions. While the research is inconclusive about the manner in which household chaos impacts the outcomes of children, it is clear that the construct has a significant impact on childhood outcomes. Therefore, it is important to develop accurate measures of the construct. Further, psychometric evidence for scores from the CHAOS has been mixed (Dumas et al., 2005). The proposed study examined psychometric properties of the CHAOS with a diverse rural sample drawn from two states.

Measuring Household Chaos

Although studies (e.g., Deater-Deckard et al., 2009; Dumas et al., 2005; Petrill et al., 2004) have demonstrated the importance of measuring chaos, current measures still warrant additional investigation. Table 2 provides information about measures of household chaos used in the literature including the domains assessed, the rating format, and the psychometric evidence for each measure.

Table 2 *Measures Used to Assess Chaos Across Reviewed Studies*

Chaos Measure	**Domains Assessed**	**Rating Format**	**Psychometric Evidence**
Chaotic Living Conditions: Cumulative Index of Residential Noise (Brown & Low, 2008)	Crowding-two or more people per bedroom Hours of TV-four or more hours of television while child is home Family instability-change since birth in caregiver's residential partner	Participants given a score of 0 to 3 based on the number of domains present	Predictive validity with child adjustment (values not reported) the index correlates with high levels of noise, overcrowding, and family instability (values not reported no other psychometric evidence was reported (Brown & Low, 2008)
Purdue Home Stimulation Inventory (Wachs & Chan, 1986)	Routines crowding decorations toys changes in toys stimulation level of toys	32 items, 9 asked to primary caregiver, 23 observed by data collector	inter-rater reliability (.85) concurrent and predictive validity with cognitive development (Wachs & Chan, 1986).
Chaotic Home System (Flouri, 2009)	number of people/ room child not living with the same parent on a continuous basis number of household moves	Parent interview	family cohesiveness scale (Cronbach's α level = .63, Flouri, 2009) no other psychometric evidence reported

	untidiness ranging from "over-tidy" to "chaotic" on 5-pt. Likert scale family cohesiveness from "rarely ever or never" to "often"—7 items on a 3-point Likert scale		
Chaos Disorganization and Instability (Vernon-Feagans, Garrett-Peters, Willoughby et al., 2012)	total number moves total number of changes in primary caregiver total number of changes in secondary caregiver total number of different people in the household number of times household members moved into or out of the household average number of hours TV on each day average household density home visit preparation cleanliness	Parent interview, data collector observation	no psychometric evidence established yet (Vernon-Feagans, Garrett-Peters, Willoughby, et al., 2012)

	neighborhood noise level		
5 Measures of Chaos (Martin et al., 2011)	*family instability* (change in caregiver, maternal relationship, or residence) *lack of routine* (mealtime frequency, rules, and rule enforcement) *Television generally on* *Crowding* (space-to-people and space to furniture ratio) *Noise* (noise level of home from inside and outside)	Parent interview, data collector observation	No psychometric data were reported (Martin, Razza,, & Brooks-Gunn, 2011)
Confusion, Hubbub, and Order Scale (CHAOS, Matheny et al., 1995)	routines and organization disorganization, confusion, and noise	15-item dichotomously scored questions, asked to participants through interviews	Matheny et al. (1995) Item-total correlations (.22 - .64) Internal consistency (Cronbach's α = .79) Test-retest stability over 1-year period (.74) Construct validity with Purdue Home Stimulation Inventory (PHSI) (R^2= .39) Dumas et al. (2005) Cronbach's α (.83)
Confusion, Hubbub, and Order Scale (CHAOS)-Spanish Translation (Haack et al., 2011)	routines and organization disorganization, confusion, and noise	15-item dichotomously scored questions, asked to participants through questionnaires	Internal consistency (.79) Convergent Validity A. PSI-SF (r = .50, p < .001)

			B. CBCL/6-18 (*r* = .40, *p* < .001) Predictive Validity-distinguishes between levels of externalizing problems (X^2 = 6.64, *p* < .01)
Shortened CHAOS (Deter-Deckard et al., 2009)	routines and organization disorganization, confusion, and noise	6-item version of full CHAOS	two factor structure: degree of household quiet and degree of household order-values not reported (Johnson, Martin, Brooks-Dunn, & Petrill, 2008)

Vernon-Feagans, Garrett-Peters, DeMarco, et al. (2012) identified several measures of household chaos that have been used in the empirical literature: (a) Chaotic Living Conditions: Cumulative Index of Residential Noise (Brown & Low, 2008), (b) Purdue Home Stimulation Inventory (PHSI; Wachs & Chan, 1986), (c) Chaotic Home System (Flouri, 2009), (d) Chaos Disorganization and Instability (Vernon-Feagans, Garrett-Peters, Willoughby, et al., 2012), (e) CHAOS: Confusion, Hubbub, and Order Scale (Matheny et al., 1995), and (f) Shortened CHAOS (Deter-Deckard et al., 2009). Further, Martin et al. (2011) used five household attributes to assess chaos. In addition to these specific measures, other studies have used elements such as noise, disorganization, moves, and overcrowding as markers of chaos that lead to poor child development outcomes (Evans, 2006).

The Chaotic Living Conditions: Cumulative Index of Residential Noise was a measure of home chaos used by Brown and Low (2008). The measure considered several conditions: (a) “two or more people per bedroom”, (b) “four or more hours per day of television while child is home”, and (c) “family instability (change since child’s birth in caregiver’s residential partner)” (p. 921). If a condition was present in the home environment, a score of 1 was assigned. Scores across all conditions were added together to create an index of home chaos. Brown and Low noted predictive validity with child adjustment and the index correlates with high levels of noise, overcrowding, and family instability; however, the authors did not report the values of these relationships. While this scale captured a few aspects of household chaos, the items omitted several key features such as family routines and organization. Therefore, using this scale to assess home chaos may have been insufficient to detect different amounts of household chaos and how such differences contributed to child outcomes.

The PHSI consisted of 32 items. Nine items were asked to the primary caregiver (e.g., number of people in the home and presence of a regular suppertime). Items that were observed (n = 23) included decorations in the child's room, if there were audio-visual responsive toys, and if there were changes in the child's toys. Some evidence for reliability and validity of scores from this measure have been published to date. For example, Wachs and Chan (1986) reported an inter-rater reliability coefficient of .85 and observed concurrent and predictive relationships with cognitive development. While this scale encompassed many aspects of household chaos, it was an observational method that required a significant amount of time and training to use. In addition, it may not have reflected the family's perception of what it was like day-to-day in their home.

The Chaotic Home System (Flouri, 2009) assessed five elements of home chaos including the number of people per room, whether or not the child was living with the same parent on a continuous basis, the number of household moves, untidiness of the home, and family cohesiveness. Items included how often the family went for walks together, went for outings together, had meals together, and went shopping together. The authors reported that the family cohesiveness 7-item survey achieved a Cronbach's α of .63; however, psychometric properties for the other scales were not reported. Also, though there was more than one aspect of chaos being assessed, the measure did not capture organization/disorganization on a daily basis.

The Chaos Disorganization and Instability (Vernon-Feagans, Garrett-Peters, Willoughby, et al., 2012) was a new measure designed to assess 10 objective indicators of household chaos. "Instability" items included (1) *the total number of times the child moved* physically to another residence, (2) *the total number of changes in the primary caregiver* (usually involves change in primary responsibility for child from mother to other adult), (3) *the total number of changes in* *the secondary caregiver* (either primary caregiver's partner or primary caregiver's mother), (4) *the total number of different people in the household*, and (5) *the total number of times household members moved into or out of the household*. Household disorganization also consisted of five dimensions including: (6) *the average number of hours that the TV was on each day*, (7) the *average household density*, (8) home visitor ratings of *home visit preparation by the household*, (9) home visitor ratings of *the cleanliness of the household*, and (10) home visitor ratings of *the neighborhood noise level around the home"* (p. 346). The psychometric properties of this measure have not been examined yet, and though this scale used objective measure of home chaos, it still did not capture the self-perceptions of household chaos on families/participants.

Martin et al. (2011) used family instability, lack of routine, pervasive television use, crowding, and noise to examine chaos. Family instability reflected changes in the child's residence and adult figures in the participants' life. Lack of routine included three items related to meal times, rules, and enforcement of family rules. Whether or not the TV was generally on somewhere in the home was reported by the mothers' of participants. Crowding was based on people- and furniture-to-space ratio. Noise was based on the amount of noise inside and outside the home. While this approach captured multiple elements of household chaos, it did not capture day-to-day instability. Further, there was no psychometric evidence to date.

Confusion, Hubbub, and Order Scale (CHAOS)

One of the earliest measures of chaos, which continues to be used in many research studies was developed by Matheny et al. (1995). They defined environmental confusion as aspects of the physical environment that are "stressful" such as "noise, crowding, and environmental traffic patterns" (p. 430). Because methods such as direct observation were costly and time consuming, Matheny et al. decided to develop a brief rating scale – the CHAOS - to assess this construct. The CHAOS includes 15 dichotomously-scored items intended to measure two facets of home chaos. The first encompassed routines and organization. The second reflected disorganization, confusion, and noise. Research has indicated that chaos is a distinct construct that can be differentiated from social or psychological constructs such as SES (Baker et al., 2003; Dumas et al. 2005; McLoyd, 1990; Sampson et al., 1999). Since its publication, the CHAOS has been used in multiple studies in the U.S. (Vernon-Feagans, Garrett-Peters, DeMarco, et al., 2012) and abroad (Haack, Gerdes, Schneider, & Hurtado, 2011; Shamama-tus-Sabah et al., 2011). Unlike other measures of home chaos, this scale conceptualized the day-to-day perception of chaos experienced by families/participants. Additionally, because it included 15 dichotomously scored items, it was easy to administer. While the scale had many positive attributes, one weakness was that there had been minimal published evidence regarding the psychometric properties of its scores.

While multiple studies (e.g., Dumas et al., 2005; Vernon-Feagans, Garrett-Peters, Willoughby, et al., 2012) have used the CHAOS as a way to capture environmental or home chaos, few researchers have examined the psychometric properties of the scale. The authors of the CHAOS, Matheny et al. (1995), examined internal consistency and test-retest stability. Based on data from a sample of 123 families, item-total correlations ranged from .22 - .64, and internal consistency (Cronbach's α) was .79. Test-retest stability over a 1-year period was .74. Matheny et al. used the Purdue Home Stimulation Inventory (PHSI) to assess evidence of construct validity between the CHAOS and direct observation (R^2= .39). They also found that lower SES and parent education were significantly related to higher CHAOS scores.

Similarly, Dumas et al. (2005) examined psychometric properties of the CHAOS with data drawn from a sample of 106 mother-child dyads. Specifically, they observed a high Cronbach's α (.83), and, after controlling for sociodemographic variables, parenting stress, and child's IQ, found that CHAOS made a significant contribution to the prediction of child performance on problem-solving tasks.

The CHAOS also was translated into other languages, including Spanish and Urdi (Haack et al., 2011; Shamama-tus-Sabah et al., 2011). Haack et al. examined the psychometric properties of the Spanish version of the scale. Specifically, internal consistency was acceptable (.79) and convergent validity coefficients with measures of parent stress (PSI-SF, r = .50) and problem behavior (CBCL, r = .40) were moderate. They also found that the scale could predict clinically significant externalizing problems. Finally, Johnson, Martin, Brooks-Dunn, and Petrill (2008) used an abbreviated (6-item) version of the scale in their research. This abbreviated scale also expanded the original dichotomous response options to a 5-point Likert scale. Factor analysis on the brief scale resulted in two factors: degree of household quiet and degree of household order. The current study, however, used the full 15-item scale and used the dichotomous metric.

Rationale, Purpose, and Hypotheses for the Proposed Study

While results from initial studies of the CHAOS were promising, further independent investigation of its psychometric properties was warranted. My dissertation examined questions related to the reliability and validity of scores from the CHAOS. Specifically, multi-group confirmatory factor analyses were conducted to examine the internal structure of the measure. Further, I examined internal consistency of the CHAOS, as well as inter-rater reliability between primary and secondary caregiver scores on the CHAOS.

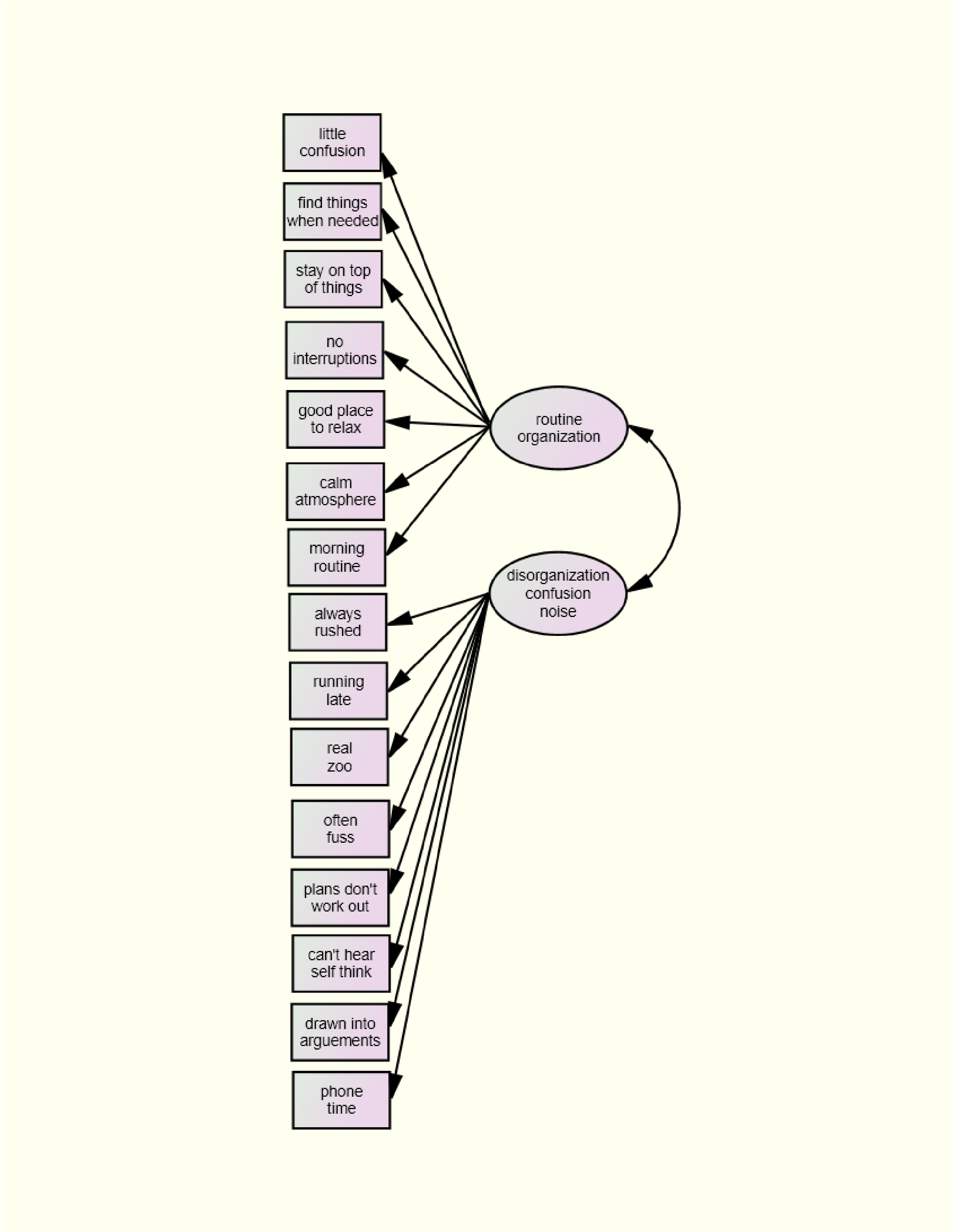
Research also has suggested that pervasive chaos in the home environment contributes to language, social, emotional, and behavioral difficulties (e.g., Vernon-Feagans, Garrett-Peters, Willoughby, et al., 2012). The studies that demonstrated these outcomes defined and conceptualized household chaos in multiple ways. However, the definition of chaos by Matheny et al. (1995) was used in the current context because their measure of this construct was the primary focus of this study. Specifically, Matheny et al. conceptualized household chaos as "environmental confusion" (p.430) and list its characteristics as "noise, crowding, and environmental traffic patterns" (p.430)

Because previous studies of chaos emphasized its pervasiveness, whether or not CHAOS scores were stable over time for the families in the study was examined. Additionally, how the CHAOS related to other key home variables was analyzed. Key home variables examined in the current study included variables like secondary caregiver presence in the home, ratings of the neighborhood noise level around the home, and the amount of time the TV is turned on in the home.

In addition to examining a broader array of psychometric properties, the current study also had a larger sample than previous psychometric studies of CHAOS. Finally, this study focused on a rural sample, which complemented the urban samples featured in the work by Dumas et al. (2005) and Johnson et al. (2008).

There were six specific hypotheses tested as part of this study. The first hypothesis was that the best-fitting structural model would reflect 15-items loading on two domains (Figure 1; Matheny et al., 1995). Second, it was predicted that the factorial structure was equivalent across groups. The third hypothesis was that the internal consistency of CHAOS scales would be moderate to high. Fourth, there would be a high correlation between primary and secondary raters on the CHAOS. Fifth, CHAOS scores demonstrated moderate to high stability over time. Sixth, there would be a moderate relationship between CHAOS scores and observation measures of home characteristics (e.g., average household density, ratings of neighborhood safety). Finally, SES, income, presence or absence of a secondary caregiver in the home, and primary caregiver employment outside of the home were expected to positively relate to CHAOS scores.

Figure 1. Two-factor model of the CHAOS (Matheny et al., 2005).



CHAPTER 2: METHOD

Participants

The Family Life Project (FLP; Vernon-Feagans, Cox, & the Family Life Project Key Investigators, in press) is a study of rural family life. Data collected in the study examine a number of variables that affect the development of children including cognitive, social, and emotional factors. There are 1,292 families who have participated in the study since their children were born in 2004. Families live in three rural counties in central Pennsylvania and three rural counties in eastern North Carolina. Stratified random sampling was used to obtain an overrepresentation of children and their families who live below the poverty line. Additionally, stratified random sampling was used to obtain an overrepresentation of children who are African American.

Data were collected in the homes of all participating families when their child was 2-, 6-, 15-, 24-, 35-, 48-, and 58-months old. The data used in the current study were collected at the ages of 35- and 48-months. Once the children reached school age, visits were conducted at their school, and a home visit was conducted in the spring of first grade. Data were collected by graduate students and full time employees of FLP. Data collectors were certified on all measures, and fidelity checks were conducted during each data collection period. Measures reflected cognitive, biological, social, and behavioral domains. Participants were from North Carolina (60.0%) and Pennsylvania (40.0%). When the study began, 100% of primary caregivers were females. Primary caregivers identified as 58.1% white, 41.4% black, .1% American Indian, .1% Chinese, .2% other Asian, and .1% Guamanian or Chamorro. Demographic data at the time points used in the current study are reported by income group in Table 3.

Table 3 *Distribution of Primary Caregiver Demographic Characteristics by Income Group*

	Low Income (*n* =754)	Middle/ High Income (*n* =414)
State		
North Carolina	66.5	47.2
Pennsylvania	33.5	52.8
Race		
Black	52.5	21.0
White	47.2	78.0
Other[a]	.1	1.0
Education Level (35-Month Visit)		
Less Than High School or GED	20.2	3.6
GED, high school graduation, some additional training	72.5	46.6
Associate's Degree or more	7.3	50.0
Education Level (48-Month Visit)		
Less than high school or GED	17.6	2.7
GED, high school graduation, some additional training	74.0	45.9
Associate's Degree or more	8.4	51.4

Note: [a]Other indicates American Indian, Chinese, other Asian, and Guamanian or Chamorro.

Measures

Confusion, Hubbub, and Order Scale (CHAOS). The CHAOS included 15 dichotomously-scored items that assess household routines, organization, confusion, and noise (see Appendix A). Sample items included "There is very little commotion in our house" and "We can usually find things when we need them." A total chaos score was calculated for each family, with higher scores indicating higher levels of household chaos. (Several items were reverse-scored.) As noted earlier, previous studies examined internal consistency, test-retest, and construct validity with the PHSI (Dumas et al., 2005; Matheny et al., 1995); however, further investigation was warranted to determine if scores from the CHAOS demonstrate sound psychometric properties.

Income Related Information. Income-to-needs ratio was used to examine a family's SES. An income-to-needs ratio "expresses a family's income as a proportion of the official federal poverty line for a family of that size" (p. 2, Acs & Gallagher, 2000). An income-to-needs ratio of 1.0 meant the family was at the federal poverty line. An income-to-needs ratio of 2.0 or below indicated that the family was of low income or "working poor."

Household Characteristics. Factors that were reflective of the people that comprise a household, as well as changes in the physical dwelling were considered in the analyses for the current study. The consistency of the person identified as the primary caregiver, whether the primary caregiver lived with his/her romantic partner, whether a secondary caregiver lived in the household, and the duration of the primary caregiver's relationship with his/her romantic partner were all asked of the primary caregiver during the home interview. People were considered to be living in the household if they "slept in the home at least three nights a week." Additionally, the number of moves the child experienced was also captured by indicating the number of moves the family underwent since the previous visit.

Primary Caregiver Employment. Changes in the primary caregiver's employment situation were used in the current study. The number of "big work changes" such as "starting or ending a job, a changed in responsibility or a change in hours" and the number of jobs the primary caregiver has stopped since the previous visit was also asked out loud by data collectors during the home interview.

Noise. The Windshield Survey is a survey used to assess household characteristics, neighborhood safety and noise level, and the validity of the FLP home interview. Items were taken from the Post-Visit Inventory (Dodge et al., 1990) used in the Fast Track Project. The current study assessed the noisiness of the participants' neighborhood using an item from the Windshield Survey. The item was phrased "The noise level in this neighborhood around this dwelling is," and was scored on a nominal scale (*Can't rate* = 0, *Very Quiet* = 1, *Average* = 2, *Noisy* = 3, *Very Noisy* = 4). Additionally, the number of hours of TV use per week was part of the primary caregiver questionnaire.

Procedure

Two home visitors traveled to each of the families to collect data in their homes. Each visit required approximately 3 hours. Home visitors included part-time graduate student assistants and full-time data collectors. Before home visitors were able to conduct home visits, they participated in rigorous training and certification on all measures. Each visit cycle (e.g., 35- and 48-month interviews) lasted 8-12 months, and each family was visited once per cycle.

The visits were comprised of interviews, questionnaires, activities for the child and parent to complete together, and activities for the child to complete alone. The home interview occurred at the start of the home visit. The items from the CHAOS were asked of the primary caregiver at the 48-month visit, and asked of both the primary and secondary caregivers at the 35-month visit. Additionally, home visitors asked about neighborhood noise level as part of the Windshield Survey. Information from the parent interviews, questionnaires, and Windshield questions were used in the current study.

Design and Data Analysis

Confirmatory Factor Analysis (CFA). CFAs were conducted to examine the structural validity of the CHAOS. CFAs were used to test a priori models of the relationships between variables and latent structures. Multi-group analyses were conducted using low and middle/high income groups. The authors' two-factor model (Figure 1) was tested with the initial sample along with alternative models (see Figures 2, 3, & 4) to determine which model demonstrated the best fit. After the overall models were tested, the similarities and differences between groups were examined.

Before these analyses were completed, data were checked to ensure they met underlying assumptions, specifically, normality, linearity, and multicollinearity. The assumption of normality was assessed using visual inspection of frequency histograms and P-P plots, and values of skew (+/- 2) and kurtosis (+/- 7) were used. Linearity was assessed through visual inspection of scatter plots. To check for outliers, the Mahalanobis distance test was used. Multicollinearity was considered by examining correlation matrices, tolerance levels, and variance inflation factors. Correlations had to be less than .90, tolerance levels greater than .10, and variance inflation factor less than 10, for the assumption of no multicollinearity to be met (Field, 2009; Kline, 2005).

Analysis of Moment Structures (AMOS) 19.0 (Student Version) software was used to estimate models. The Maximum Likelihood method was used to test all models. Kline (2005) suggested a minimum of four fit indices to evaluate model fit: model chi-squared, root mean squared error of approximation (RMSEA), Bentler's comparative fit index (CFI), and standardized mean squared error residuals (SRMR). The model chi-squared statistic examined the overall fit of the model. The higher the value for chi-squared, the worse the fit of the model was for the data. Chi-squared tests assumed large sample sizes and multivariate normality. The RMSEA was considered the "badness of fit" of the model (p. 138; Kline, 2005) because the smaller the value, the better the fit of the model. It favored simpler models, and the further the model fit deviated from the null hypothesis, the larger the values became. The CFI examined the improvement in goodness of fit relative to the baseline model (i.e., the index examined the relationship between the data and the proposed model), and values greater than .90 were considered indicative of good fit. The standardized mean squared error residuals (SRMR) examined the standardized difference between observed and predicted residuals. The closer the value was to 0, the better the model fit. SRMR values below .10 are generally accepted as an indication of good fit (Kline).

There are several other indices that were considered in addition to the four indices recommended by Kline (2005). The GFI examined the variance and covariance jointly explained in the model, and values higher than .95 are generally considered a good fit. The AIC and BCC were predictive fit indices that examined how the observed covariance matrix differed from the predicted covariance matrix; therefore, the model favors larger populations. These indices favored simpler models, and lower values indicated better fit.

Figure 2. Three-factor model of the CHAOS

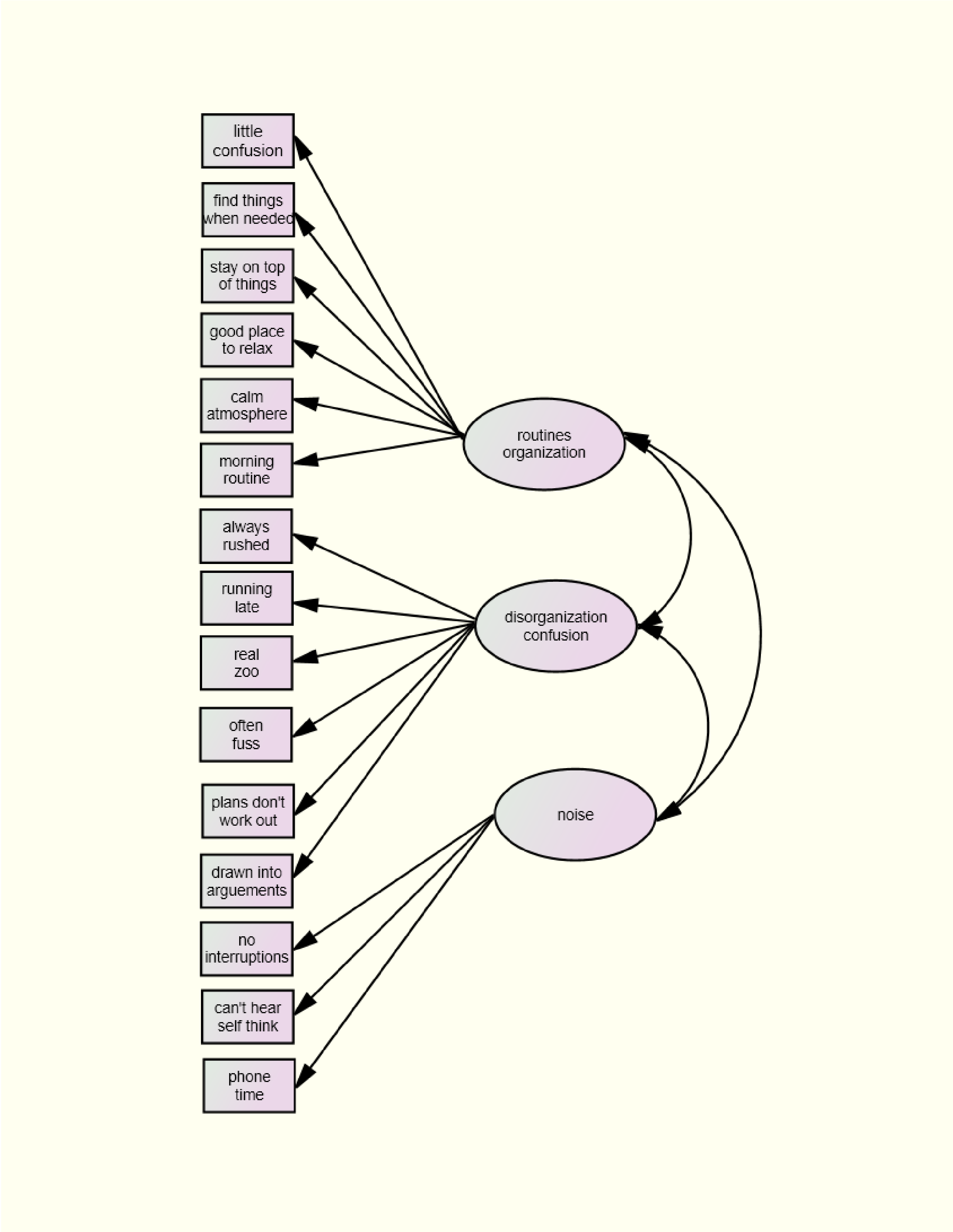


Figure 3. One-factor factor model of the CHAOS.

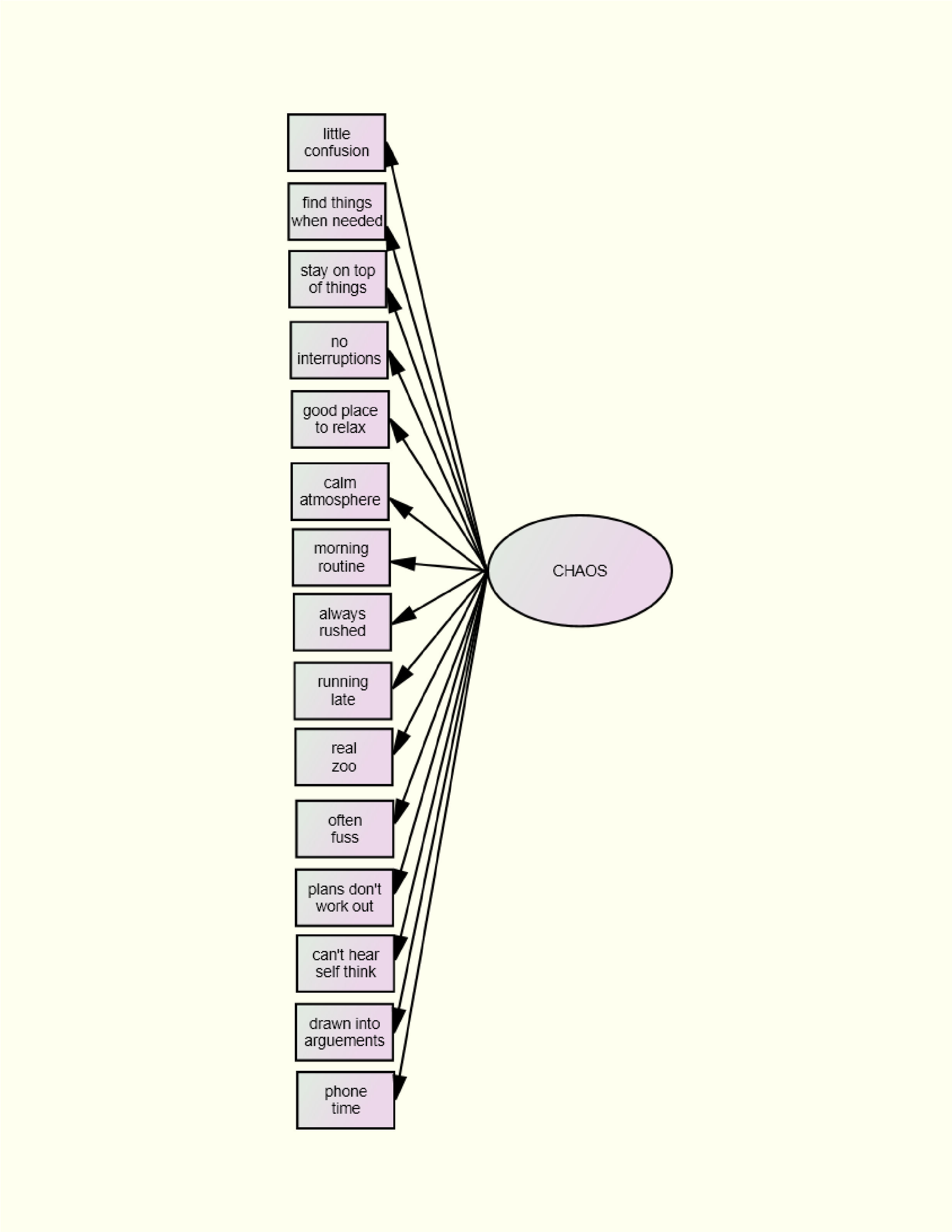
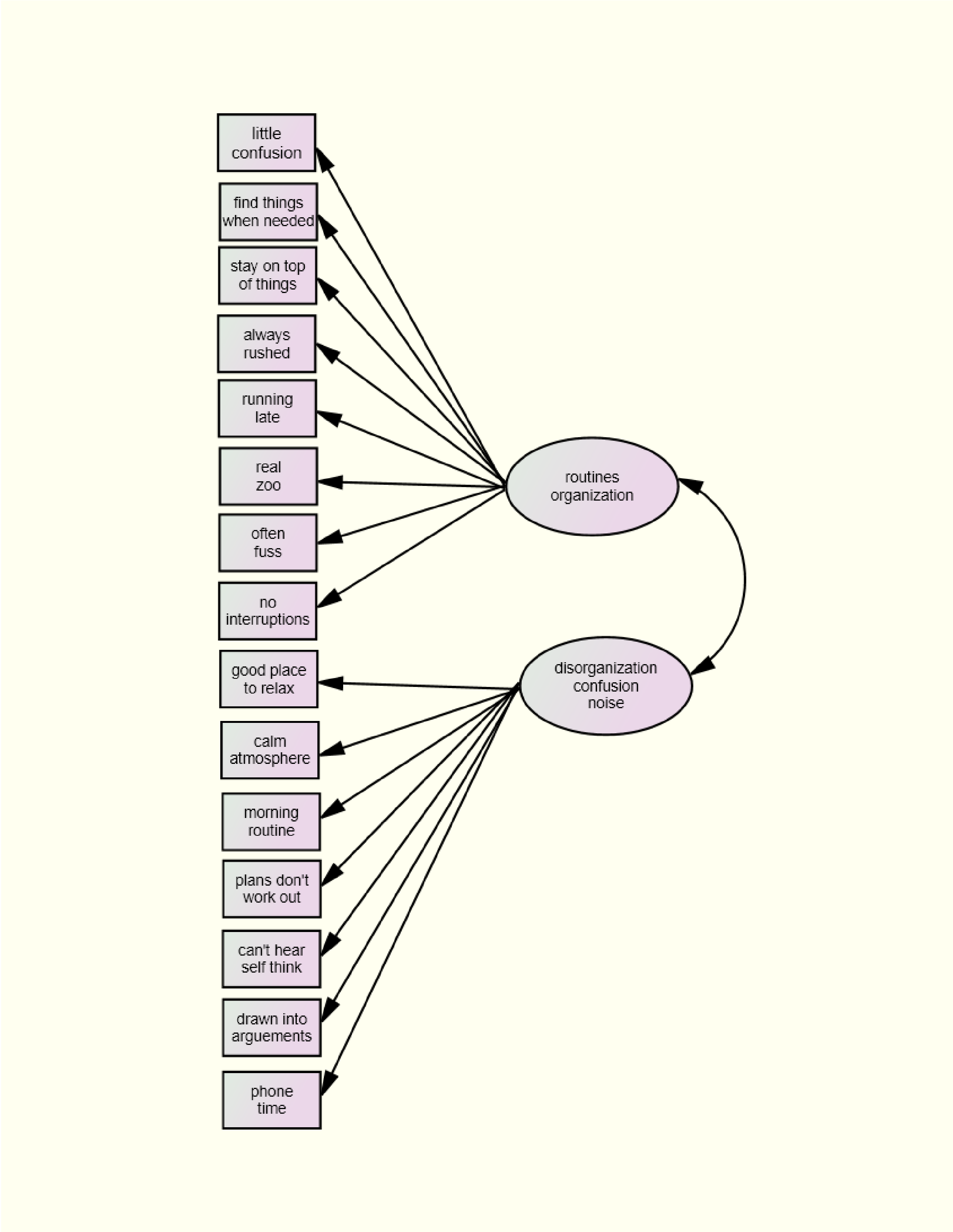


Figure 4. Atheoretical model of the CHAOS.



The NFI and TLI examined the goodness of fit of the model relative to the null model. For these indices, values greater than .95 were considered a good fit, and values above .90 were considered an acceptable fit (Hu & Bentler, 1999). The NFI is better for larger sample sizes, so it was a good choice for the current sample (Kline, 2005). Because the sample in the proposed study was large, using the NFI, AIC, BCC, and Chi-square was appropriate. Based on recommendations by Kline (2005), other indices include RMSEA, CFI, SRMR, GFI, and TLI.

The procedure used in the current study followed the 3-step process for multi-group analysis outlined by Byrne (2010). First, the model (e.g., one-factor, two-factor) was run for both groups (i.e., low income, middle/high income) individually. For each group, model fit and modification indices were examined. Modification indices that were high for each group were each added to the model and fit indices for each of the resulting models were examined. The model that best fit the data across both groups was selected for the next set of analyses. Then, the model was tested using a combined sample, and the resulting model was referred to as the configural or baseline model. The final step was to test for structural and measurement invariance. This process involved restricting and unrestricting parameters to determine which items and factor covariances were equivalent (invariant) across groups. To determine equivalence, chi-squared differences and differences in CFI were calculated.

Reliability Analyses. Because not all participants had a second rater, 64% of the sample (N=743) were included in the reliability analyses. Of those cases, 68.5% were the romantic partners of the primary caregiver. The remainders were someone else living in the home who played a significant role in providing care for the child, usually the maternal grandmother. The Kuder-Richardson Formula-20 was used to assess the internal consistency of items on the CHAOS. Cohen's Kappa coefficients and intra-class correlations were used to assess the inter-rater reliability of scores between primary and secondary caregivers at the 35-month visit. R^2 was used to assess practical significance. For reliability, Sattler (2001) recommended the following coefficient standards: .80 or higher for clinical decision making, .70-.79 as relatively reliable, .60-.69 as marginally reliable, and below .60 as unreliable. When evaluating Kappa coefficients, Landis and Koch (1977) suggested the following criteria: < 0 (no agreement), 0 - .20 (slight agreement), .21-.40 (fair agreement), .41 - .60 (moderate agreement), .61 - .80 (substantial agreement), and > .81 (perfect agreement). Fleiss (1981) used .75 as excellent agreement, .40-.75 as far to good agreement, and .40 as poor agreement for assessing Kappa coefficients.

Analysis of Score Change over Time. To examine whether or not individual CHAOS scores vary across time points relative to SES, a repeated measures 2 x 2 ANOVA was used to compare scores at different time points (35- and 48-months) and income levels (low and middle/high) using SPSS software. Previous researchers had not examined CHAOS scores across time points. Assumptions such as sphericity (using Mauchaly's test), normality, homogeneity of variance, and linearity were tested. A .05 statistical significance level was used, and w^2 was used to assess for practical significance. Field (2009) indicated .01 as a small effect size, .06 as a medium effect size, and .14 as a large effect size when using w^2.

CHAOS, Environmental, and Familial Variables. In addition, regression models were used to examine how income-to-needs ratio, noisiness of the neighborhood, number of moves since the last visit, whether primary caregiver was the same since the last visit, whether or not the primary caregiver was living with his or her romantic partner, whether there was a secondary caregiver living in the household, relationship duration, number of work changes, number of jobs stopped since last visit, and number of hours TV was on per week related to CHAOS scores. A significance level of .05 was used to assess statistical significance. The following values were used to interpret correlation coefficients: less than .10 (small), .10 - .50 (medium), greater than or equal to .50 (large), and squared correlations (R^2) were evaluated using the following values: less than .09 (low), greater than or equal to .09 and less than .25 (moderate), and greater than or equal to .25 (large, Cohen, 1992). The assumptions of normality, linearity, and homogeneity of variance were assessed. The following equation was tested: $Y_i = (b_0 + b_1X_{i1} + b_2X_{i2} + b_3X_{i3 \ldots}$ $b_7X_{i10}) + \varepsilon_i$

Y_i = CHAOS score

X_{i1} = income-to-needs ratio

X_{i2} = noisiness of the neighborhood

X_{i3} = number of moves since the last visit

X_{i4} = primary caregiver same as previous visit

X_{i5} = living with partner

X_{i6} = presence or absence of a secondary caregiver

X_{i7} = relationship duration

X_{i8} = number of work changes

X_{i9} = number of jobs stopped since last visit

X_{i10} = number of hours the TV is on per week

The income variable (X_{i1}) was continuous income-to-needs ratio. Noisiness of the neighborhood (X_{i2}) utilized a 4-point interval scale. The number of moves since the last visit was a whole number (X_{i3}). Whether or not the primary caregiver was the same as the previous visit (X_{i4}) was dummy-coded as 0 (no) and 1 (yes). Whether or not the primary caregiver was living with his or her partner (X_{i5}) was dummy-coded as 0 (no) and 1 (yes). Presence or absence of a secondary caregiver (X_{i6}) was dummy-coded as 0 (no secondary caregiver) and 1 (secondary caregiver present). The relationship duration (X_{i7}) was the amount of time the primary caregiver was romantically involved with his/her romantic partner. The number of work changes (X_{i8}) was based on the number of times the primary caregiver started or ended a job, had a change in hours, or had a change in responsibilities. The number of jobs stopped since last visit (X_{i9}) was a whole number. The number of hours of TV per week (X_{i9}) was based on primary caregiver self-report.

Chapter 3: Results

Structural Validity

Descriptive Statistics and Testing of Assumptions. Before testing assumptions for the primary analyses, missing data were identified and examined. Specifically, 5.9% and 9.2% of the data were missing from the 35-month and 48-month data collection periods, respectively. Roth and Switzer (1999) purported that if less than 10% of data were missing and the sample was large, multiple ways of working with missing data are permissible (e.g., deletion or imputation). For the current study, the Estimation Maximization Algorithm (EM) was used. Descriptive statistics are reported in Table 4. Skew and kurtosis values of subtest raw scores were calculated to determine whether they met the assumption of normality. Kurtosis values were less than 7, and skew values were almost all less than 2, suggesting the data approximated univariate normality (Field, 2009). In addition, given the large sample size in the current study, any violations of the normality assumption likely had negligible impact on the analyses (Hair, Black, Babin, Anderson, & Tatham, 2005). Based on a visual inspection of scatterplots between subtest raw scores, the assumption of linearity appeared to be met.

Mahalanobis' distances were calculated to determine the presence of any outliers, and four potential outliers were identified. No meaningful differences were found between the results of the analyses with and without the outliers; thus, all data were included in the analyses (Hair et al., 2005). Multi-group confirmatory factor analyses were run using low and middle/high income groups. Inter-item correlations are reported in Table 5.

Table 4 *Means, Standard Deviations, Skew, and Kurtosis Values for CHAOS Scores*

Item	*M*	*SD*	Skew	Kurtosis
1. There is very little commotion in our house	.43	.48	.28	-1.88
2. We can usually find things when we need them	.87	.33	-.2.23	3.06
3. We almost always seem to be rushed	.47	.49	.11	-1.94
4. We are usually able to stay on top of things	.82	.37	-1.75	1.12
5. No matter how hard we try, we always seem to be running late	.41	.48	.38	-1.82
6. It's a real zoo in our home	.28	.44	.98	-.99
7. At home we can talk to each other without being interrupted	.51	.49	-.05	-1.96
8. There is often a fuss going on at our home	.30	.45	.90	-1.14
9. No matter what our family plans, it usually doesn't seem to work out	.18	.37	1.72	1.04
10. You can't hear yourself think in our home	.21	.40	1.43	.10
11. I often get drawn into other people's arguments at home	.21	.40	1.44	.14
12. Our home is a good place to relax	.75	.42	-1.18	-.55
13. The telephone takes up a lot of our time at home	.16	.36	1.90	1.70
14. The atmosphere in our home is calm	.66	.46	-.70	-1.47
15. First thing in the day, we have a regular routine at home	.81	.38	-1.60	.64

Table 5 *Inter-item Correlations by Income Group*

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. Little commotion	-	.20	.18	.13	.12	.32	.24	.32	.05	.11	.17	.28	.02	.32	.04
2. Find things when needed	.16	-	.22	.33	.19	.28	.17	.23	.22	.13	.14	.19	.09	.19	.11
3. Always seem rushed	.28	.16	-	.27	.48	.30	.23	.25	.21	.21	.22	.26	.16	.25	.03
4. Able to stay on top of things	.12	.18	.20	-	.32	.28	.13	.15	.19	.11	.12	.25	.07	.22	.09
5. Always running late	.03	.19	.50	.24	-	.30	.18	.20	.29	.21	.15	.20	.15	.14	.05
6. Real zoo	.38	.20	.26	.30	.23	-	.25	.47	.26	.35	.38	.32	.16	.32	.02
7. Talk without interruption	.27	.12	.23	.13	.09	.20	-	.29	.22	.17	.22	.32	.04	.30	.07
8. Often a fuss at home	.35	.16	.32	.20	.18	.46	.27	-	.28	.26	.38	.35	.09	.27	.07
9. Plans don’t seem to work out	.03	.21	.18	.19	.15	.11	.06	.12	-	.27	.30	.24	.17	.09	.08
10. Can’t hear self think	.21	.21	.20	.13	.18	.34	.16	.29	.13	-	.28	.25	.11	.17	.01
11 Drawn into arguments	.10	.14	.13	.0	.09	.10	.16	.17	.20	.18	-	.25	.18	.19	.09
12. Good place to relax	.26	.05	.28	.15	.15	.33	.27	.45	.14	.24	.23	-	.12	.43	.17
13. Telephone takes up time	.00	.14	.11	.06	.10	.04	.05	.01	.09	.09	.16	.01	-	.04	.06
14. Atmosphere is calm	.37	.23	.35	.19	.19	.39	.29	.47	.10	.18	.18	.49	.08	-	.16
15. Regular routine	.09	.20	.07	.14	.19	.08	.04	.06	.01	.06	.06	.08	.08	.17	-

Note: Values below the diagonal represent correlations for the middle/high income group and values above the diagonal represent the low income group.

Analysis. Four confirmatory factor analyses were performed on the covariance matrix. All analyses were multi-group analyses for the purpose of examining fit across low and middle/high income subgroups. In the one-factor model, all items were indicators of one chaos factor (Figure 3). For the two-factor model, Routine/Organization and Disorganization/Confusion/Noise (Figure 1) were the two latent factors. In the three-factor model, items related to noise were loaded onto a separate factor. (The three-factor model, however, did not yield a positive definite matrix, and therefore was not considered further.) The atheoretical model forced unrelated items onto 2 factors (Figure 4). In each of the multi-factor models, the factors were correlated. For all models, the latent variables were scaled by setting their variance to 1.00 (Albright & Park, 2009). The one-factor model was determined to be identified based on the three-indicator rule in that there were at least three variables (items) on one latent factor, each variable only loaded on one latent factor, and measurement errors were uncorrelated. The two-factor model was determined to be identified based on the two-indicator rule in that there were at least two correlated constructs and two measures per latent factor (Kline, 2005).

One-Factor Model. For the one-factor model, all items loaded onto one latent factor identified as household chaos. Fit indices of the model indicated the SRMR, RMSEA, and GFI demonstrated adequate fit (Table 6). The other indices (i.e., AIC, BCC, CFI, NFI, and TLI), however, did not indicate a good fit for the data. Modifications were incorporated to see if they improved overall model fit. Factor loadings are reported in Table 7 and ranged from .15 - .67 for the low-income group and .10 - .61 for the middle/high income group. The SRMR, RMSEA, and GFI improved slightly as a result of the modifications (Table 8). The CFI, NFI, and TLI did not indicate a good fit for the data in the initial model, but improved as modifications were added

Table 6 *Fit Indices for Alternative CHAOS Models*

	χ^2	DF	SRMR	RMSEA	GFI	CFI	NFI	TLI	AIC	BCC
One-Factor	779.79	180	.06	.05	.91	.80	.76	.72	899.79	903.53
Two-Factor	701.39	178	.05	.05	.92	.83	.78	.79	825.39	829.26
Atheoretical	764.35	178	.06	.05	.91	.80	.76	.77	888.35	892.22

Note. All χ^2 values are statistically significant at $p < .01$.

Table 7 *Factor Loadings for Each Model by Income Group*

	Low Income			Middle/High Income		
	One-Factor	Two-Factor	Atheoretical	One-Factor	Two-Factor	Atheoretical
1. Little commotion	.42	.46	.42	.50	.51	.33
2. Find things when needed	.41	.41	.43	.33	.35	.51
3. Always seem rushed	.51	.51	.52	.53	.54	.54
4. Able to stay on top of things	.41	.42	.42	.35	.35	.36
5. Always running late	.46	.47	.47	.36	.38	.37
6. Real zoo	.67	.69	.67	.62	.62	.62
7. Talk without interruption	.46	.49	.45	.41	.42	.41
8. Fuss going on at home	.60	.62	.60	.67	.68	.69
9. Plans don’t seem to work out	.44	.46	.45	.23	.24	.24
10. Can’t hear self think	.44	.46	.45	.43	.44	.42
11. Drawn into arguments	.50	.53	.52	.28	.28	.29
12. Good place to relax	.58	.65	.61	.60	.61	.63
13. Telephone takes up time	.22	.24	.23	.10	.11	.11
14. Atmosphere is calm	.50	.59	.51	.67	.70	.70
15. Regular routine	.15	.19	.17	.18	.19	.19

Table 8 *Fit Indices for One-Factor CHAOS Model and Subsequent Modifications*

Modification	χ^2	DF	SRMR	RMSEA	GFI	CFI	NFI	TLI	AIC	BCC
1) 1-Factor	779.83	180	.06	.05	.91	.80	.76	.77	899.83	903.58
2) e3 ⇔ e5	609.59	178	.05	.05	.93	.86	.81	.83	733.59	737.45
3) e12 ⇔ e14	559.84	176	.05	.04	.94	.87	.83	.85	687.84	691.83
4) e2 ⇔ e4	509.54	174	.05	.04	.94	.89	.84	.86	641.64	645.76
5) e1 ⇔ e9	477.82	172	.05	.04	.95	.90	.85	.88	613.82	618.06

suggesting that the model became a better fit compared to the null, and the correlations improved within the data as modification indices were added.

After examining the configural model, multi-group equivalence was assessed. Fit indices for the one-factor model are reported by group in Table 9. Examining the fit indices of the model indicated that a revised Model 1 with the errors for Items 3 and 5 co-varied should be used as the baseline model. Even though adding multiple error covariances could lead to an even better fit, it is important to only add modifications that can be supported theoretically (Byrne, 2010). In this case, Items 3 and 5 both tapped into the concept of running late and rushing. While this model only achieved adequate fit based on a priori criteria, it was used for multi-group analysis as it demonstrated the best fit of the models tested.

Fit indices of the configural model are reported in Table 10. Tests of equivalence when all factors were constrained (i.e., set to be equivalent across groups) indicated that the two models were not equivalent across all items (Table 10). Tests of equivalence included change in chi-squared and difference in CFI. For the difference in CFI, a .01 cut- score was recommended (Byrne, 2010). Tests of equivalence indicated that only Items 10 and 13 were equivalent for this model (Table 10).

Internal Consistency. The Kuder-Richardson Formula-20 (KR-20) was used to assess the internal consistency for the best-fitting one-factor model across groups (α = .79 [overall]) and by group (.81 [low income], .76 [middle/high income]). Item-total correlations are reported in Table 11 and ranged from .17 - .59 for the low-income group and .13 - .53 for the middle/high income group.

Table 9 *Fit Indices for One-Factor Models by Income Group*

Model	χ^2	df	CFI	RMSEA	RMSEA 90% CI	ECVI
Low Income						
1. Hypothesized one-factor model	470.57	90	.81	.08	.07, .08	.71
2. Model 1 with One Error Covariance Specified (Items 3, 5)	376.75	89	.85	.07	.06, .07	.58
3. Model 2 with One Error Covariance Specified (Items 1, 5)	462.81	89	.82	.08	.07, .08	.70
4. Model 3 with One Error Covariance Specified (Items 12, 14)	427.84	89	.83	.07	.06, .08	.65
Middle/High Income						
1. Hypothesized one-factor Model	309.18	90	.79	.08	.07, .09	.89
2. Model 1 with One Error Covariance Specified (Items 3, 5)	232.75	89	.86	.06	.05, .07	.74
3. Model 2 with One Error Covariance Specified (Items 1, 5)	290.51	89	.80	.07	.07, .08	.85
4. Model 3 with One Error Covariance Specified (Items 12, 14)	296.67	89	.80	.08	.07, .09	.87

Table 10 *Summary of Goodness-of-Fit Statistics for Tests of Multi-group Invariance*

Model Description	Equality Constraints Imposed	χ^2	df	$\Delta\chi^2$	Δdf	*p*	CFI	ΔCFI
1. Configural Model	None	609.57	178	--	--	--	.86	--
2. Model A	All	685.24	192	75.68	14	< .001	.84	.02
Model B	Item 1	638.77	179	29.20	1	< .001	.85	.01
Model C	Item 2	615.50	179	5.94	1	< .05	.85	.00
Model D	Item 3	634.78	179	25.21	1	< .001	.85	.01
Model E	Item 4	616.50	179	6.94	1	< .01	.85	.00
Model F	Item 5	616.57	179	7.01	1	< .01	.85	.00
Model G	Item 6	631.68	179	22.12	1	< .01	.85	.01
Model H	Item 7	624.52	179	14.96	1	< .001	.85	.00
Model I	Item 8	643.23	179	33.67	1	< .001	.84	.01
Model J	Item 9	623.04	179	13.47	1	< .001	.85	.00
Model K	Item 10	611.74	179	2.17	1	NS	.86	.00
Model L	Items 10, 12	640.05	180	30.48	2	< .001	.85	.01
Model M	Items 10, 13	611.39	180	1.18	2	NS	.86	.00
Model N	Items 10, 13, 14	669.28	181	59.72	3	< .001	.84	.02
Model O	Items 10, 13, 15	614.37	181	4.80	3	< .05	.86	.00

Note: Each model is compared to the configural model.

Table 11 *Pearson Correlations of Items with Total CHAOS Scores and Latent Construct Scores by Income Group*

	Low Income			Middle/High Income		
	CHAOS Total Score	Routine Organization	Disorganization Confusion Noise	CHAOS Total Score	Routine Organization	Disorganization Confusion Noise
1. Little commotion	.35	.35		.35	.39	
2. Find things when needed	.41	.34		.34	.28	
4. Able to stay on top of things	.36	.31		.34	.27	
7. Talk without interruption	.41	.36		.32	.34	
12. Good place to relax	.54	.50		.46	.42	
14. Atmosphere is calm	.47	.49		.53	.55	
15. Regular routine	.19	.17		.24	.19	
3. Always seem rushed	.49		.45	.47		.48
5. Always running late	.45		.44	.38		.40
6. Real zoo	.59		.54	.49		.42
8. Fuss going on at home	.55		.46	.52		.42
9. Plans don’t seem to work out	.42		.42	.23		.25
10. Can’t hear self think	.43		.40	.39		.37
11. Drawn into arguments	.44		.45	.27		.24
13. Telephone takes up time	.17		.24	.13		.14

Two-Factor Model. For the two-factor model, SRMR, RMSEA, and GFI demonstrated adequate fit (Table 6). Factor loadings are reported in Table 8 and ranged from .19 - .69 for the low-income group and .11 - .70 for the middle/high income group. Modifications were incorporated into the model to see if they improved fit. All indices improved slightly as modifications were added to the model (Table 12).

After examining the configural model, multi-group equivalence was assessed along with subsequent model modifications. Fit indices for the two-factor model are reported by group in Table 13. As with both one-factor models, fit indices indicated that Model 1 with the errors for Items 3 and 5 co-varied should be used as the baseline model. While this model only demonstrated adequate fit, it was used for multi-group analysis as it demonstrates the best fit of the models tested. Fit indices of the configural model are reported in Table 14.

Tests of equivalence, when all factors were constrained (i.e., set to be equivalent across groups), indicated that the two models were not equivalent across all items (Table 14). The test of equivalence, when all items in Factor 2 were constrained as equal, demonstrated that the items in Factor 2 were not equivalent across groups. Test of equivalence, when all items in Factor 1 were constrained as equal, demonstrated that the items in Factor 1 were not equivalent across groups. Tests of equivalence at the item level indicated that nine items were equivalent across groups: Items 1, 2, 5, 6, 7, 8, 10, 12, and 15. Results of the tests of equivalence are reported in Table 15. Structural equivalence between factors was also examined; however, results indicated that the covariance between the two factors were not equivalent across groups (Table 14).

Table 12 *Fit Indices for Matheny et al. (1995) CHAOS Two-Factor Model and Subsequent Modifications*

	Fit Indices									
Modification	χ^2	DF	SRMR	RMSEA	GFI	CFI	NFI	TLI	AIC	BCC
1) CHAOS (Matheny et al., 1995)	701.39	178	.05	.05	.92	.83	.78	.79	825.40	829.26
2) e3 ⇔ e5	534.70	176	.05	.04	.94	.88	.83	.86	662.70	666.69
3) e2 ⇔ e4	483.45	174	.05	.04	.95	.90	.85	.88	615.45	619.57
4) e4 ⇔ e5	449.05	172	.05	.04	.95	.91	.86	.89	585.05	589.29
5) e1 ⇔ e9	427.39	170	.04	.04	.95	.91	.87	.89	567.39	5 71.75
6) e14 ⇔ e9	405.56	168	.04	.04	.96	.92	.84	.90	549.56	544.05
7) e5 ⇔ e9	392.71	166	.04	.03	.97	.92	.88	.90	540.71	545.32
8) 4 → 3	374.16	164	.04	.03	.96	.93	.88	.91	526.16	530.90
9) 1 to 13	361.31	162	.04	.03	.96	.93	.89	.91	517.31	522.17

Table 13 *Fit Indices for Two-Factor Model to Determine a Baseline Model*

Model	χ^2	df	CFI	RMSEA	RMSEA 90% CI	ECVI	SRMR
Low Income							
1. Hypothesized two-factor model (Matheny et al., 1995)	395.94	89	.84	.07	.06, .08	.61	.05
2. Model 1 with One Error Covariance Specified (Items 3, 5)	303.46	88	.89	.06	.05, .06	.49	.05
3. Model 2 with One Error Covariance Specified (Items 1, 5)	362.74	88	.86	.06	.06, .07	.57	.05
4. Model 3 with One Error Covariance Specified (Items 12, 14)	363.47	88	.86	.06	.06, .07	.59	.05
Middle/High Income							
1. Hypothesized two-factor Model (Matheny et al.)	305.38	89	.79	.08	.07, .09	.89	.06
2. Model 1 with One Error Covariance Specified (Items 3, 5)	231.12	88	.86	.06	.05, .07	.72	.06
3. Model 2 with One Error Covariance Specified (Items 1, 5)	296.59	88	.80	.08	.07, .09	.88	.06
4. Model 3 with One Error Covariance Specified (Items 12, 14)	289.08	88	.80	.07	.07, .08	.86	.06

Table 14 *Summary of Goodness-of-Fit Statistics for Tests of Multi-group Two-Factor Model Invariance*

Model Description	Equality Constraints Imposed	χ^2	df	$\Delta\chi^2$	Δdf	*p*	CFI	ΔCFI
1. Configural Model	None	534.65	176	--	--	--	.88	--
2. Model A	All	615.90	191				.86	
Model B	Factor 2	600.66	184	66.00	8	< .001	.86	.02
Model C	Factor 1	553.61	183	18.95	7	< .01	.88	.00
Model D	Item 1	535.14	177	.49	1	NS	.88	.00
Model E	Items 1, 2	540.45	178	5.80	2	NS	.88	.00
Model F	Items 1, 2, 4	544.75	179	10.09	3	< .05	.88	.00
Model G	Items 1, 2, 7	541.67	179	7.01	3	NS	.88	.00
Model H	Items 1, 2, 7, 12	542.70	180	8.05	4	NS	.88	.00
Model I	Items 1, 2, 7, 12, 14	548.92	181	14.27	5	< .05	.88	.00
Model J	Items 1, 2, 7, 12, 15	543.00	181	8.34	5	NS	.88	.00
Model K	Items 1, 2, 7, 12, 15, 3	543.75	182	9.09	6	NS	.88	.00
Model L	Items 1, 2, 7, 12, 15, 3, 5	547.39	183	12.73	7	NS	.88	.00
Model M	Items 1, 2, 7, 12, 15, 3, 5, 6	548.81	184	14.16	8	NS	.88	.00
Model N	Items 1, 2, 7, 12, 15, 3, 5, 6, 8	550.51	185	15.86	9	NS	.88	.00
Model O	Items 1, 2, 7, 12, 15, 3, 5, 6, 8, 9	585.12	186	50.46	10	< .001	.87	.01
Model P	Items 1, 2, 7, 12, 15, 3, 5, 6, 8, 10	552.13	186	17,48	10	NS	.88	.00
Model Q	Items 1, 2, 7, 12, 15, 3, 5, 6, 8, 10, 11	572.23	187	37.58	11	< .001	.87	.01
Model R	Items 1, 2, 7, 12, 15, 3, 5, 6, 8, 10, 13	555.97	187	21.32	11	< .05	.88	.00
3. Structural Model	Covariance between two latent factors	632.06	192	97.40	16	< .001	.85	.03

Note: Each model is compared to the configural model.

Internal Consistency The KR-20 was used to assess the internal consistency for the two factors across groups. For the Routines and Organization subscale, α = .65 and .62 for the low and middle/high income groups, respectively. For the Disorganization, Confusion, and Noise subscale, α levels were .75 and .64 for low and middle/high income groups. Item-total correlations with latent factors are reported in Table 11 and ranged from .17 - .50 (Routines and Organization) and .24 - .54 (Disorganization, Confusion, and Noise) for the low income group and .19 - .55 (Routines and Organization) and .14 - .48 (Disorganization, Confusion, and Noise) for the middle/high income group.

Atheoretical Model. For the atheoretical model, items were arbitrarily loaded onto two factors. Fit indices of the model indicate the SRMR, RMSEA, and GFI demonstrated an adequate fit (Table 6). Factor loadings are reported in Table 7 and ranged from .17 - .67 for the low-income group and .11 - .70 for the middle/high income group. Because the model did not provide adequate fit, factor loadings were low, and factors were not theoretically supported, no modifications were made to the model.

Inter-rater Reliability of 15-Item CHAOS Scores

Cohen's Kappa coefficients, intraclass correlations, and percent exact agreement were used to assess the inter-rater reliability between the primary and secondary caregivers at the 35-month time point. All values are reported in Table 15. Kappa coefficients ranged from .06 - .16. Intraclass correlations ranged from .08 - .44, and percent exact agreement ranged from 61.9 - 82.0. The overall percent exact agreement for the scale was 70.8.

Table 15 *Reliability Indices (N=743)*

Item	Kappa Coeffiecient	Intraclass Correlation	% Exact Agreement
1. Little commotion	.16	.43	66.1
2. Find things when needed	.14	.42	82.0
3. Always seem rushed	.09	.24	59.6
4. Able to stay on top of things	.11	.14	77.9
5. Always running late	.16	.44	67.2
6. Real zoo	.13	.37	70.3
7. Talk without interruption	.14	.35	61.9
8. Fuss going on at home	.12	.33	69.7
9. Plans don't seem to work out	.07	.05	75.6
10. Can't hear self think	.07	.08	70.5
11. Drawn into arguments	.10	.25	73.2
12. Good place to relax	.06	.18	71.5
13. Telephone takes up time	.13	.43	80.5
14. Atmosphere is calm	.08	.24	63.5
15. Regular routine	.07	.21	72.9

Variations in 15-Item CHAOS Scores by SES Across Time Points

To examine the stability of CHAOS scores across time points relative to SES, a repeated measures factorial 2 (time: 35- and 48-months) x 2 (income levels: low income and middle/high income) ANOVA was run using SPSS software. The correlation between scores across time was .66. Because the repeated measures variables used in the model had only two levels, the assumption of sphericity was met (Field, 2009). Results indicated statistically significant difference in CHAOS scores across time points, F (1, 1166) = 43.87, p < .05, η^2 = .036, such that CHAOS scores were higher at the 35-month visit (M = 4.36, SD = 3.19) compared to the 48-month visit (M = 3.82, SD = 3.15). There was not a statistically or practically significant difference in CHAOS scores across time points based on income level, F (1, 1166) = .68, p > .05, η^2 = .001. The main effect for income level was also not statistically significant, F (1, 1166) = 2.13, p > .05, η^2 = .002.

CHAOS, Environmental, and Familial Variables

Descriptive Statistics and Assumptions. Prior to conducting the proposed regression analyses, assumptions were checked. Mean, standard deviations, skew and kurtosis values for the variables are presented in Table 16. Even though some values indicated non-normality these were dichotomous items and most of the sample answered the item the same way (e.g., whether or not the child was living with their original primary caregiver). Further, normality was also examined using histograms superimposed with normal curves. Visual inspection indicated that the values approximated normality. All data were independent observations. Linearity was demonstrated through the examination of residual plots. Tolerance values ranged between .74 and .98. Because these values were greater than .1 and VIF values are below 10, there is no evidence of multicollinearity. Pearson correlations for chaos-related variables are listed in Table 17.

Table 16 *Mean, Standard Deviation, Skew, and Kurtosis for Familial and Environmental Variables*

	M	*SD*	*Skew*	*Kurtosis*
Income to Needs Ratio	1.87	1.56	2.13	8.49
CHAOS Sum Score	4.36	3.19	.74	-.02
Noisiness of Neighborhood	1.88	.62	.68	1.76
Number of moves since last visit	.48	.74	2.00	6.16
Relationship duration	1.14	.30	2.29	4.57
Big changes in work situation	.36	.47	.56	-1.55
Number of jobs stopped since last interview	.64	.48	1.01	6.07
Hours a day TV is on	8.06	5.83	1.41	1.38

Table 17 *Correlations Between Familial and Environmental Variables by Income Group*

	1	2	3	4	5	6	7	8	9	10	11
Income to Needs Ratio	-	-.12**	.07	-.09*	-.02	.14**	.12**	.05	.12**	-.16**	.14**
CHAOS Sum Score	.01	-	-.17**	-.09*	.04	.19**	.13**	.02	.01	.04	-.23**
Noisiness of Neighborhood	.09**	-.17**	-	.12**	-.02	-.06	-.04	-.05	.06	-.10*	.12**
Number of moves since last visit	-.00	-.09**	.07*	-	-.12**	.01	-.16**	-.13**	.18**	.32**	.06
PR same as last visit	-.04	-.02	.04	-.21**	-	.06	.11*	-.15**	.03	.02	.01
PR living with partner	.21**	.22**	-.18**	.03	-.09**	-	.62**	.13**	.00	.08	-.12**
SR in household	.12**	.28**	-.19**	-.04	-.02	.76**	-	.07	-.05	-.03	-.01
Relationship duration	.06	-.01	.03	-.07*	-.12**	.07*	-.00	-	-.06	-.07	-.14**
Big changes in work situation	-.08*	.01	.02	.05	.01	-.06*	-.04	.00	-	-.01	-.01
Number of jobs stopped since last interview	-.11**	-.20**	.00	.32**	-.01	-.06	-.10**	.03	.03	-	-.03
Hours a day TV is on	.20**	-.19**	.09**	.09**	-.12**	.02	-.01	-.02	-.02	-.05	-

*Correlation significant at .05, **Correlation significant at .01 Note: Values below the diagonal are for low income group and values above the diaganol are for middle/high income group.

Analysis. A simultaneous regression was conducted with the familial and environmental variables. The findings were statistically significant F (10, 743) = 9.32, $p < .05$, accounting for 11.1% of the variance in CHAOS scores for the low income sample, and the findings were statistically significant F (10, 403) = 4.93, $p < .05$, accounting for 10.9% of the variance in CHAOS scores for the middle/high income group. Specific results for each regression analysis are reported in Tables 18 (low income) and 19 (middle/high income). For both income groups, primary caregiver living with his/her romantic partner, number of jobs stopped since the previous data collection time point, and the amount of TV usage per day positively related to scores on the CHAOS. For the low income group, there also was a statistically significant relationship between noisiness of the neighborhood and scores on the CHAOS. Income-to-needs ratio and changes in primary caregiver work situation were related to CHAOS scores for the high income group only.

Follow-Up Analyses: Consideration of a 13-Item CHAOS

Given the low item-total correlations (< .3) and low factor loadings for both the low and middle/high income groups on Items 13 and 15, the one-, two-, and atheoretical models were subsequently run with Items 13 and 15 removed. The model fit indices from these analyses are reported in Appendix C, the factor loadings are reported in Appendix D, and modifications of the 13-item models are reported in Appendices E and H. Both the one- and two- factor version of the CHAOS scale demonstrated a good fit for the data after multiple modifications. However, although the 13-item versions of the CHAOS demonstrated good fit for the data after several modifications, adding more error covariances to any model will increase the fit of the model, and it is only appropriate to add error covariances that have a theoretical justification (Byrne, 2010).

Additionally, multi-group analyses were run on both the one- and two-factor models. The one-factor model had more equivalent items in the 13-item version than the 15-item version. Specifically, eight items (1, 2, 4, 6, 7, 8 10, & 12) were equivalent across groups (Appendices F & G). For the two-factor model, tests of equivalence at the item level indicated that nine items (1, 3, 6, 7, 8, 10, 12, & 14) were equivalent across groups (Appendices I & J). None of the 13-item models demonstrated a significantly better fit than the 15-item models.

Table 18 *Results of Regression for Familial and Environmental Variables on CHAOS Scores for Low Income Families*

	B	*SE B*	β	T	*p*-value	ΔR^2	R^2_{adj}
Overall Model						.33	.11
Income-to-Needs Ratio	.02	.22	.00	.09	.926		
Noisiness of Neighborhood	.62	.19	.12	3.21	.001**	.01	
Number of Moves Since Last Visit	-.03	.16	-.01	-.19	.853		
PR Same as Last Visit	-.01	.87	.00	-.01	.99		
Living With Partner	2.17	.43	.28	5.05	.000**	.05	
SR in Household	-.53	.39	-.08	-1.39	.167		
Relationship Duration	.03	.02	.04	1.19	.234		
Big Changes in Work Situations	-.41	.24	-.06	-1.72	.086		
Number of Jobs Stopped Since Last Visit	-.58	.25	-.08	-2.30	.022*	.01	
TV Usage/Day	.09	.02	.18	4.92	.000**	.04	

**p* significant at .05, **p* significant at .01 level

Table 19 *Results of Regression for Familial and Environmental Variables on CHAOS Scores for Middle/High Income Families*

	B	*SE B*	β	T	*p*-value	ΔR^2	R^2_{adj}
Overall Model						.33	.11
Income-to-Needs Ratio	-.23	.10	-.12	-2.44	.015*	.01	
Noisiness of Neighborhood	.18	.26	.03	.70	.485		
Number of Moves Since Last Visit	-.54	.28	-.10	-1.9	.056		
PR Same as Last Visit	-.89	1.05	-.04	-.85	.396		
Living With Partner	1.80	.64	.18	2.81	.005**	.02	
SR in Household	.24	.78	.02	.30	.764		
Relationship Duration	.01	.02	.03	.58	.562		
Big Changes in Work Situations	.91	.32	.14	1.87	.004**	.01	
Number of Jobs Stopped Since Last Visit	-.82	.34	-.12	-2.39	.017*	.03	
TV Usage/Day	.09	.03	.14	2.78	.006**	.02	

p significant at .05, *p* significant at .01 level

CHAPTER 4: DISCUSSION

The purpose of the study was to examine the reliability and validity of scores from the Confusion, Hubbub, and Order Scale (CHAOS). Confirmatory Factor Analyses (CFAs) were conducted to examine the structural validity of the CHAOS. Results of these analyses indicated that there was no clear best-fitting model for the data and 60% of items were equivalent across income groups. Internal consistency was similar to previous research as α values were high for both the overall scale and within both factors in the two-factor model. Inter-rater reliability between primary and secondary caregivers was low to moderate. In addition, a repeated measures 2 x 2 ANOVA indicated that CHAOS scores were higher at a follow-up home visit; however, CHAOS scores were not different between low and middle/high income families. Regression analysis indicated that the primary caregiver's partner living in the household, number of recent job changes, and number of hours of TV per week were significantly related to CHAOS scores for both the low and middle/high income group.

Factor Structure of the CHAOS

The first hypothesis was that there would be a two-factor structural model reflecting the domains of Routines/Organization and Disorganization/Confusion/Noise (Matheny et al., 1995). While the authors of the CHAOS suggested this structural model, they did not conduct a factor analysis to provide evidence for this model. In the present study, fit indices were strongest for the two-factor model; however, they did not appear significantly different from the one- and atheoretical models. In fact, some indices for the one factor model were slightly better than the indices for the two-factor model. Further, none of the hypothesized models met all of the a priori criteria for evidence of good fit. The SRMR, RMSEA, and GFI indices all indicated an adequate or good fit for all of the models, but the CFI, NFI, and TLI did not indicate good fit for any of the models. However, the best fitting model is the two factor model, and this model may have some utility as the two factors may relate differently to other variables.

In addition, when the models were examined for equivalence, there were several differences in the loadings of the individual items across groups. Only Item 10 (can't hear self think) was equivalent across groups in both the one- and two-factor models. The items that were not equivalent for either of the models included Item 3 (always seem to be rushed), Item 4 (able to stay on top of things), Item 9 (plans don't seem to work out), Item 11 (drawn into arguments), and Item 14 (atmosphere is calm). Items 4 and 9 both have to do with plans and scheduling. It is possible that families in the different income groups have different attitudes and practices when it comes to making plans and following through with commitments. It also appears families across income groups experience conflict differently based on the content of Item 11.

Atmosphere (Items 14) and pace (Item 3) also appear to be experienced differently based on income level. This information raises questions about the organizational structure of the construct of household chaos proposed by the authors. Although some of the a priori fit criteria were met, the structure of the scale is not equivalent across income groups, and differences in scores related to SES should be interpreted with caution.

Internal Consistency

The second hypothesis was that the scale would demonstrate moderate to high internal consistency. The Cronbach's α value in the current study was high for the overall model, consistent with those found in the work by Matheny et al. (1995), Dumas et al. (1995), and Haack et al. (2011). Cronbach's α values were also high for each of the two factors suggested by Matheny et al. Item-total correlations in the current study were moderate. These ranges are similar to the findings in the study conducted by Matheny et al. where most individual items correlated with total CHAOS scores moderately (.3 or higher). In the current study, two items were below .3, indicating that these items may not be functioning in a similar manner to the other items on the scale. These items reflect the time spent on the telephone and whether or not the family has a typical morning routine.

There are several reasons why these items may not be functioning as intended. For example, raters may not feel the telephone interrupts their plan, but rather is integral to household functioning. Further, with the increase of cell phones, text messages, emailing, and voicemail, many people are better able to choose a convenient time for them to answer the phone/communicate with others. In regard to the item about routine, perhaps families do not have an expectation of a regular routine when their children are as young as they were in the current study. As kids start formal schooling and need to complete time sensitive activities such as catching the bus, families may establish more of a routine. Matheny et al. (1995) also found a low item-total correlation for the item that encompasses whether or not the family has a typical morning routine. Given the consistency across both studies, this item may not reflect the concept of household chaos as much as the other items. However, the overall high internal consistency as well as the high factor level internal consistency is a strength of the scale and suggests that the constructs in the scale relate to each other and to the two factors suggested by Matheny et al.

Inter-rater Reliability

The third hypothesis was that there would be a high correlation between primary and secondary caregiver scores on the CHAOS. Based on the criteria by Fleiss (1981), Landis and Koch (1977), and Sattler (2009), the observed Kappa coefficients are low. While Kappa coefficients are low, they are considered a stringent reliability index. Based on the criteria specified by Sattler (2009), the intra-class correlations also are low. The percent exact agreement was 70%. Graham, Milanowski, and Miller (2012) indicated that between 75% and 90% agreement is sufficient; however they noted that the average agreement in the studies they reviewed was 70%. Thus, there is not strong evidence for inter-rater reliability. There are several explanations for why raters would have such different scores. One possibility is that they experience a qualitatively different set of activities as a result of being present in the home at different times of the day. For example, if one caregiver is home during the morning and another at night, there could be fundamental differences in the types of activities families engage in during those times, thus impacting their perceptions of the home environment.

The differing experiences of the two caregivers could lead to different ratings. For example, if one caregiver grew up in a highly structured home and the other one grew up in a chaotic environment, they may have different frames of reference for their current household. While they may get more similar over time, different frames of reference may influence responses on the CHAOS. Further, the activity level in the responders' experiences outside the home on a regular basis (e.g., work or volunteer activities) may also influence their perception of chaos. If one responder has a very chaotic and stressful work place or a very calm workplace, difference may result in varying perspectives of how he/she perceives his/her experiences at home. Even though there may be plausible reasons why respondents in the current study consistently rated CHAOS items differently, the reliability of scores between raters was still low. Therefore, the scale may be a better indicator of the way a rater *experiences* environmental chaos rather than a rating of *actual* household chaos, which may or may not make a difference in outcomes related to household chaos. For research purposes, reliability evidence is at best minimally-sufficient to justify use as an overall indicator of CHAOS. If used, it would be prudent to have at least two raters if possible, so that the agreement between raters can be reported (Graham et al., 2012).

Change in CHAOS Scores Over Time

It also was hypothesized that there would be differences in CHAOS scores over time based on SES groups. Results indicated that while CHAOS scores were higher (indicating more household chaos) at the 35-month time point, but that there no interaction between income level and time. Further there was no main effect for income level. This is consistent with some previous findings indicating that environmental chaos is a factor independent of SES (e.g., Dumas et al., 2005; Vernon-Feagans, Garrett-Peters, DeMarco, et al., in press), but inconsistent with Evans et al. (2005) who found children in poverty experienced more home chaos than their middle- or high-income counterparts. However, there could be variation across income level in the items endorsed on the chaos scale, but not differences in overall chaos scores. Perhaps differences did not materialize because there is not enough variability in CHAOS scores for a difference in scores to emerge by group; however, differences were seen across time points so this may not be the case.

Another potential explanation that a difference may not have been evident across income groups is because the target children in the study were all very young, which in itself is chaotic. Previous literature does indicate that as the age of children in a household increases, the household chaos decreases (Lichter & Wellington, 2010). This is likely due to the amount of time and attention parents need to provide to young children. Because young children have less predictable schedules and require more direct supervision, parents cannot focus on routines and organization as easily. This increases the level of chaos in the home.

CHAOS, Environmental, and Familial Variables

The relationships between CHAOS scores and other indicators of environmental instability also were examined in the current study. Results indicated that whether or not the primary caregiver was living with their romantic partner, the number of job changes since the last visit, and the number of hours the TV was on in a week were statistically significantly related to CHAOS scores for both the low and middle/high income group. A secondary caregiver may provide more structure to the environment and offer an extra pair of hands to help routines. Changes in work situation, especially changes in hours, can disrupt family schedules and be a stressor that causes increased tension. Having the TV on for many hours creates ambient noise, which contributes to household chaos. Because these relationships were significant across income levels, this suggests that these variables may be strongly related to CHAOS scores regardless of income level.

The noisiness of the neighborhood also was statistically significantly related to CHAOS scores for the low income group. Whether or not a neighborhood is noisy may effect CHAOS scores because of auditory confusion created by the extra noise. The differences in auditory stimulation within the low income group may be greater than in the middle high income group. For example, in more urban/city environments, people with lower incomes tend to live in apartments, trailer homes, or single-family homes with minimal separation. Because neighbors may live closer together, they may hear each other more. Conversely, people who have lower incomes and live in isolated very rural settings may experience a calmer environment. Changes in work situation and income-to-needs ratio were also statistically significantly related to CHAOS scores for the middle/high income group. These results suggest that the middle/high income families experience more disruption from changes in financial, routine, and schedule than low-income families. One potential explanation of this finding is that low-income families are more accustomed to having changes in economic situations and have a better understanding of resources available to them. As a result, they are more resilient to such changes than middle/high income families. Nonetheless, the significance of income-to-needs ratio could also be an artifact of how the subsamples were constructed. Specifically, because a 2.0 income-to-needs ratio was used as the cut-off between the low and middle/high income groups, and the study has an overrepresentation of families who live in poverty, it may be that the middle/high income group actually is comprised of a significant number of families that are "near poor". Therefore, the middle/high income group may experience more chaos with changes in work situations because it makes the difference between living in poverty or not. Additionally, there is a large variation in the income-to-needs ratio of this group.

13-Item CHAOS

In light of the questionable contribution of two items to the CHAOS, exploratory analyses were conducted using a 13-item version of the CHAOS. Results of these analyses yielded initial fit indices that were similar to those observed for the 15-item version of the scale, and overall fit improved with modifications. Ultimately, removing the items with low item-total correlations did not help overall model fit, suggesting that there are other reasons for the lack of model fit. Because there was a lack of convergence on a particular model, the model that was endorsed theoretically by the author's was the model used for the remaining analyses; however, because this model still did not demonstrate a good fit across indices, using the CHAOS should be used with caution in future research. Because the items are not equivalent across income groups, it may be important to use the CHAOS with samples that are relatively homogeneous with regard to income. This finding also raises questions regarding the appropriateness of retaining certain items (e.g., 2 [find things when needed], 5 [always running late]) across income levels.

Limitations

There are several limitations of the current study. There were some missing data (5.9%) in the study that needed to be imputed. In addition, the sample also over represented families who live in poverty or have low levels of income; however, analyses were conducted by income group to examine possible differences. Additionally, because the Family Life Project studies a rural population, these data may not generalize to urban populations. However, there are variations of rurality within the sample such that some families do live in a distal suburban or town setting. Further, participating families have very young children and thus only represented one phase of family life. Finally, because not all families had two caregivers participate, there were fewer participants in the inter-rater reliability analysis. There could be differences in household characteristics between those with and without secondary caregivers that may limit generalizability of results

Conclusions/Implications Regarding the CHAOS

While there are some limitations to the methodology of the present study, perhaps the most significant limitation is the scale itself. There were several positive psychometric characteristics of the scale. For example, a priori criteria for factor structure were met across models. Additionally, there was a high internal consistency for the overall model as well as both factors suggested by Matheny et al. However, there was no convergence of evidence on a single best fitting model. However, the two-factor model achieved a slightly better fit, and the two factors might correlate with other variables differently. More specifically, the routines and organization factor may relate to outcome variables differently than the disorganization, confusion, and noise items. An additional issue regarding internal structure was that the items do not fit the model in the same way across income groups. Another limitation of the model is that there was not a strong demonstration of reliability of scores across raters. Overall, the observed mixed evidence for reliability and validity of CHAOS scores suggests it should not be used for high stakes decision-making. To use the scale in research, it will be important to think about the characteristics of the target sample and obtain ratings the information from more than one respondent to establish reliability of scores.

Perhaps psychometric evidence is limited because the constructs captured by the scale are not experienced the same way for families of different economic backgrounds. Similarly, there may be some aspects of chaos experienced by families in both income groups that the scale does not measure. For example, one aspect of chaos is the number of people moving in and out of the house, which occurs more in lower income families (Lareau, 2003). In families with lower income levels, financial circumstances can lead multiple adults (and children) to move in and out of the household at different time points. Further, with increases in the number of people in the household, there could be a higher room to people ratio in this income group. Person-to-room ratio is another characteristic of household chaos not captured by the scale. While there are plausible explanations for these income-related differences, there are clear differences in the structure of the scale based on income level. Therefore, it is particularly important to not use the CHAOS when trying to examine differences across income levels because differences observed statistically could be an artifact of the scale rather than true differences in the level of household chaos being experienced. However, there is no difference in total CHAOS scores between groups, which suggests the differences between income groups involve which specific items are endorsed, rather than the number of items endorsed.

In addition to the facets of chaos a family experiences, a family's perception of the chaos they experience could be situation dependent. Therefore, the dichotomous format of the scale may not be sufficiently sensitive to allow varying perceptions on a day to day basis, while still allowing the household chaos in general to be captured. For example, changes in life circumstances around the time that the CHAOS questions are asked could influence responses (e.g., someone changing jobs or a new baby in the family). Further, there could be individual differences in a family's perception of chaos. Specifically, some people may have a higher tolerance level for household chaos. Increasing the continuum of responses may allow there to be some variation in perceptions of household chaos, while at the same time providing a more accurate general idea of the level of household chaos experienced by the family.

An additional limitation is that the dichotomous response format may not fully capture the concepts that the authors of the studies intended because it may be difficult for participants to commit to yes or no. It is possible that a family's overall chaos score is higher or lower than their actual experience of household chaos due to the dichotomous response format for each item. If study participants were able to respond with "sometimes", "occasionally", or other points on a response continuum, the user may get a clearer picture of the amount of chaos a family experiences.

The difficulties with the dichotomous items of the scale, as well as its situation-dependent nature are both reflective of another limitation of the scale, that it is self-report. The fact that the scale is based on self-report can be an asset because it is more reflective of how families actually perceive their routines, organization, and noise. However, it would also be easy for families to provide more socially desirable responses or be excessively negative if they are administered the scale at a particularly stressful moment.

A further limitation is that the scale does not reflect recent advances in technology that may contribute to household chaos. Many families have more video games, more complex phones, the Internet and other technology than they did when the scale was created in 1995. Thus, there could be new facets of chaos that are not considered within the CHAOS. Increases in the amount of technology available may make it more difficult for parents and their children to interact with each other, which could lead to both language and socio-emotional difficulties. Additional technology devices may also make it more difficult for all members of the family to focus, which could create difficulties for families to maintain a routine leading to higher levels of chaos. However, as previously discussed, it also could be true that the prevalence of technology today may actually lower the day-to-day perception of chaos because caregivers are better able to choose when they respond to interactions outside the home than they could when there were landline phones. Emailing, text messaging, voicemail, and even caller id can make it easier for families to choose whether or not to interrupt activity to attend to someone else. Because technology is now such a huge part of many levels of the multilevel, multidimensional environment proposed by Bonfrenbrenner, this topic needs to be investigated further to determine the impact of technology on both household chaos and child development outcomes.

Finally, while the scale purports to include a measure of noise, it is not really a decibel level (i.e., noise), but rather a different construct within the scale, for example, auditory stimulation. While auditory stimulation may make it difficult for children and adults to concentrate, have conversations, and listen to each other, calling it noise is not accurate. Because noise is not really part of the scale, it may be a critical aspect of chaos that the CHAOS is not capturing.

Recommendations for Development of an Alternative Measure of Chaos

Based on the limited psychometric evidence justifying the use of the CHAOS, it is important that researchers develop new measures for assessing household chaos. Even though the CHAOS scale does not have very strong evidence for the reliability and validity of scores, it is clear from examining the literature that household chaos, measured in multiple ways, including the CHAOS scale, does have a moderate relationship with child outcomes. Even if the scale was refined, studies that show the relationship between household chaos and childhood outcomes are based on limited measures of chaos. Additional tools for assessing household chaos must be developed, refined, and validated to ensure that researchers can adequately (and accurately) study this construct.

When addressing the limitation that the CHAOS scale is not equivalent across income groups, it is also important to consider whether people of varying SES experience the negative outcome associated with SES in the same way. For example, it has been suggested that chaos often has more harmful effects on children from middle or upper income families because for children in lower income families, there is already a high level of disorder, so chaos cannot contribute to it that much more (e.g., Farmer, 1989). Conversely, for children from middle/high income families it is more atypical to have higher levels of adversity, thus it is more likely to contribute more heavily to negative outcomes. This idea was supported by the current study since changes in work situation and income to needs ratios seemed to relate more strongly to CHAOS scores for middle/high income families than for families in the lower income group.

In addition to outcomes, there may be differences in the way families of differing incomes experience CHAOS itself. For example, sticking to a schedule like organizing children's extra-curricular activities is an activity that is more common in middle class families

(Lareau, 2003). Further, some other elements of chaos are not captured by the CHAOS, like people moving in and out of the house, person to room ratio, and neighborhood characteristics, such as danger, noise, and population density, may be more typical of families of lower socioeconomic backgrounds. It would be helpful to develop a scale that captures aspects of chaos that impact families across SES.

In addition to varying SES interpretations, the scale also does not account for activity resulting from increased technology use in many homes. For example, video games, more television channels, and family members using their own individual communication devises all increase the amount of activity occurring in households; however, some of this new technology may also increase the ability of families to maintain their routines and organization because they have more of a choice about when to communicate. Future researchers should both develop measurement tools and examine the effects of the increase use of technology in home.

When measuring the construct, it may be helpful to collect both subjective and objective information and use an expanded continuum of response options. For example, decibel level and person-to-room ratio could be observed by researchers. The concept of people coming in and out of the household on a regular basis will also be a good construct to capture as it may impact the amount of chaos a family experiences. Finally, the relatively brief and dichotomous nature of the scale, may not allow a full representation of the level of household chaos. Instead, it may be useful to score items on a Likert scale, as was done in previous studies (e.g., Johnson et al., 2005).

Directions for Future Research Involving the Impact of Chaos

Once more effective tools for examining household chaos are developed and refined, it is important to continue to study this construct and its relationship with the functioning of children and families. If the scale could be more fine-tuned, researchers could look for effects of a chaos threshold. For example, they could investigate whether there is such a thing as too much or too little chaos as Wachs and Evans (2010) suggested. Whether or not there is a threshold, parent training may be a useful tool for helping parents maximize helpful stimulation for their child, while at the same time reducing the amount of harmful chaos their child experiences. The impact of such interventions should be examined by future researchers.

Further research needs to be conducted to determine the relationships between parental pathology, child pathology, and household chaos. It may be that parent pathology produces chaos. However, an alternative explanation is that child pathology leads to parenting that is not conducive for maintaining order. From an ecological perspective, it is most likely that these variables interact in a reciprocal manner. Path analysis between parental pathology, child pathology, and chaos should be conducted to determine how these variables relate to each other. Additionally, further path analyses could be conducted to determine whether or not providing parents with outside resources may impact this pathway.

Conclusions

The purpose of the study was to examine reliability and validity evidence for CHAOS scores. Results raise questions about the psychometric properties of the measure. Even though CHOAS scores did vary across time points, there was high internal consistency for both the overall items in the scale as well as each of the two factors, and several familial and environmental characteristics significantly related to CHAOS scores, there was also a lot of evidence indicating that the CHAOS has some reliability and validity issues. Specifically, none of the models tested met a priori criteria indicative of good fit with the data, and most of the items did not function equivalently across groups. Additionally, inter-rater reliability was low.

Based on the results of this study, there appear to be several ways the current scale could be modified to increase its utility, such as including items related to traffic patterns, noise, and new technology. Because evidence for psychometric properties of the CHAOS was mixed at best, use of the CHAOS is not recommended at this time.

References

Acs, G., & Gallahger, M. (2000). Income inequality among America's children. *National Survey of America's Families.* Series B. Retrieved from http://www.urban.org/

Adam, E. (2004). Beyond quality: Parental and residential stability and children's adjustment. *Current Directions in Psychological Science, 13(5),* 210-213. doi: 10.1111/j.0963-7214.2004.00310.x

Adam, E. K., & Chase-Landsdale, L. (2002). Home sweet home(s): parental separations, residential moves, and adjustment problems in low-income adolescent girls. *Developmental Psychology, 38 (5),* 792-805. doi: 10.1037//0012-1649.38.5.792.

Albright, Jeremy J., and Hun Myoung Park. (2008).*Confirmatory Factor Analysis Using Amos, LISREL, Mplus, and SAS/STAT CALIS*. Technical Working Paper. The University Information Technology Services (UITS) Center for Statistical and Mathematical Computing, Indiana University.

Bada, H. S., et al. (2008). Importance of stability of early living arrangements on behavior outcomes of children with and without prenatal drug exposure. *Journal of Development Behavior Pediatrics, 29,* 173-182. doi: 10.1097/DBP.0b013e3181644a79.

Baker, B. L., McIntyre, L. L., Blacher, J., Crnic, K., Edelbrock, C., & Low, C. (2003). Pre-school children with and without developmental delay: Behavior problems and parenting stress over time. *Journal of Intellectual Disability Research, 47*, 217-230. doi: 10.1046/j.1365-2788.2003.00484.x.

Billows, M., Gradisar, M., Dohnt, H., Johnston, A., McCappin, S., & Hudson J. (2009). Family disorganization, sleep hygiene, and adolescent sleep disturbance. *Journal of Clinical Child and Adolescent Psychology, 38,* 745-752. doi: 10.1080/15374410903103635.

Bronfenbrenner, U. (1999). Environments in developmental perspective: Theoretical and operational models. In S. L. Friedman & T. D. Wachs (Eds.), Measuring environment across the lifespan: Emerging methods and concepts (pp. 3-28). Washington, DC: American Psychological Association Press.

Bronfenbrenner, U., & Evans, G. W. (2000). Developmental science in the 21st century: emerging questions, theoretical models, research designs and empirical findings. *Social Development, 9 (1),* 115-125. doi: 10.1111/1467-9507.00114.

Brown, E. D., & Low, C. M. (2008). Chaotic living conditions and sleep problems associated with children's responses to academic challenge. *Journal of Family Psychology, 22,* 920-923. doi: 10.1037/a0013652.

Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming.* New York, NY: Routledge.

Chen, E., Cohen, S., & Miller, G. E. (2010). How low socioeconomic status affects 2-year hormonal trajectories in children. *Psychological Science, 21,* 31-37. doi: 10.1177/0956797609355566.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112,* 155-159. doi:10.1037/0033-2909.112.1.155

Coldwell, J., Pike, A., & Dunn, J. (2006). Household chaos-links with parenting and child behaviour. *Journal of Child Psychology and Psychiatry, 47(11),* 1116-1122. doi: 10.1111/j.1469-7610.2006.01655.

Corapci, F., & Wachs, T. D. (2002). Does parental mood of efficacy mediate the influence of environmental chaos upon parenting behavior? *Merrill-Palmer Quarterly, 48,* 182-201. doi: 10.1353/mpq.2002.0006.

Deater-Decker, K., Mullineaux, P. Y., Beckman, C., Petrill, S. A., Schatschneider, C., Thompson, L. A. (2009). Conduct problems, IQ, and household chaos: a longitudinal multi-informant study. *The Journal of Child Psychology and Psychiatry, 50,* 1301-1308. doi: 10.1111/j.1469-7610.2009.02108.x

Dodge et al. (1990). *The Fast Track Project: Post Visit Inventory*. Retrieved from http://fasttrackproject.org/techrept/p/pvc/

Dumas, J. E., Nissley, J., Nordstrom, A., Smith, E. P., Prinz, R. J., & Levine, D. W. (2005). Home chaos: sociodemographic, parenting, interactional, and child correlates. *Journal of Clinical Child & Adolescent Psychology*, *34*, 93-104. doi: 10.1207/s15374424jccp3401_9.

Evans, G. W. (2006). Child development and the physical environment. *Annual Review of Psychology*, *57*, 423-451. doi: 10.1146/annurev.psych.57.102904.190057.

Evans, G. W., Gonnella C., Marcynyszyn, L. A., Gentile, L., & Salpekar, N. (2005). The role of chaos in poverty and children's socioemotional adjustment. *Psychological Science, 16,* 560-565. doi: 10.1111/j.0956-7976.2005.01575.x

Evans, G. W., Hygge, S., & Bullinger, M. (1995). Chronic noise and psychological stress. *Psychological Science, 6,* 333-338. doi: 10.1111/j.1467-9280.1995.tb00522.x

Evans, G. W., Kliewer, W., & Martin, J. (1991). The role of the physical environment in the health and well being of children. In H. E. Schroeder (Ed.), *New directions in health psychology assessment* (pp. 127-157). New York, NY: Hemisphere.

Evans, G. W., Lepore, S. J., Shejwal, B., & Palsane, M. N. (1998). Chronic residential crowding and children's well-being. *Child Development, 69,* 1514-1523. doi: 10.2307/1132129.

Evans, G. W., Maxwell, L. E., & Hart, B. (1999). Parental language and verbal responsiveness to children in crowded homes. *Developmental Psychology, 35(4),* 1020-1023. doi: 10.1037//0012-1649.35.4.1020.

Evans, G., & Wachs, T. (2010). Chaos and its influence on children's development: An ecological perspective. Washington, DC: American Psychological Association.

Farmer, E. M. Z. (1991). *Maladjustment in the life course: Young adult outcomes of maladjusted boys.* (Doctoral dissertation). Retrieved from Dissertations and Theses. (9206650).

Fleiss, J. L. (1981). *Statistical methods for rates and proportions* (2nd ed.). New York: John Wiley.

Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.) London: Sage.

Flouri, E. (2009). Strong families, tidy houses, and children’s values in adult life: are “chaotic,” “crowded,” and “unstable” homes really so bad?. *International Journal of Behavioral Development, 33,* 496-505. doi: 10.1177/0165025409340090.

Fontaine, N. M. G., McCrory, E. J. P., Boivin, M., Moffitt, T. E., & Viding, E. (2011). Predictors and outcomes of joint trajectories of callus-unemotional traits and conduct problems in childhood. *Journal of Abnormal Psychology, 120,* 730-742. doi: 10.1037/a0022620.

Forssman, L., Eninger, L., Tillman, C. M., Rodriguez, A., & Bohlin, G. (2010). Cognitive functioning and family risk factors in relation to symptoms behaviors of adhd and odd in adolescents. *Journal of Attention Disorders, 20,* 1-11. doi: 10.1177/1807054710385065.

Graham, M., Milanowski, A., & Miller, J. (2012). *Measuring and prompting inter-rater agreement of teacher and principal performance ratings.* [Bulletin]. Center for Educator Compensation Reform.

Gregory, A. M., Eley, T. C., O’Connor, T. G., Rijsdijk, F. V., & Plomin, R. (2005). Family influences on the association between sleep problems and anxiety in a large sample of pre-school aged twins. *Personality and Individual Differences, 39*, 1337-1348. doi: 10.1016/j.paid.2005.06.008.

Haack, L. M., Gerdes, A., Schneider, B. W., & Hurtado, G. D. (2011). Advancing our knowledge of adhd in latino children: psychometric and cultural properties of spanish-versions of parental/family functioning measures. *Journal of Abnormal Child Psychology, 39*, 33-43. doi: 10.1007/s10802-010-9441-y.

Hair, J. F., Black, B., Babin, B., Anderson, R. E., & Tatham, R. L. (2005). *Multivariate Data Analysis* (6th Ed.). Upper Saddle River, NJ: Prentice Hall.

Hanscombe, K. B., Haworth, C. M. A., Davis, O. S. P., Jaffee, S. R., & Plomin, R. (2011). Chaotic homes and school achievement: a twin study. *Journal of Child Psychology and Psychiatry*, *52*, 1212-1220. doi: 10.1111/j.1469-7610.2011.02421.x.

Hart, S. A., Petrill, S. A., Deater-Deckard, K., & Thompson, L. A. (2007). SES and CHAOS as environmental mediators of cognitive ability: A longitudinal genetic analysis. *Intelligence, 35*, 233-242. doi: 10.1016/j.intell.2006.08.004

Hu, L., & Bentler, P. M. (1999). Cut off criteria for fit indices in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1-55. doi:10.1080/10705519909540118

Johnson, A. D., Martin, A., Brooks-Gunn, J., & Petrill, S. A. (2008). Order in the house! Associations among household chaos, the home literacy environment, maternal reading ability, and children's early reading. *Merrill-Palmer Quarterly, 54*(4), 445-472. doi: 10.1353/mpq.0.0009.

Kline, R. B. (2005). *Principles and practice of structural equation modeling.* New York, NY: The Guilford Press.

Kretschmer, T., & Pike, A. (2009). Young children's sibling relationship quality: distal and proximal correlates. *Journal of Child Psychology and Psychiatry, 50,* 581-589. doi: 10.1111/j.1469-7610.2008.02016.x.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33 (1),* 159-174. doi:10.2307/2529310.

Lareau, Annette. (2003). *Unequal childhoods: Class, race, and family life.* London: University of California Press, Ltd.

Licter, D. T., & Wethington, E. (2010). Chaos and the diverging fortunes of american children: a historical perspective. In G. Evans & T.Wachs (Eds.), *Chaos and its influence on children's development: An ecological perspective* (255-264). Washington, DC: American Psychological Association

Marcynyszyn, L. A., Evans, G. W., & Eckenrode, J. (2008). Family instability during early and middle adolescence. *Journal of Applied Developmental Psychology, 29(5),* 380-392. doi: 10.1016/j.appdev.2008.06.001

Martin, A., Razza, R. A., & Brooks-Gunn, J. (2011). Specifying the links between household chaos and preschool children's development. *Early Child Development and Care, 1*, 1-17. doi: 10.1080/03004430.2011.605522.

Matheny, A. P., Wachs, T. D., Ludwig, J. L., Phillips, K. (1995) Bringing Order out of Chaos: Psychometric Characteristics of the Confusion, Hubbub, and Order Scale. *Journal of Applied Developmental Psychology, 16,* 429-444. doi: 10.1016/0193-3973(95)90028-4.

Maxwell, L. E. (2003). Home and school density effects on elementary school children: The role of spatial density. *Environment and Behavior, 35*, 566-578.

doi: 10.1177/0013916503035004007

McLoyd, V. (1990). The impact of economic hardships on Black families and children: Psychological distress, parenting and socioemotional development. *Child Development, 61,* 311-346. doi: 10.2307/1131096.

Milan, S., Pinderhughes, E. and the Conduct Problems Prevention Research Group (2006). Family instability and child maladjustment trajectories during elementary school. *Journal of Abnormal Child Psychology, 34(1),* 43-56. doi: 10.1007/s10802-005-9007-6.

Mokrova, I., O'Brien, M., Calkins, S., & Keane, S. (2010). Parental adhd symptomology and ineffective parenting: the connecting link of home chaos. *Parenting: Science and Practice, 10,* 119-135. doi: 10.1080/15951909032l2844.

Nelson, J. A., O'Brien, M., Blankson, A. N., Calkins, S. D., & Keane, S. P. (2009). Family stress and parental response to children's negative emotions: tests of the spillover, crossover, and compensatory hypotheses. *Journal of Family Psychology, 23 (5),* 671-679. doi: 10.1037/a0015977.

Petrill, S. A., Pike, A., Price, T., & Plomin, R. (2004). Chaos in the home and socioeconomic status are associated with cognitive development in early childhood: environmental mediators identified in a genetic design. *Intelligence, 32,* 445-460. doi: 10.1016/j.intell.2004.06.010.

Raviv, T., Kessenich, M., & Morrison, F. J. (2004). A mediational model of the association between socioeconomic status and three-year-old language abilities: The role of parenting factors. *Early Childhood Research Quarterly, 19,* 528–547. doi: 10.1016/j.ecresq.2004.10.007

Roth, P. L., & Switzer, F. S., (1999). Missing data: Instrument-level heffalumps and item-level woozles. *Research Methods Forum.* Retrieved from: http://www.aom.pace.edu.

Sameroff, A. (2010). Dynamic developmental systems: Chaos and order. In G. Evans & T.Wachs (Eds.), *Chaos and its influence on children's development: An ecological perspective* (255-264). Washington, DC: American Psychological Association

Sampson, R. J., Morenoff, J. D., & Earls, F. (1999). Beyond social capital: Social dynamics of collective efficacy for children. *American Sociological Review, 64,* 633-660. doi: 10.2307/2657367

Sattler, J. (2001). *Assessment of Children: Cognitive Applications* (4th ed.). San Diego: Jerome Sattler, Publisher, Inc.

Shamama-tus-Sabah, S., Gillani, N., & Wachs, T. D. (2011). Relation of home chaos to cognitive performance and behavioral adjustment of pakistani primary school children. *International Journal of Behavioral Development, 35,* 507-516. doi: 10.1177/0165025411406852

Smith, E. P., Connell, C. M., Wright, G., Sizer, M., Norman, J. M., Hurley, A., et al. (1997). An ecological model of home, school, and community partnerships: Implications for research and practice. *Journal of Educational and Psychological Consultation, 8,* 339-360. doi: 10.1207/s1532768xjepc0804_2

Smith, E. P., Prinz, R. J., Dumas, J. E., & Laughlin, J. E. (2001). Latent models of family processes in African American families: Relationships to child competence, achievement, and problem behavior. *Journal of Marriage and Family, 63,* 967-980. doi: 10.1111/j.1741-3737.2001.00967.x

Trouton, A., Spinath, F. M., & Plomin, R. (2002). Twins early development study (TEDS): a multivariate, longitudinal genetic investigation of language, cognition and behavior problems in childhood. *Twin Research, 5,* 444-448. doi: 10.1375/13690520906255.

Vernon-Feagans, L., Cox, M., & the Family Life Project Key Investigators (in press). The family life project: an epidemiological and development study of young children living in poor and rural communities. *Society for Research in Child Development.*

Vernon-Feagans, L., Garrett-Peters, P. DeMarco, A. & Bratsch, M. Children Living in Rural Poverty: The Role of Chaos in Early Development. (2012). In V. Maholmes & R. King (Eds). *The Oxford Handbook of Poverty and Child Development* (pp. 448-466). New York, NY: Oxford University Press.

Vernon-Feagans, L., Garrett-Peters, P.T., Willoughby, M.T., Mills-Koonce, W.R., & The Family Life Project Key Investigators (2012). Poverty, chaos, and parenting: Predictors of early language development. *Early Childhood Research Quarterly, 27,* 339-351. doi: 10.1016/j.ecresq.2011.11.001.

Wachs, T. D. (1993) Nature of relations between the physical and social microenvironment of the two-year-old child. *Early Development and Parenting, 2*, 81-87. doi: 10.1002/edp.2430020202

Wachs, T. D., & Chan, A. (1986). Specificity of environmental action, as seen in environmental correlates of infants' communication performance. *Child Development, 57,* 1464-1474. doi: 10.2307/1130424

Wachs, T. D., & Evans, G. W. (2010). Chaos in context. In G. Evans & T.Wachs (Eds.), *Chaos and its influence on children's development: An ecological perspective* (255-264). Washington, DC: American Psychological Association.

Wachs, T. D., & Camli, O. (1991). Do ecological or individual characteristics mediate the influence of the physical environment upon maternal behavior? *Journal of Environmental Psychology*, *11*, 249-264. doi: 10.1016/S0272-4944(05)80186-0

Wahler, R. G., & Dumas, J. E. (1989). Attention problems in dysfunctional mother-child interactions: An interbehavioral model. *Psychological Bulletin*, *105,* 116-130. doi: 10.1037//0033-2909.105.1.116

Appendix A

Confusion, Hubbub, and Order Scale (CHAOS)

1. There is very little commotion in our house

2. We can usually find things when we need them

3. We almost always seem to be rushed

4. We are usually able to stay on top of things

5. No matter how hard we try, we always seem to be running late

6. It’s a real zoo in our home

7. At home we can talk to each other without being interrupted

8. There is often a fuss going on at our home

9. No matter that our family plans, it usually doesn’t seem to work out

10. You can’t hear yourself think in our home

11. I often get drawn into other people’s arguments at home

12. Our home is a good place to relax

13. The telephone takes up a lot of out time at home

14. The atmosphere in our home is calm

15. First thing in the day, we have a regular routine at home

Appendix B

Windshield Survey

1. Family's preparation for session and organization of session.
   0 = Can't rate   1 = Surprise/Difficulty   2 = Aware but unprepared
   3 = Aware/Ready   4 = Good Hosts

2. Primary Respondent's receptivity toward visitors.
   0 = Can't rate   1 = Very Uncomfortable   2 = Distant but Polite
   3 = Average Friendliness   4 = Very warm

3. Secondary Respondent's receptivity toward visitors.
   0 = Can't rate   1 = Very Uncomfortable   2 = Distant but Polite
   3 = Average Friendliness   4 = Very warm   5=Not Applicable, no secondary respondent

4. How much difficulty did you have in completing this interview?
   0 = Can't rate   1 = Very Smooth   2 = Slight Difficulty
   3 = Some Difficulty   4 = Great Difficulty

5. Do you have reason to doubt the validity of this interview and home visit?
   0 = Can't rate   1 = Probably Valid   2 = Respondent Responses Possibly Invalid
   3 = Definitely Reasons to Doubt Validity

6. How clean is this dwelling?______________________________
   0 = Can't rate   1 = Very Dirty   2 = Slightly Dirty   3 = Messy   4 = Clean

7. How safe is the interior of this building?
   0 = Can't rate   1 = Obviously dangerous   2 = Slightly Dangerous
   3 = Average   4 = Above Average Safety

8. How many rooms are in this dwelling?______________________________
   0 = Can't rate   1 = 1 or 2   2 = 3 or 4   3 = 5 or 6   4 = >6

9. How safe is the area outside of this building?
   0 = Can't rate   1 = Obviously dangerous   2 = Slightly Dangerous
   3 = Average   4 = Above Average Safety

10. The street on which this dwelling is located is:? _____________
    0 = Can't rate   1 = Mainly residential   2 = Mixed Resid & Commercial
    3 = Mostly Commercial   4 = Rural or Agricultural

11. The noise level in this neighborhood around this dwelling is: __________
    0 = Can't rate   1 = Very Quiet   2 = Average   3 = Noisy   4 = Very Noisy

12. The safety of the neighborhood around this dwelling is:? _____________

0 = Can't rate 1 = Very Safe/Crime Free 2 = Average for This City
3 = Unsafe 4 = Very Unsafe/High Risk

Appendix C

Fit Indices for Alternative 13-Item CHAOS Models

	χ^2	DF	SRMR	RMSEA	GFI	CFI	NFI	TLI	AIC	BCC
13-Item										
One-Factor	662.29	130	.06	.06	.92	.81	.78	.78	766.29	769.12
Two-Factor	902.62	128	.06	.06	.93	.84	.82	.80	1010.62	1012.31
Atheoretical	651.03	128	.06	.06	.92	.82	.78	.78	759.03	761.96

Appendix D

Factor Loadings for 13-Item Scale

	Low Income			Middle/High Income		
	One-Factor	Two-Factor	Atheoretical	One-Factor	Two-Factor	Atheoretical
1. Little commotion	.43	.49	.43	.50	.48	.51
2. Find things when needed	.41	.37	.41	.32	.38	.32
3. Always seem rushed	.50	.53	.51	.53	.52	.53
4. Able to stay on top of things	.41	.37	.42	.35	.39	.36
5. Always running late	.46	.46	.46	.36	.44	.36
6. Real zoo	.67	.70	.68	.62	.67	.63
7. Talk without interruption	.46	.48	.46	.41	.47	.41
8. Fuss going on at home	.61	.65	.61	.68	.64	.69
9. Plans don’t seem to work out	.44	.33	.45	.23	.39	.23
10. Can’t hear self think	.44	.46	.45	.42	.46	.42
11. Drawn into arguments	.50	.47	.51	.27	.44	.29
12. Good place to relax	.58	.64	.60	.60	.63	.64
13. Telephone takes up time	--	--	--	--	--	--
14. Atmosphere is calm	.50	.63	.51	.67	.62	.70
15. Regular routine	--	--	--	--	--	--

Appendix E

Fit Indices for One-Factor 13-item CHAOS Model and Subsequent Modifications

	Fit Indices									
Modification	χ^2	DF	SRMR	RMSEA	GFI	CFI	NFI	TLI	AIC	BCC
1) CHAOS 1factor	662.29	130	.06	.06	.92	.81	.78	.78	766.29	769.12
2) e3 ↔ e5	488.75	128	.05	.05	.94	.87	.84	.85	596.75	599.69
3) e12 ↔ e14	438.68	126	.05	.05	.94	.89	.86	.87	550.68	553.73
4) e2 ↔ e4	387.11	124	.049	.04	.95	.91	.87	.88	503.11	506.26
5) e1 ↔ e9	354.90	122	.05	.04	.96	.92	.88	.90	474.90	478.16
6) 5 → 4	314.37	120	.04	.04	.96	.93	.90	.91	438.37	441.74

7) e5 ⟺ e9	297.14	118	.04	.04	.96	.94	.90	.92	425.14	428.62
8) e1 ⟺ e5	279.11	116	.04	.04	.96	.94	.91	.92	411.11	414.70
9) 9 → 14	256.56	114	.04	.03	.97	.95	.92	.93	392.56	396.26
10) e2 ⟺ e12	240.47	112	.04	.03	.97	.96	.92	.94	380.47	384.28
11) e10 ⟺ e14	225.62	110	.04	.03	.97	.96	.93	.94	369.62	373.53

Appendix F

Fit Indices for One-Factor 13-Item Model to Determine a Baseline Model

Model	χ^2	Df	CFI	RMSEA	RMSEA CI	ECVI	SRMR
Low Income							
1. Hypothesized one-factor model	406.53	65	.89	.08	.08, .09	.61	.06
2. Model 1 with One Error Covariance Specified (Items 3, 5)	311.28	64	.87	.07	.06, .08	.49	.05
Middle/High Income							
1. Hypothesized one-factor Model	255.56	65	.81	.08	.07, .10	.75	.06
2. Model 1 with One Error Covariance Specified (Items 3, 5)	177.26	64	.88	.07	.05, .08	.56	.06

Appendix G

Goodness-of-Fit Statistics for Tests of Multi-group One-Factor 13-Item Model Invariance: A Summary

Model Description	Equality Constraints Imposed	χ^2	df	$\Delta\chi^2$	Δdf	*p*	CFI	ΔCFI
1. Configural Model	none	488.75	128	--	--	--	.87	--
2. Model A	all	564.20	141	75.45	13	< .001	.85	.02
Model B	Item 1	490.06	129	1.31	1	NS	.87	.00
Model C	Items 1, 2	496.14	130	7.39	2	< .05	.87	.00
Model D	Items 1, 3	490.19	130	1.44	2	NS	.87	.00
Model E	Items 1, 3, 4	495.34	131	6.58	3	NS	.87	.00
Model F	Items 1, 3, 4, 5	499.33	132	10.58	4	< .05	.87	.00
Model G	Items 1, 3, 4, 6	496.87	132	8.11	4	NS	.87	.00
Model H	Items 1, 3, 4, 6, 7	497.50	133	8.74	5	NS	.87	.00
Model I	Items 1, 3, 4, 6, 7, 8	499.12	134	10.37	6	NS	.87	.00
Model J	Items 1, 3, 4, 6, 7, 8, 9	530.65	135	41.90	7	< .001	.86	.01
Model K	Items 1, 3, 4, 6, 7, 8, 10	500.51	135	11.76	7	NS	.87	.00
Model L	Items 1, 3, 4, 6, 7, 8, 10, 11	517.75	136	29.00	8	< .001	.87	.01
Model M	Items 1, 3, 4, 6, 7, 8, 10, 12	500.52	136	11.77	8	NS	.87	.00
Model N	Items 1, 3, 4, 6, 7, 8, 10, 12, 14	511.60	137	22.85	9	< .01	.87	.01

Note: Each model is compared to the configural model

Appendix H

Fit Indices for Two-Factor 13-Item CHAOS Model Subsequent Modifications

	Fit Indices									
Modification	χ^2	DF	SRMR	RMSEA	GFI	CFI	NFI	TLI	AIC	BCC
1) CHAOS 13 items 2-factor	597.90	128	.06	.06	.92	.84	.80	.80	705.90	708.83
2) e3 ⇔ e5	426.94	126	.05	.05	.94	.90	.86	.87	538.94	541.99
3) e2 ⇔ e4	374.61	124	.05	.04	.95	.91	.88	.89	490.61	493.76
4) 5 → 4	327.50	122	.04	.04	.96	.93	.89	.91	447.50	450.77
5) e9 ⇔ e5	307.21	120	.04	.04	.96	.94	.90	.92	431.21	434.58
6) e1 ⇔ e9	287.41	118	.04	.04	.96	.94	.91	.92	415.41	418.88

7) 9 → 14	261.69	116	.03	.03	.97	.95	.91	.93	393.69	397.28
8) e2 ⟺ e12	238.51	114	.03	.03	.97	.96	.92	.94	374.51	378.21
9) e1 ⟺ e5	222.62	112	.03	.03	.97	.96	.93	.95	362.62	366.42
10) e14 ⟺ e10	207.69	110	.03	.03	.97	.97	.93	.95	351.69	355.60

Appendix I

Fit Indices for Two-Factor 13-item Model to Determine a Baseline Model

Model	χ^2	Df	CFI	RMSEA	RMSEA CI	ECVI	SRMR
Low Income							
1. Hypothesized one-factor model	345.01	64	.85	.08	.07, .08	.53	.06
2. Model 1 with One Error Covariance Specified (Items 3, 5)	250.51	63	.90	.06	.06, .07	.41	.05
Middle/High Income							
1. Hypothesized one-factor Model	252.66	64	.81	.08	.07, .09	.10	.06
2. Model 1 with One Error Covariance Specified (Items 3, 5)	176.20	63	.88	.07	.06, .08	.08	.06

Appendix J

Goodness-of-Fit Statistics for Tests of Multigroup Two-factor 13-Item Model Invariance: A Summary

Model Description	Equality Constraints Imposed	χ^2	df	$\Delta\chi^2$	Δdf	p	CFI	ΔCFI
1. Configural Model	none	426.916	126	--	--	--	.90	--
2. Model A	all	504.45	139	77.53	13	< .001	.87	.02
Model B	Factor 2	489.21	133	62.29	7	< .001	.88	.02
Model C	Factor 1	445.58	132	18.70	6	< .01	.89	.01
Model D	Item 1	427.27	127	.36	1	NS	.90	.00
Model E	Items 1, 2	433.18	128	6.26	2	< .05	.89	.00
Model F	Items 1, 4	432.95	128	6.03	2	< .05	.89	.00
Model G	Items 1, 7	428.84	128	1.83	2	NS	.90	.00
Model H	Items 1, 7, 12	429.56	129	2.65	3	NS	.90	.00
Model I	Items 1, 7, 12, 14	434.98	130	8.07	4	NS	.89	.00
Model J	Items 1, 7, 12, 14, 3	435.31	131	8.39	5	NS	.89	.00
Model K	Items 1, 7, 12, 14, 3, 5	440.04	132	13.13	6	< .05	.89	.00
Model L	Items 1, 7, 12, 14, 3, 6	438.97	132	11.78	6	NS	.89	.00
Model M	Items 1, 7, 12, 14 3, 6, 8	439.62	133	12.71	7	NS	.89	.00
Model N	Items 1, 7, 12, 14, 3, 6, 8, 9	475.84	134	48.92	8	< .001	.88	.01
Model O	Items 1, 7, 12, 14, 3, 6, 8, 10	442.00	134	15.08	8	NS	.89	.00
Model P	Items 1, 7, 12, 14, 3, 6, 8, 10, 11	463.72	135	36.81	9	< .001	.89	.01
Structural model	Covariance between two latent factors	519.76	140	92.84	14	< .001	.87	.03

Note: Each model is compared to the configural model.

VITA

Sarah Wollersheim Shervey
Saw325@psu.edu
651-492-9493

EDUCATION

The Pennsylvania State University

Graduate Student, Doctoral Candidate 8/08-8/13

Currently completing graduate studies in the
School Psychology Ph.D. Program (5th year)
Received M.Ed. (Spring 2010), Ph.D. (exp. August 2013)

PROFESSIONAL CERTIFICATION

Pennsylvania

Certified School Psychologist 5/12-present

EMPLOYMENT EXPERIENCES

Sarah Reed Children's Center

APA-approved pre-doctoral intern 8/12-7/13

Conducting therapy, psychoeducational evaluations, and consultation at a residential treatment facility and public elementary school in Erie

The Pennsylvania State University, The Prevention Research Center

Graduate Research Assistant, Family Life Project 5/09-8/12

Collect data and conducted statistical analyses for longitudinal research project.

The Pennsylvania State University, CEDAR Mobile Clinic

Graduate Research Assisstant, Dr. James DiPerna 8/08- 6/09

Observed the impact of Linking Energizers to Academic Performance (Project LEAP) on subsequent learning and behavior

PRACTICUM EXPERIENCES

The Meadows Psychiatric Hospital

Psychology Practicum Student 10/11-5/12

Facilitated and co-facilitated individual and group therapy sessions and completed psychosocial evaluations

State College Area School District

School Psychology Practicum Student 8/10-6/11

Conducted Response to Intervention (RTI) reading groups, administered AIMSweb progress monitoring assessments, completed psycho-educational evaluations of students in preschool, elementary, and secondary grades

The Pennsylvania State University CEDAR School Psychology Clinic

Student Clinician and Supervisor 6/09 – 5/12

Conducted and supervised psycho-educational evaluations to determine appropriate psychological and educational services