UC Berkeley
UC Berkeley Electronic Theses and Dissertations

Title

Do Strengths Matter?: The Role of Strengths in Improving Mental Health Outcomes for Youth with Emotional and Behavioral Problems In Contact with a Public System of Care

Permalink

https://escholarship.org/uc/item/9mn2d86c

Author

Accomazzo, Sarah Maria

Publication Date

2014

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

Do Strengths Matter?:
The Role of Strengths in Improving Mental Health Outcomes for Youth with Emotional and Behavioral Problems In Contact with a Public System of Care

By

Sarah Maria Accomazzo

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Social Welfare

in the

Graduate Division

of the

University of California, Berkeley

Committee in Charge:

Professor Eileen Gambrill, Chair
Professor Steven P. Segal
Professor Valerie Shapiro
Professor Sophia Rabe-Hesketh

Fall 2014

Do Strengths Matter?:
The Role of Strengths in Improving Mental Health Outcomes for Youth with Emotional and Behavioral Problems In Contact with a Public System of Care

© 2014

by

Sarah Maria Accomazzo

Abstract

Do Strengths Matter?:

The Role of Strengths in Improving Mental Health Outcomes for

Youth with Emotional and Behavioral Problems

In Contact with a Public System-of-Care

By

Sarah Maria Accomazzo

Doctor of Philosophy in Social Welfare

University of California, Berkeley

Professor Eileen Gambrill, Chair

The aim of this dissertation was to examine the extent of association, if any, between strengths (including those of the client, the client's family, and the client's environment) and psychiatric symptom severity for "youth" (children and adolescents aged 5-22) in contact with a mental health and/or substance abuse agency in a public System of Care. This dissertation examined a public systems dataset that included administrative assessment data for every client aged 5-22 (n=2049) who entered into San Francisco's Children's System of Care (SF-CSOC) from June 2010-August 2013. The assessment data from SF-CSOC was collected by clinicians at entry to the system (initial assessment), six months after initial assessment, and twelve months after initial assessment. Behavioral/social learning theory and related research in applied and experimental settings (e.g. Goldiamond, 1974; Madden, 2013; Staats, 2012) would suggest that growth in strengths would be associated with reduction in symptoms. However, time order cannot be determined with the current data and thus only associational relationships are discussed.

This dissertation is a "three paper" dissertation including a dissertation introduction, three inter-related papers, and a dissertation conclusion. The dissertation introduction provides a background to the entire dissertation and is intended to orient to the reader to terminological and conceptual issues relevant to all three papers (each paper might also cover some of the literature included in the dissertation introduction, as relevant to the specific paper).

Paper one reviews the literature on prevalence of emotional and behavioral problems among youth in the United States, orients the reader to the history and current policies of Systems of Care (SOC), and reviews the literature on strength-based assessment. Then, paper one reports the prevalence of clinician-reported strengths in San Francisco's public System of Care (SF-SOC), explores the psychometric properties of San Francisco's Child and Adolescent Strengths and Needs (SF-CANS) assessment's Strengths section, and describes three alternative modeling options to explore group differences (by gender, age, and ethnicity) of youth strengths.

Youth assessed at entry to San Francisco's System of Care were reported to enter with considerable strengths, similar to other studies reporting on prevalence of strengths in a System of Care (e.g. CMHI, 2011; Dunleavey et al., 2011). The confirmatory factor analysis found that the hypothesized two factor model was a better fit for the data than a single factor model, though the fit of the two factor model was only adequate and the interpretation of the factors were questionable. Individual strength items were assessed as the most informative for future analyses, compared to the factors or the composite score.

Paper two examines whether, at initial assessment, client strengths are associated with psychiatric symptoms. Higher levels of family relationships, peer relationships, and educational strengths were significantly associated with lower levels of psychiatric symptoms, after accounting for variation due to clinician, program, reported trauma exposure and youth demographics.

Paper three reviews some studies that report on the relationship between strengths and mental health outcomes in Systems of Care samples. This paper then examines 1) the average change in psychiatric symptom severity twelve months after entry (initial assessment) to San Francisco's Children's System of Care (SF-CSOC), 2) whether initial strengths are associated with rate of change in psychiatric symptoms, 3) whether rate of change in strengths at six month and twelve month assessment are associated with rate of change in psychiatric symptoms, and 4) explores the contribution of variations related to clinician, youth, and time. Results suggest that self-reported growth in peer relationships, relationship permanence, and education strengths twelve months after entry to SF-CSOC was associated with larger reported reductions in symptoms at twelve months, adjusting for clustering, time, and youth demographics. Only associational relationships are reported; causal direction cannot be established in the current data. The percentage of youth referred for services, where they were referred from and to, the percentage who went to services, type, quantity, and quality of interventions offered and used (if any), and drop-out rates and reasons are not reliably tracked in SF-CSOC.

This dissertation focuses on one aspect of service delivery in public mental health systems, namely the role of strengths, if any, in Systems of Care for youth with emotional and behavioral problems. The dissertation conclusion discusses overall themes that emerged, notes limitations relevant to the entire dissertation, and discusses implications for social workers and others.

Dedication

In honor of my loved ones.

You celebrate life's joys and face life's challenges with passion and courage.

You have taught me so much.

I am grateful for you.

Table of Contents

Introduction to the Dissertation: Youth with Emotional and Behavior Problems in a System of Care (SOC): How are Strengths Viewed?.....1

Paper 1: Strengths of Youth in San Francisco's System of Care: Measurement Questions and Implications for Analysis…….23

Introduction
Method
Results
Discussion

Paper 2: Strengths and Psychiatric Symptom Severity of Youth Entering San Francisco's Public System of Care……..45

Introduction
Method
Results
Discussion

Paper 3: Youth Strengths, Psychiatric Symptom Severity, and Clinician/Client Effects Six and Twelve Months after Initial Assessment in San Francisco's Public System of Care……...62

Introduction
Method
Results
Discussion

Conclusion of the Dissertation……92

References…….....97

Appendices……..103

INTRODUCTION TO THE DISSERTATION

Youth with Emotional and Behavioral Problems in a System of Care:
How Are Strengths Viewed?

Sarah Accomazzo

School of Social Welfare
University of California, Berkeley

Introduction Contents

Youth with Emotional and Behavior Problems in a System of Care (SOC): How are Strengths Viewed?

"Mental Health" and the Medicalization of Misery……...3
Terminology
Towards a Definition of a Strength
A Note on the Directionality of the Relationship between Client Strengths and Psychiatric Symptoms

The Federal Policy Response to Youth with Emotional and Behavior Problems: The System of Care (SOC) Concept and a Focus on Strengths……..6
The Development of the Systems-of-Care (SOC) Concept.
System of Care (SOC) Concept: Current Definition, Core Values, and Principles
Evaluating the System of Care (SOC) Approach
A Focus on Strengths in the System of Care Concept and Related Policy Documents

Theoretical conceptualizations of Strengths in Systems of Care (SOC): The Under-Valued Contribution of Behavioral Theory and Related Research…….14
Behavioral Theory.
Behavioral Theory's Constructional Approach and Implications for the Directionality of the Relationship Between Strengths and Psychiatric Symptoms in Systems of Care (SOC)

Measuring Youth Strengths in a System of Care: Strengths-based Assessment, the CANS, and San Francisco's Children's System of Care……...15
The Behavioral and Emotional Rating Scale (BERS).
The Child and Adolescent Needs and Strengths (CANS) Assessment
Limitations of the CANS
Dissertation Setting: San Francisco's Children's System of Care (SF-CSOC) and the SF-CANS
SF-CSOC
SF-CANS
Limitations of SF-CANS
Limitations of Administrative Data

In Summary……..22

Youth with Emotional and Behavior Problems in a System of Care (SOC): How Are Strengths Viewed?

The aim of this dissertation was to examine the extent of association, if any, between strengths (including those of the client, the client's family, and the client's environment) and psychiatric symptom severity for "youth" (children and adolescents aged 5-22) in contact with a mental health and/or substance abuse agency in a public System of Care. This dissertation examined a public systems dataset that included administrative assessment data for every client aged 5-22 (n=2049) who entered San Francisco's Children's System of Care (SF-CSOC) from June 2010-August 2013. Assessment data were collected by clinicians at initial assessment (entry to the system), six months after initial assessment, and twelve months after initial assessment. Behavioral/social learning theory and related research in applied and experimental settings (e.g. Goldiamond, 1974; Staats, 2012) would suggest that growth in strengths would be associated with reductions in symptoms. However, given the data used, time order cannot be established and directionality cannot be determined. Also, several variables of interest, such as use of medication and interventions actually received, were not available in this administrative data.

This dissertation is a "three paper" dissertation which includes a dissertation introduction, three inter-related papers, and a dissertation conclusion. The dissertation introduction provides a background to the entire dissertation and is intended to orient to the reader to various terminological and conceptual issues relevant to all three papers (each paper might also include some of the information from the dissertation introduction, as is relevant to the specific paper). Paper one describes the prevalence of reported strengths in San Francisco's public System of Care, explores the psychometric properties of the assessment's Strengths domain to learn if the domain is unidimensional or multidimensional, and describes three alternative modeling options to explore group differences (by gender, age, and ethnicity) of youth strengths. Paper two examines whether, at initial assessment, client strengths are associated with psychiatric symptoms. Paper three explores: 1) the average estimated change in psychiatric symptom severity within twelve months of initial assessment, 2) whether initial strengths are associated with rate of change in psychiatric severity within twelve months of initial assessment, 3) the contribution to variation by clinician, youth, and time, and 4) whether change in strengths over twelve months is associated with change in psychiatric symptoms. The dissertation conclusion discusses overall themes that emerged from the findings of the three papers, limitations that are relevant to the entire dissertation, and implications for social workers and others.

"Mental Health" and the Medicalization of Misery.

There are two main orientations to understanding and intervening with emotional and behavioral problems. One approach views deviant behavior as pathological and focuses on eliminating, alleviating, or otherwise overcoming distress, as in psychiatric approaches. A constructional approach views deviant behavior as functional and focuses on understanding these functions and drawing on and creating positive repertoires of behavior (e.g. Madden, 2013; Schwartz & Goldiamond, 1975). There has long been a medicalization of distress, encouraged by the "biomedical industrial complex" (Conrad, 2007; Gomory, Wong, Cohen, & Lacasse, 2011; Szasz, 1961; Szasz, 2007). The latter promotes an emphasis on brain malfunction, overlooking environmental and social factors. The bio-medical industrial complex has monopolized mental health and substance abuse services. The biomedical industrial complex is defined as:

> reinforcing and interlocking connections between the pharmaceutical, biotechnological, and medical industries that – together with academic experts in the helping professions, governmental funding and regulatory bureaucracies…and professional and family lobbies – that promote and support a biomedical model of psychosocial distress and disability" (Gomory, et al., 2011, pg. 137).

Psychosocial distress and misbehavior (deviant or socially disruptive behavior) are attributed to biological brain-based disorders, while environmental and social factors are downplayed (e.g. Szasz, 1987, 2007).

The biomedical industrial complex is exemplified by the reliance, in the United States, on the Diagnostic and Statistical Manual of Mental Disorders, now in its fifth edition (DSM-5, American Psychological Association, 2013). The DSM furthers the medicalization of distress by "supporting the positive psychiatric project of codifying human suffering into disease-like categories" (Rapley, Moncrieff, & Dillon, 2011, pg. 1). The DSM's use of the term "mental disorder" (a pseudonym for mental illness) frames human distress and troubling and deviant behavior within a psychiatric, pathological paradigm. Public funding streams in the United States, such as Supplemental Security Income (SSI) and Medicare, require the receipt of an official DSM diagnosis. A survey of 69 social work psychopathology course syllabi from 58 schools of social work found that only six (8.5%) included any assignments or readings that were critical of the DSM or a biomedical approach to classification of mental distress and maladaptive behavior (Lacasse & Gomory, 2003). Despite increasing critiques (e.g. Conrad, 2007; Gambrill, 2013a; Illouz, 2008; Kutchins & Kirk, 1997; Szasz, 1987, 2007), including rejection of the DSM as a classification system and a call for a totally brain-based classification system from the Director of the National Institute of Mental Health (Insel, 2011). (See also Frances, 2010.) The DSM remains the predominant classification system used to label adult and youth emotional and behavioral problems in the United States. The focus on strengths in federal youth mental health policy (e.g. Liu, 2011) developed, in part, as a reaction to an overemphasis, especially in American culture, on pathology and related deficit-based views of deviant behavior.

Terminology.

Many of the terms used in this dissertation are 1) referenced across a variety of literatures, but defined differently in each literature, 2) used inconsistently or unclearly defined within one particular literature, and 3) used interchangeably with terms that are conceptually related but have different meanings. In order to avoid conceptual confusion, this dissertation attempts to be clear about the definition in use in each section of the paper. This dissertation uses "emotional and behavior problems" to describe the expression, through emotions and behaviors, of distress, madness, or suffering, that disrupt youth functioning or development in one or more major life domains (e.g. family, school, peer relationships). This dissertation uses "mental health" to describe an overall experience of emotional and behavioral well-being, not merely the absence of disorder (WHO, 2014). Expanded discussions of other key terms and concepts are provided throughout this introduction in order to orient the reader. The definition of a "strength" is discussed below in detail due to its relevance to the rest of the dissertation.

Towards a Definition of a Strength. Terminological consensus has not been reached as far as the definition of a "strength." Conceptually related research has been conducted by researchers in different disciplines using different terminology. Researchers in the areas of behavioral theory and applied behavioral analysis (e.g. Goldiamond, 1974; Madden, 2013;

Staats, 2012), resilience (e.g. Masten, 2007), positive youth development (e.g. Catalano, Berglund, Ryan, Lonczak, & Hawkins, 2004), developmental psychopathology (e.g. Luthar & Brown, 2007), social policy (e.g. Jenson, Alter, Nicotera, & Forrest-Bank, 2013), youth violence (e.g. Ozer, Richards, & Kliewer, 2004), and public mental health services (e.g. Graves, 2007; Lyons, 2009) have all conducted research related to the role of positive youth attributes, skills, and environments in protecting against the negative effects of risks and/or promoting health. Though conceptually related, these literatures use different terminology (e.g. strengths, alternative available repertoires, new alternative repertoires, resilience, assets, resources, competencies, protective factors, promotive factors), resulting in some conceptual confusion. (For example, see Probst (2009) for a discussion of terminological confusion in the strengths-based practice literature; see Luthar (2000) and Fergus & Zimmerman (2006)'s critique of the resilience terminology).

This dissertation defines youth strength as: a characteristic, attribute, repertoire, ability, thought, skill, behavior, or resource, located at the individual, interpersonal, or environmental level, that may reinforce and/or construct desired behaviors, protect against the negative effects of risk, and/or promote health and well-being. Nearly anything can be a strength, and a strength for one individual is not necessarily a strength for another individual. Strengths may change over time, interact with each other, and/or have different impacts over time and contexts. In the context of mental health service delivery for children and adolescents with emotional and behavioral problems, strengths may be discussed as 1) characteristics for assessment that describe the youth's current repertoire of strengths 2) characteristics for intervention that may help a child or adolescent develop and thrive, or 3) service delivery outcomes.

Examples of strengths include (but are not limited to) insight into oneself, resources and relationships available in the youth and family's social network and community, talents, cultural identity and practices, spirituality, high levels of commitment to school, opportunities for employment in the community. (For more examples, see Crews et al., 2007; Gambrill, 2013b; Goldiamond, 1974; Jenson & Frasier, 2011; Layng, 2009; National Wraparound Initiative website, www.nwi.pdx.edu/wraparoundbasics.shtml; Saleebey, 2009; Schwartz & Goldiamond, 1975; Staats, 2012.)

A note on the directionality of the relationship between client strengths and psychiatric symptom severity. Behavioral/social learning theory and related research in applied and experimental settings would suggest a directional relationship between strengths and symptoms, where an increase in strengths results in a decrease in symptoms. For example, in paper three, it was expected that an increase in reported youth strength items between initial assessment and twelve month assessment would be associated with a reduction in psychiatric symptoms, adjusting for other covariates. The opposite directional model is a possibility, for example, that a reduction in psychiatric symptoms increases strengths. Perhaps strengths and symptoms operate in both directions, contingent on a given context. It is not possible to establish directionality in this dissertation, due to the nature of the data; this administrative dataset was collected for non-research purposes through clinician self-report, without randomization or control procedures. Since information on strengths and symptoms was collected at the same time, order cannot be determined. Also, the dataset is missing many variables that might impact both strengths and symptoms, such as medication used and interventions offered and engaged in, and these could be responsible for the reduction in symptoms. Future analyses could explore statistical evidence for directionality, for example by using a cross-lagged panel model (e.g. Finkel, 1995) or propensity score matching techniques (e.g. Rosenbaum & Rubin, 1983).

The Federal Policy Response to Youth with Severe Emotional and Behavior Problems: The System of Care (SOC) Concept and a Focus on Strengths

Emotional and behavioral problems are common during childhood and adolescence. In the United States, it is estimated that one-third to one-half of young people will meet criteria for a diagnosis as described in the Diagnostic Statistical Manual-IV-TR (2000) at least once before the age of 21 (e.g. Costello, Erkanli, Fairbank, & Angold, 2003; O'Connell et al., 2009; Kessler et al., 2012). (The Diagnostic Statistical Manual (DSM; 2014) is the predominant classification system for emotional and behavioral problems in the United States, as discussed above.) Related research suggests that approximately 10% of youth in the United States will qualify as having "severe" emotional and behavioral problems (resulting in severe impairments or disruptions in one or more life functioning domains or developmental tasks) during any given year (Kessler et al., 2012; Merikangas et al., 2010).

The federal government has faced criticism for its lack of a comprehensive legislative strategy to address severe emotional and behavioral problems among youth as a major public health threat (e.g. Knitzer, 1982; Lourie & Hernandez, 2003). When the total annual economic costs of mental, emotional, and behavioral problems for persons under 25 are calculated to include burden of disease, crime-related costs, and health services expenditures, is estimated that emotional and behavioral problems in childhood and adolescence cost the public upwards of $200 billion each year, or about $2,380 per individual, including "about $500 in health service costs and $1,900 in health, productivity, and crime-related costs" (O'Connell et al., 2009, pg. 231). Though service rates have overall increased in the past ten years, it is estimated that approximately half of youth who meet criteria for an emotional or behavioral problem receive inadequate or no mental health intervention at all (Kessler et al., 2012; Merikangas et al., 2010; President's New Freedom Commission, 2003). It has been estimated that 28%-85% of youth drop out "prematurely" (or before the main problem has been reported resolved by youth, caregiver, or clinician) (e.g. Garcia & Weisz, 2002).

Over the past thirty years, the cornerstone of the federal government's attempts to respond to children and adolescents ("youth") with emotional and behavioral problems has been the development of the System of Care (SOC) approach (e.g. Stroul & Friedman, 1986; CMHI, 2011). Intended as a guiding philosophy and values framework for mental health and substance abuse service provision for youth with emotional and behavioral problems, the SOC approach has been implemented in 173 communities throughout the United States and its territories since 1984, with a total of approximately two billion dollars in federal funding. To the knowledge of this author, there has not been a meta-analysis or systematic review of the effectiveness of SOCs in improving outcomes for youth and their families, and individual studies have had mixed results (e.g. Bickman, 1996; CMHI, 2011; Foster, Stephens, Krivelyova, & Gamfi, 2007), as discussed below. The following sections review the history and current definition, values, and principles of the System of Care (SOC) concept.

The Development of the Systems of Care (SOC) Concept.

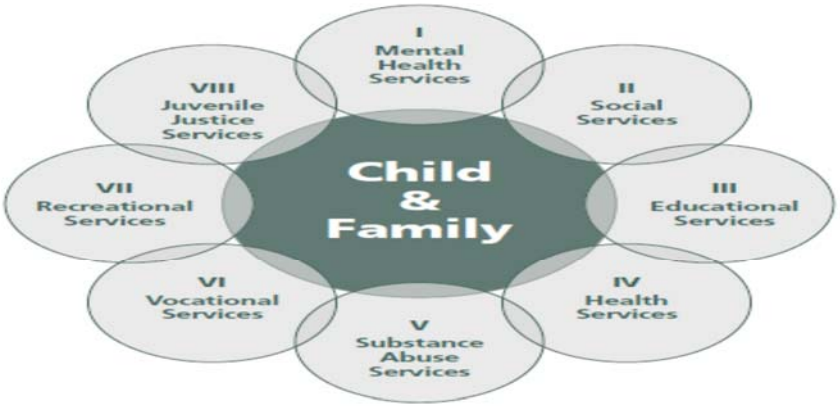
In 1982, Jane Knitzer and colleagues, with funding from the Children's Defense Fund (CDF), published a policy report entitled, *Unclaimed Children: the Failure of Public Responsibility to Children in Need of Mental Health Services.* The authors sent surveys to state mental health officials in fifty states, conducted interviews with various stakeholders, including youth, family members, and the leadership of state mental health agencies, and analyzed existing state and federal policies. The report featured profiles of youth who were "seriously disturbed,"

for example, youth whose emotional and behavioral problems resulted in their involvement in a public service system, and who typically faced multiple other issues, including prior hospitalizations, involvement in child welfare or juvenile justice settings, poverty, developmental disabilities, discrimination, or scarce environmental resources (Knitzer, 1982, pg. x). The final report concluded that youth with severe emotional and behavioral problems and their families had been abandoned by health professionals and policy makers to a fragmented array of ineffective, and even dangerous, mental health services. The report called on public agencies to "reclaim responsibility" (pg. xiii) for this population.

In 1984, in response to *Unclaimed Children* and other advocacy efforts, the US Congress appropriated funds through Public Law 102-321 and created the Children and Adolescent Service System Program (CASSP), designed to address youth mental health issues (Stroul, Blau, & Friedman, 2010). CASSP initiated "a participatory process, involving multiple and diverse stakeholders including policy makers, service providers, agency administrators, technical assistance providers, family members, advocates, leaders in cultural competence, researchers, and others" (Stroul et al., 2010, pg. 1) to determine a course of action for the federal response, out of which grew the System of Care concept. The original "System of Care" (SOC) concept, core values, and philosophy were further articulated in Stroul & Friedman's (1986, 1994) book, *A System of Care for seriously emotionally disturbed children and youth.*

Stroul & Friedman (1986, 1994) specified a System of Care (SOC) as a framework to guide mental health service delivery for children and adolescents with severe emotional and behavioral problems (Stroul et al., 2010). The goal was to "reform child-serving services" (pg. 1) and provide more effective and efficient services that adhered to a specific philosophy and values, compared to the existing, fragmented array of services (Stroul et al., 2010). Instead of specifying procedures or intervention steps, the System of Care concept provided a set of core values and principles that Stroul & Friedman (1986) claimed could be flexibly applied to different community contexts, regardless of types of services, amount of services, and system, agency, provider, and client characteristics. The original SOC concept emphasized the importance of integrating and linking different youth service systems, encouraged partnerships with youth, families, and communities, and provided care in the least restrictive environment (Stroul & Friedman, 1986). The 1994 edition additionally emphasized "child-centered" and "community-based" approaches, family collaboration, and community care focused on youth strengths instead of disabilities (language drawn in part from the Disability Rights movement). (See figure one for the original System of Care framework provided in Stroul & Friedman (1986, pg. 30).

Figure One. System of Care Framework from Stroul & Friedman (1986).



From Stroul, B. & Friedman, R. (1986 rev ed). *A system of care for children and youth with severe emotional disturbances* (rev. ed., p. 30). Washington, DC: Georgetown University Child Development Center, National Technical Assistance Center for Children's Mental Health.

In 1992, Congress amended Public Law 102-321 and created the Substance Abuse and Mental Health Services Administration (SAMHSA), housed within the U.S. Department of Health and Human Services, to oversee mental health and substance abuse efforts in the United States. Since then, the Center for Mental Health Services (CMHS) (under SAMHSA) has run the Comprehensive Children and Adolescents Community Mental Health Services Program, also known as the Children's Mental Health Initiative (CMHI) (CMHI, 2011, pg. i). CMHI awards matching System of Care (SOC) grants to "states, local communities, U.S. territories, Indian tribes and tribal organizations, and other American Indian/Alaska Native communities" (pg. i) to develop, implement, sustain, and evaluate SOCs in their communities. Eligible communities can receive SOC funding for up to six years, with the federal government matching three dollars to every one local dollar for the first three years, a dollar for dollar match in the fourth year, and a two to one dollar match in the fifth and sixth year. Since 1992, 173 communities have received CMHI funding to implement a SOC in their community. For example, San Francisco City and County received the six year grant from 2002-2008.

The System of Care (SOC) Concept: Current Definition, Core Values, and Principles.

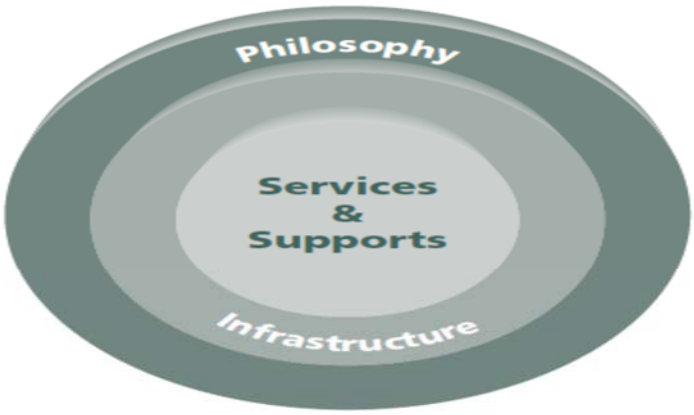
An updated System of Care (SOC) definition, core values, and principles was developed over the past fifteen years. In 2005, a "consortium of professional, family members, and youth" (Stroul, Friedman, & Sondheimer, 2008, pg. 4) met to begin the official process. A special issue of the *Journal of Evaluation and Program Planning* (Hodges & Ferreira, 2010) was devoted to debating the updated concept, reviewing the empirical literature related to SOCs, and offering suggestions for next steps in implementation and evaluation. Based on these and other efforts, in 2010, the Georgetown University Center for Child and Human Development and the Substance Abuse and Mental Health Services Administration (SAMHSA) released an policy brief (Stroul et al., 2010), entitled "Updating the System of Care Concept and Philosophy."

The updated definition, core values, and guiding principles expanded on the original concept in several ways (see Figure 2 for the updated SOC diagram). For example, the updated concept attempted to be more inclusive of youth with a range of emotional and behavioral issues by changing the language from "severe emotional disturbance" to "severe mental health challenges." The Children's Mental Health Initiative (CMHI) National Evaluation Annual Report (2011) defined "severe mental health challenges" as:

> an emotional, socio-emotional, behavioral, or mental disorder diagnosable under the *DSM–IV–TR* or its *ICD–9–CM* equivalents, or *the Diagnostic Classification of Mental Health and Developmental Disorders of Infancy and Early Childhood–Revised* (DC:0–3R);
> a disability in functioning at home, school, or community, or in a combination of these settings; or a level of functioning such that intervention is required that involves two or more community agencies providing services in the areas of mental health, education, child welfare, juvenile justice, substance abuse, or primary health care; the identified disability present for at least 1 year or, on the basis of diagnosis, severity, or multiagency intervention (Appendix C, pg. 2).

In order to demonstrate an increased commitment to family and youth collaboration and partnership, the language was updated from "child centered and youth focused" to "family-driven and youth-guided." To move away from a medical model language of "medical necessity," youth and family "strengths and needs" was emphasized. Other changes included an attempt to increase the focus on cultural and linguistic issues, emphasizing the role of "natural supports" in addition to the more traditional health services, and to emphasize the challenges of comparing SOCs as a single unit (see Stroul et al., 2010, pgs. 3-5, for a complete summary of the 2010 updates to the 1986 SOC definition, core values, and principles).

Figure 2. Updated System of Care Concept.



From Stroul, B. (2002). *Issue Brief: Systems of care: A framework for systems reform in children's mental health.* Washington, DC: Georgetown University Center for Child and Human Development, National Technical Assistance Center for Children's Mental Health.

The System of Care (SOC) concept was summarized in a federal policy brief (Stroul et al., 2010). The updated definition is:

> A spectrum of effective, community-based services and supports for children and youth with or at risk for mental health or other challenges and their families, that is organized into a coordinated network, builds meaningful partnerships with families and youth, and addresses their cultural and linguistic needs, in order to help them to function better at home, in school, in the community, and throughout life (Stroul, Blau, & Friedman, pg. 6).

The policy brief also articulated three updated Core Values to which SOCs must adhere:

1. Family driven and youth guided, with the strengths and needs of the child and family determining the types and mix of services and supports provided. 2. Community based, with the locus of services as well as system management resting within a supportive, adaptive infrastructure of structures, processes, and relationships at the community level. 3. Culturally and linguistically competent, with agencies, programs, and services that reflect the cultural, racial, ethnic, and linguistic differences of the populations they serve to facilitate (pg. 6).

The policy brief also articulated thirteen updated Guiding Principles on which SOC service implementation and delivery should be based:

> Systems of Care are designed to:
> 1. Ensure availability and access to a broad, flexible array of effective, community-based services and supports for children and their families that address their emotional, social, educational, and physical needs, including traditional and nontraditional services as well as natural and informal supports.
> 2. Provide individualized services in accordance with the unique potentials and needs of each child and family, guided by a strengths-based, wraparound service planning process and an individualized service plan developed in true partnership with the child and family.
> 3. Ensure that services and supports include evidence-informed and promising practices, as well as interventions supported by practice-based evidence, to ensure the effectiveness of services and improve outcomes for children and their families.
> 4. Deliver services and supports within the least restrictive, most normative environments that are clinically appropriate.
> 5. Ensure that families, other caregivers, and youth are full partners in all aspects of the planning and delivery of their own services and in the policies and procedures that govern care for all children and youth in their community, state, territory, tribe, and nation.
> 6. Ensure that services are integrated at the system level, with linkages between child-serving agencies and programs across administrative and funding boundaries and mechanisms for system-level management, coordination, and integrated care management.
> 7. Provide care management or similar mechanisms at the practice level to ensure that multiple services are delivered in a coordinated and therapeutic manner and that children and their families can move through the system of services in accordance with their changing needs.
> 8. Provide developmentally appropriate mental health services and supports that promote optimal social-emotional outcomes for young children and their families in their homes and community settings.
> 9. Provide developmentally appropriate services and supports to facilitate the transition of youth to adulthood and to the adult service system as needed.
> 10. Incorporate or link with mental health promotion, prevention, and early identification and intervention in order to improve long-term outcomes, including mechanisms to identify problems at an earlier stage and mental health promotion and prevention activities directed at all children and adolescents.

11. Incorporate continuous accountability and quality improvement mechanisms to track, monitor, and manage the achievement of System of Care goals; fidelity to the System of Care philosophy; and quality, effectiveness, and outcomes at the system level, practice level, and child and family level.
12. Protect the rights of children and families and promote effective advocacy efforts.
13. Provide services and supports without regard to race, religion, national origin, gender, gender expression, sexual orientation, physical disability, socio-economic status, geography, language, immigration status, or other characteristics, and ensure that services are sensitive and responsive to these differences (pg. 6).

Currently, there are two main websites devoted to the System of Care concept and its implementation and evaluation throughout the United States. The System of Care Home (http://www.tapartnership.org/systemsOfCare.php) is a resource for communities that are currently implementing a SOC or that are interested in implementing the SOC concept and philosophy in the future. For example, the website describes implementation, evaluation, and sustainability of SOCs and provides policy documents and practice resources on these topics. The Community Mental Health Initiative digital archive (http://www.cmhi-library.org/) is a centralized space for resources, reports, and policy documents.

Evaluating the System of Care (SOC) Approach.

Individual studies evaluating the System of Care (SOC) approach have had mixed results (e.g. Bickman, 1996; Bickman, Noser, & Summerfelt, 1999; CMHI, 2011; Foster et al., 2007). (To the knowledge of this author, there has not been a meta-analysis or a systematic review of the effectiveness of SOC communities in improving outcomes for youth and their families.) An early SOC comparison study (a study that makes at least some attempt to include a comparison group) compared three military areas in Fort Bragg, North Carolina (n=984), one of whom had replaced their existing children's health care framework with a System of Care approach (Bickman, 1996). The System of Care group reported increased use of services, less restrictive levels of care, and higher client satisfaction with services. However, the SOC and comparison groups reported the same rate of overall improvement on psychiatric symptoms, and the SOC reported higher costs than the non-SOC sites. In a five year follow up study, consistent with previous findings, no differences on overall improvement were seen between the groups (e.g. Bickman, Lambert, Andrade, & Penaloza, 2000). The Fort Bragg study was critiqued for lack of random assignment to a SOC approach or a non-SOC approach and failure to document type and effectiveness of services provided during the course of the study (Bickman, 1996). A follow-up study was conducted using random assignment of children and families (n=350) to either a pre-existing System of Care (SOC) or to a control group who sought services on their own (Bickman et al., 2000). This study found that youth and families received more case management and home visits in the SOC, but replicated that there were no significant differences between the two groups in psychiatric symptom reduction over time (Bickman, 2000).

The System of Care (SOC) approach was intended as a guiding framework, with each SOC adhering to a shared philosophy, core values, and guiding principles. Each SOC was intended to have a context-specific array of services and supports, tailored to meet its community's specific needs (Stroul et al., 2010). Perhaps other SOCs were more effective than those included in Bickman et al.'s (1996, 2000) studies. Alternatives for evaluating System of Care effectiveness have been proposed, including a focus on 1) the type and quality of services (e.g. Salzer & Bickman, 1999), 2) the implementation of, and fidelity to, SOC principles and

values (e.g. Stephens et al., 2004), 3) understanding and integrating common elements (e.g. Garland, Bickman, & Chorpita, 2010) and common factors (e.g. Wampold, 2010), and 4) compiling evidence of effectiveness through site-specific evaluations on a range of outcomes (e.g. Foster et al., 2007).

Since 2002, the Children's Mental Health Initiative (CMHI) has conducted an ongoing evaluation of funded System of Care (SOC) communities (n=173) (CMHI, 2011 pg. ii). The 2011 National Evaluation of SOCs reported data from sixty-eight SOC communities funded between 2006 and 2011. Youth from each SOC community are evaluated on multiple, specific outcomes, including emotional and behavior strengths (e.g. interpersonal skills, family involvement), emotional and behavioral problems/needs (e.g. anxiety and depressive symptoms, suicide attempts, contact with law enforcement, arrest history), education outcomes, physical health, and exposure to potentially traumatic events. Outcome data is compiled by CMHI's National Evaluation team. For example, 3,004 youth who entered a SOC in 2005-2006 were tracked over 36 months. Twenty-seven percent of these youth reported a significant reduction in their overall level of behavioral and emotional problems (as measured on the Child Behavioral Checklist, Achenbach & Rescorla, 2000) within six months and 40% reported a clinically significant reduction within thirty months (pg. 12). Caregivers reported that 25% of youth (as measured on the Behavioral and Emotional Rating Scale-2, Parent Report, Buckley & Epstein, 2004) had a significant increase in strengths within six months and 40% had a significant increase in thirty months. Regression to the mean likely contributed to these changes. The percentage of youth referred for services, where they were referred from and to, the percentage who went to services, the type, quantity, and quality of interventions offered and used (if any), and drop-out rates and reasons, are not reported in the Annual Reports. No data are provided on associations between caregiver and youth reports, which may differ. (National Evaluation results, including findings and limitations of the evaluation, are discussed in more detail in paper three.)

A Focus on Strengths in the System of Care Concept and Related Policy Documents.

Embedded in both the original and updated System of Care (SOC) concept is the notion that services for youth with emotional and behavioral problems should include a focus on strengths. The term "strength" is specifically used in the updated SOC concept (Stroul, et al., 2010). For example, the first Core Value states that SOCs should provide: "family-driven and youth-guided care, with the strengths and needs of the child and family determining the types and mix of services and supports provided" (Stroul et al., 2010, pg. 5). The Guiding Principles state that SOCs should "provide individualized services in accordance with the unique potentials and needs of each child and family, guided by a strengths-based, wraparound service planning process" (pg. 5). The CMHI (2011) Annual Report to Congress states that the SOC philosophy is "based on service plans that are individualized, strengths based, and evidence informed" (pg. i) and that "CMHI-funded SOCs build on the individual strengths of participating children, youth, and families to address their service needs" (pg. ii). The System of Care website, part of SAMHSA's Technical Assistance Partnership for Children and Family Mental Health (TAP), similarly states that SOCs "…design mental health services and supports that…build on the strengths of individuals, and that address each person's cultural and linguistic needs" (http://www.tapartnership.org/SOC/SOCintro.php, para.1).

Other youth mental health policy documents have also used language that reflect a focus on strengths. For example, in 2000, the Report of the Surgeon General's Conference on Children's Mental Health stated as an action step: "Increase accessible, culturally competent,

scientifically-proven services that are sensitive to youth and family strengths and needs" (Goal 4), though strengths are not defined. The Center for Mental Health Services' Strengthening *Parenting and Enhancing Child Resilience* (2007) report encouraged "a strength-based family approach to promoting mental wellness and preventing mental health problems among at-risk children and youth" (pg. 1). The term "strengths-based" is not defined, but the report describes preventive interventions focused on increasing parent, caregiver, and family capacities in order to prevent youth emotional and behavioral problems (for examples, see pgs. 18-27 of the report). Similarly, the Institute on Medicine's 2009 report stated that "Interventions designed to promote mental, emotional, and behavioral health both frequently involve directly strengthening children's competencies and positive mental health or strengthening families, schools, or communities" (pg. 110) and describes interventions (e.g. Good Behavior Game, Kellam, Brown, et al., 2008; Seattle Social Development Program, Hawkins, Kosterman, Catalano, Hill, & Abbott, 2008; Triple P Positive Parenting Program, Prinz, Sanders, Shapiro, Whitaker, & Lutzker, 2009), but the report states that "improved knowledge pertaining to the conceptualization and assessment of developmental competencies is needed to better inform interventions" (pg. 110).

There is little direction from federal policy documents on how to define youth strengths in the context of a System of Care (SOC) or how to conceptualize the inclusion of strengths. In the 2011 Children's Mental Health Initiative (CMHI) Annual Report on the national evaluation of SOCs, "strengths-based" is defined as: "the priority that services should attend to the needs and strengths of the child and individual family members" (Appendix C, pg. 2). Besides this, in the Annual Report and the *Updating the System of Care Concept* policy brief (Stroul et al., 2010) youth strengths are not defined, examples of youth strengths are not listed, and practice competencies for a focus on strengths are not articulated at a clinician or a system level. The 2011 Children's Mental Health Initiative (CMHI) Annual Report, the Community Mental Health Initiative (CMHI) digital archive (http://www.cmhi-library.org/), which archives reports, policy documents, and research articles related to the ongoing national evaluation of CMHI System of Care grantees, does not include a "Strengths" sub-heading in its archival search page (though it does include a "Resiliency" sub-heading, including fourteen articles as of November 19th, 2014, two of which include the term "strengths" in the article title). Though the term "strength" is used in all these documents, there is more attention given to how to reduce problems than to the question of exactly how SOC communities use strengths to help improve service delivery, on a clinical, administrative, system, or policy level.

Also, though federal System of Care (SOC) policy documents specifically mention youth strengths, as discussed above, theoretical reasoning for focusing on strengths in addition to needs, problems, or symptoms are not articulated in the federal policy documents. For example, the CMHI (2011) Annual Report on the National Evaluation of Systems of Care does not provide a theoretical framework for the inclusion of strengths in the evaluation. Many studies that report on client strengths, and/or use strengths variables in their analyses, do not provide a theoretical explanation for including strengths. For example, a "strengths" composite score is included in Sieracki, Leon, Miller, & Lyons' (2008) multi-level models, but theory is not discussed in the introduction or literature review of the paper

Theoretical Conceptualizations of Strengths in Systems of Care (SOC): The Under-Valued Contribution of Behavioral Theory and Related Research.

Behavioral theory and related research in applied and experimental contexts over decades provide a rationale for a focus on strengths whereby increasing strengths (e.g. positive alternative repertoires) decreases psychiatric symptoms (e.g. Goldiamond, 1974; Madden, 2013; Staats, 2012). These approaches are typically misunderstood and/or ignored (Gambrill, 2013a,b; Thyer, 2005) and are rarely described in strengths literature outside of the behavioral literature. Approaches that emphasize strengths also draw on resilience theories and frameworks, health promotion approaches, positive psychology approaches, social capital theories, empowerment theories, and the recovery and peer support literatures, among others.

Behavioral Theory and Related Research.

Empirical research related to behavioral theory shows that behavior is acquired and altered via environmental learning processes (Fisher, Groff, & Roane, 2011; Madden, 2013; Staats, 2012). Behavior is understood in the context of environmental contingencies, defined as "the relationship of behavior to its consequences and the relation of the behavior to the antecedent conditions under which this behavior occurs" (Schwartz & Goldiamond, 1975, pg. 17). Certain consequences are contingent on certain behaviors, given a particular environmental context. Behavioral interventions focus on strengths, drawing on core principles of behavioral theory and related research (e.g. see literature on Applied Behavioral Analysis, Madden, 2013). Applied Behavioral Analysis (ABA) is "a scientific approach to the study of behavior paying careful attention to the social validity of concerns addressed and related outcomes. This applies to clients as well as to their significant others, those who influence clients and who may be affected by interventions" (Gambrill, 2013a, pg. 121). Empirical studies within ABA apply behavioral principles to develop, implement, and test clinical interventions with a goal of client behavior change.

Behavioral approaches emphasize that behavior is functional; it serves a purpose for the individual (Layng, 2009). Maladaptive behavior is understood through the behavior's function - the environmental contingencies that reinforce the behavior (Layng, 2009). If behaviors are viewed as functional, even distressful or seemingly bizarre behaviors can be understood as a creative, though perhaps costly, responses to difficult life circumstances (Gambrill, 2013a, pg. 132). Decades of research provide evidence for the effectiveness of interventions based on behavioral views of behavior change (e.g. Positive Behavior Supports, Horner, Sugai, & Anderson, 2010, early and intensive intervention for children with autism spectrum disorder, Smith, 2010, and behavioral treatment for enuresis, Houts, 2006). (See also Madden, 2013.)

Behavioral Theory's "Constructional Approach" and Implications for the Relationship between Strengths and Psychiatric Symptoms.

Behavioral theory's constructional approach provides theoretical and empirical grounds for focusing on strengths. In describing the difference between a pathological approach and a constructional approach, Goldiamond (Schwartz & Goldiamond, 1975) wrote:

> One of the aims of the [constructional] interviewer is to give the client confidence that he can assume control over the solution to his own problems. What the client may interpret as failure and deficit on his part the therapist may actually interpret as signs of success and strengths…therapists…have noted how difficult it often is to get

> patients to indicate their strengths, or what is good about them. This difficulty is not necessarily a sign of malfunctioning. Rather it suggests that the patient is in tune with a culture that is heavily oriented pathologically and has applied this orientation to himself… (pg. 78-79, brackets added).

This approach does not ignore the existence of problems, it shifts the focus to identifying related environmental contingencies and using and constructing positive alternative repertoires as the means of reducing problems. A repertoire is defined as a strength, skill, asset, and environmental or relational contingency (Goldiamond, 1974, pg. 80). In the constructional approach, behaviors are changed through an increase in "strengths" (alternative available repertoires) and/or the construction of "strengths" (new positive alternative repertoires).

For example, Goldiamond (1974) used the term "strength" throughout his discussion of the constructional approach and its applications to reducing emotional and behavioral problems. Goldiamond (1974) describes that behavioral theory's constructional approach emphasizes self-determination of the client (pgs. 73-74), empowering the client to solve his/her own problems (pg. 77), the clinician's attention to what is going right in addition to what is going wrong (pg. 79-81), collaboration with the client (pg. 73), the role of context and the environment (pg. 69), the idea that all behavior typically serves a function even if it is not immediately apparent to either the clinician, the client, or family (pg. 80), and turning away from a focus on deficits and pathology (pg. 78-79). Clients' strengths (e.g. available alternative repertoires) are drawn on to reduce maladaptive behavior and symptoms. And, strengths (e.g. new positive alternative repertoires) are constructed to reduce maladaptive behavior and symptoms.

Measuring Youth Strengths in a System of Care: Use of the CANS in San Francisco's Children's System of Care

A focus on youth strengths in Systems of Care (SOC) requires reliable and valid measurement tools that measure strengths in addition to client problems. Assessments that identify the presence of current strengths and/ or identify the potential for the construction of future strengths will be termed "strengths-based assessments" for the purpose of this dissertation. Within this broad definition, a variety of assessment measures are included, such as Goldiamond's Constructional Questionnaire (Schwartz & Goldiamond, 1975, pgs. 79-81), the Family Assessment Form (FAF; McCroskey & Keezan, 1997); the Behavioral and Emotional Rating Scale (BERS) (Buckley, Ryser, Reid, & Epstein, 2006), the Child and Adolescent Needs and Strengths Assessment (CANS, Lyons 2009), and the Devereaux Student Strengths Assessment (DESSA; Lebuffe, Shapiro, & Naglieri, 2009). Before turning to the Child and Adolescent Needs and Strengths (CANS) assessment (used to collect this dissertation's data), the Behavioral and Emotional Rating Scale, Second Edition, Parent and Youth Scales (BERS-2-Y; BERS-2-P; Buckley & Epstein, 2004), will be briefly discussed below because it is used in the National Evaluation of Systems of Care (e.g. CMHI, 2011).

The Behavioral and Emotional Rating Scale (BERS).

The Behavioral and Emotional Rating Scale, Second Edition, Parent and Youth Scales (BERS-2-Y; BERS-2-P; Buckley & Epstein, 2004) is currently used to measure strengths in the national evaluation of Systems of Care communities conducted by the Children's Mental Health Initiative (CMHI), a federal program administered by SAMHSA and the Department of Health and Human Services. Both the BERS-2 Parent and Youth scales are designed to measure five

domains of social-emotional, individual-level strengths, including interpersonal strengths, family involvement, intrapersonal strengths, school functioning, and affective strengths. Each domain's score can be reported individually or compiled into a BERS Strengths Index Score, a composite score of the five domains (Buckley & Epstein, 2004). The five subscales can range from 0-20 and have a mean standard score of 10 (standard deviation: 3). The Strength Index may range from 0-200 and has a mean of 100 (standard deviation: 15), including below 70 (very poor strengths); 70 to 79 (poor strengths); 80 to 89 (below average strengths); 90 to 110 (average strengths); 111 to 120 (above-average strengths); 121 to 130 (superior strengths); above 130 (very superior strengths). All scales are written at a fifth-grade reading level and take approximately ten minutes to complete. The BERS was normed on a nationally representative sample and has been claimed to demonstrate acceptable reliability and validity (e.g. Buckley & Epstein, 2004). For example, the Strengths Index demonstrated internal consistency with coefficients ranging from .95-.97, and the BERS-2 scales correlated negatively with a self-report scale measuring youth emotional and behavioral problems (discriminant validity) (Buckley & Epstein, 2004, pg. 24).

The Child and Adolescent Needs and Strengths (CANS) Assessment

The Children and Adolescent Needs and Strengths (CANS) is a free tool, copyrighted by a non-profit, the Praed Foundation (Praed Foundation, 2014, para. 1). The CANS was designed, by Dr. Jon Lyons, for use in public systems with youth with emotional and behavioral problems (Lyons, 2009). According to Jon Lyons' biography on the Chapin Hall policy center website, versions of the CANS are currently in use in fifty states, there are currently statewide applications in 36 states, and versions of the CANS are used on every continent except Antarctica (http://www.chapinhall.org/experts/john-lyons).

Every community has the opportunity to develop their own version of a CANS tool by constructing a tailored CANS tool from a bank of hundreds of CANS items. For example, a community that served youth sex trafficking victims might choose to include an extended set of questions on potentially traumatic experiences and legal issues that are specific to this population, while another community might choose to leave those out in order to shorten the assessment. Every clinician who administers the CANS, regardless of location or community, is expected to complete a CANS training and demonstrate an inter-rater reliability of at least .70 on CANS case vignettes (compared to expert ratings) and to be re-tested annually (the percentage who are actually re-tested annually is unknown) (Praed Foundation, 2014).

The CANS was designed as a "total outcome management tool," (Lyons, 2004). While many measurement tools are designed specifically for research purposes (e.g. the BERS-2, described above), the goal of a total outcome management tool "is not simply to measure, but to actively manage outcomes by providing 'clinically actionable' information that can be used to alter the course of treatment for an individual patient" (Toche-Manley, Nankin, & Dietzen, 2004, pg. 52). The CANS was intended to be primarily useful to public system clinicians, staff, and clients, with a secondary benefit of providing data for mental health services researchers. The CANS was intended to align more easily with medical charts and billing databases in public systems than research-focused tools, but still capture a wider breadth of data than that typically found in medical charts and billing databases (Lyons, 2009).

The CANs was developed based on Communimetrics Theory (Lyons, 2009), or "measurement as communication" (pg. 30). Communimetrics Theory emphasizes the importance

of measurement tools in helping to improve communication across levels of care in public systems, according to six core principles:

1. Each item has implications for differential action
2. Levels of each item are immediately translatable into action
3. Measurement must remove the context, including:
    a. Services already in place
    b. Culture
    c. Development
4. Measurement is descriptive and minimizes cause–effect assumptions.
5. Observation windows can be trumped by the action levels.
6. Information integration (Lyons, 2009, pg. 30).

Communimetrics Theory emphasizes construct validity at the item level and de-emphasizes construct validity at the domain level, and thus the CANS includes items that might share variance, but have different clinical implications, and thus are all useful in clinical decision making. Lyons (2009, pg. 106) gives the example of the substance abuse item, which, in a sample of 6,010 youth, had a zero or even low negative correlation with some of the other items in the Emotional and Behavioral Needs domain (e.g. a correlation of -.03 with the psychosis item; a correlation of .01 with the anxiety item). Despite the low correlation with other items in the domain, Communimetrics Theory would suggest that the substance use item be left in because it has different clinical and intervention implications than the other items (Lyons, 2009, pg. 106). Though this opens the CANS to critiques of being overly lengthy, proponents would argue that the four point scoring strategy reduces the amount of time it takes to score each item and that the potential benefit of quicker and more accurate clinical decision making is worth the extra length. For example, in a sample of 544 youth in a child welfare system, use of the CANS decision making algorithm was shown to be better predictor of need for residential treatment than a multi-disciplinary team decision-making process (Chor, McClelland, Weiner, Jordan, & Lyons, 2012).

The CANS has been shown to perform reasonably well on some traditional psychometric standards. Versions of the CANS have been shown to be reliable at the item level, for example, one study found an intra-class correlation of .85 between researchers and clinicians in sixty randomly drawn case files. (Anderson et al., 2003). Acceptable item-total correlations were demonstrated in a study with a sample of 6010 youth in a public system, for the strengths domain (Cronbach's alpha = .70) and the psychiatric symptom severity domain (Cronbach's alpha = .71) (Lyons, 2009). Also, versions of the CANS have demonstrated convergent validity. For example, CANS scores were shown to be correlated (r=.63) with another self-report measure of children's emotional and behavioral problems, the Child and Adolescent Functional Assessment Scale (e.g. Lyons et al., 2004). Versions of the CANS have demonstrated discriminant validity. For example, a version of the CANS accurately distinguished youth (n=150) who were placed in inpatient care and youth who were placed in community-based outpatient care youth (Anderson et al., 2001).

Limitations of the CANS. The CANS has limitations, which are briefly mentioned here and discussed in detail below in the context San Francisco's version of the CANS. The CANS was designed as a clinician self-report tool, and there is no youth report or parent report scale to act as a check on the clinician's report. There is no consistent way to track number of contacts

and/or number of hours the clinicians spends on contacting stakeholders before filling out the CANS. Observational methods of data collection (such as videotaping assessments) have not been used, either in validity studies of the CANS or in the collection of the data used in this dissertation.

Also, though the CANS may be considered a strengths-based assessment, most of the items measure problems, symptoms, and deficits. For example, the version of the CANS used in San Francisco's Children's System of Care included one Strengths domain, with nine strengths items, out of fifteen total domains (San Francisco CANS User Manual, 2010). (In the scoring guidelines described in the San Francisco CANS User Manual (2010), the fourteen problem domains, including the psychiatric symptoms domain, are designed to assess severity of problems instead of counting frequency of problems.) Another issue is that some strengths that have been suggested as related to improved functioning (e.g. self-efficacy, quality of neighborhood or community resources) are not included. Given that the CANS has not been normed on a national sample and considering that each community might be using a different version of the CANS, it is challenging to make comparisons between communities that use the CANS.

Dissertation Setting: San Francisco's Children's System of Care (SF-CSOC) and the San Francisco CANS (SF-CANS).

The setting for this dissertation is San Francisco's Children's System of Care (SF-CSOC). SF-CSOC and the version of the CANS used in SF-CSOC are described below.

San Francisco's Children's System of Care (SF-CSOC). The setting for this dissertation is San Francisco's Children's System of Care (SF-CSOC). Every System of Care (SOC) may be different depending on the context of the SOC and population being served, and every Child and Adolescent Needs and Strengths Assessment (CANS) may be different, depending on which items were selected to meet that particular community's needs. San Francisco City and County first received Community Mental Health Initiative (CMHI) funding to develop and implement a SOC in 2002. SF-CSOC was funded by CMHI for the full six year cycle. Since then, SF-CSOC programs have been funded by a variety of both private (e.g. foundation grants) and public funding sources at the local (e.g. San Francisco County General Fund), state (e.g. California's 2005 Mental Health Services Act), and federal level (e.g. Medi-Cal, California's Medicaid Program). Currently, SF-CSOC is housed within the Community Behavioral Health Services division of the San Francisco City and County Department of Public Health.

SF-CSOC programs provide mental health and substance abuse services for youth aged 5-18 (and youth involved in the foster system up to age 22), and their families, who are San Francisco residents and whom are alleged to have emotional and behavioral problems. In 2012-2013, there were nearly 50 programs designed to provide a continuum of care from most restrictive to least restrictive, including: detention, inpatient, intensive case management, day treatment programs, intensive outpatient, and outpatient programs. The interventions provided in SF-CSOC programs are currently untracked at a system level, but may range from more traditional mental health services (e.g. individual psychotherapy or group therapy) to less traditional mental health services (e.g. drumming group, youth leadership courses). Youth receiving services in an SF-CSOC program may be involved in one or more child-serving systems, including the child welfare system, the juvenile justice system, and special education in public or private school settings. Youth may receive services in SF-CSOC from clinicians with varied degrees and certifications (e.g. BA, BSW, CCCP, MFT, MSW, LCSW, PsyD, PhD) and who may or may not have lived experience of emotional and behavioral problems. Youth are

referred for an initial assessment to a SF-CSCO program from a variety of sources that are currently untracked in this public system, but which include schools, family resource centers, community centers, and juvenile justice programs.

San Francisco CANS (SF-CANS). San Francisco's Children's System of Care (SF-CSOC) implemented the San Francisco CANS (SF-CANS) as its primary assessment tool in 2008. Every community who uses the CANS can develop their own version of the tool by picking from a bank of CANS items. In 2008, SF-CSOC went through a group process, involving consultation with subject experts and family members, in order to adopt the items currently in use on the San Francisco CANS (SF-CANS). The SF-CANS includes fifteen domains with more than one hundred items that were designed to identify type and acuity of clinical and psychosocial problems and strengths (CANS 5-18 Scoring Manual, 2010). SF-CANS takes approximately two hours to fill out online, though this is not currently tracked in SF-CSOC. Each clinician is trained to use the CANS and is re-tested annually. In San Francisco, the system-wide average inter-rater reliability of clinicians who administer the CANS is .78. Each CANS domain contains multiple items, and all CANS items are scored on a 4-point likert scale. As mentioned earlier, the fourteen problem domains are designed to assess severity of problems (instead of counting frequency of problems). For example, the San Francisco CANS Scoring Manual (2010, pgs. 4-10) describes the scoring for the psychiatric symptoms domain's nine items (psychosis, impulse/hyper, depression, anxiety, oppositional, conduct, substance use, somatization, and anger control): 0 = no evidence of any needs, 1 = a dimension that requires monitoring watchful waiting, or preventive activities, 2= a dimension that requires action to ensure that this identified need is addressed/child meets criteria for a DSM diagnosis in this category, 3 = a dimension that requires immediate or intensive action.

The Youth Strengths domain is designed to track the presence of accessible youth strengths in the past thirty days that could be used as the basis of a strengths-based plan (San Francisco CANS User Manual, 2010, pg. 3). This domain consisted of six items: Family/Caregiver Relationship Strength, Peer and Non-Family/Caregiver Interpersonal, Relationships Strength, School System and Educational Plan Strength, Extracurricular Activities and Talents Strength, Spiritual/Religious Beliefs and/or Involvement Strength, and Relationship Permanence Strength. These items were scored as 3 = strengths exist that can be used as a centerpiece for a strength-based plan, 2 = strengths exist but require some strength building efforts in order for them to serve as a focus of a strength-based plan, 1 = strengths have been identified but they require significant strength building efforts before they can be effectively utilized in as a focus of a strength-based plan, 0 = efforts are needed in order to identify potential strengths for strength building efforts. To aid clinicians in reliable and accurate scoring, a detailed San Francisco CANS User Manual (2010) was provided. The scoring for each individual item was further explicated in this manual. For example, the family/caregiver relationship strength was described on page 58 of the San Francisco CANS User Manual:

> This item refers to all biological or adoptive relatives with whom the child or youth remains in contact along with other individuals in relationships with these relatives.
>
> 3 Significant family strengths. This level indicates a family with much love and mutual respect for each other. Family members are central in each other's lives. Child is fully included in family activities.

2 Moderate level of family strengths. This level indicates a loving family with generally good communication and ability to enjoy each other's company. There may be some problems between family members. Child is generally included.

1 Mild level of family strengths. Family is able to communicate and participate in each other's lives; however, family members may not be able to provide significant emotional or concrete support for each other. Child is often not included in family activities.

0 This level indicates a child with no known family strengths. Child is not included in normal family activities (Please see Appendix A for the other items in the San Francisco CANS (5-18) User Manual).

Upon first contact with a San Francisco Children's System of Care program, a case is opened for the youth. After a case is opened, the assigned clinician has thirty days to submit a CANS initial assessment online (though specific procedures may vary by agency and number of assessments successfully submitted within the time frame is not reliably tracked). During these thirty days, the clinicians are expected to identify and contact multiple stakeholders in the youth's life (e.g. parents, caretakers, other family, teachers, caseworkers, etc.) to gather as much information as possible (though number of contacts, and whether contact was made in person or by telephone, is not currently tracked in SF-CSOC). An initial assessment is estimated to take around two hours to enter into the online database, but total amounts of time vary and are untracked in this system. Re-assessments must be completed at least every six months while the youth's case is marked "open" in the San Francisco Children's System of Care (SF-CSOC) database (actual number of assessments completed and number of contacts prior to re-assessment is unknown) and when the youth terminates contact with SF-CSOC. Reasons for termination are not reliably tracked, but could include dropping out, moving out of area, or completing a course of therapeutic intervention successfully, partially successfully, or unsuccessfully. Since August 1st, 2010, CANS data have been stored in an online, HIPAA-compliant database. De-identified data were extracted for this study by a data manager and stored by the researcher according to procedures approved by the institutional review company, Ethical and Independent Review Services.

Limitation of the SF-CANS. SF-CANS data relies solely on clinician self-report, meaning that CANS data is always at risk of clinician bias and documentation error. There is no youth report or parent report to act as a check on the clinician's report. Relatedly, direct observation (such as videotaping assessments) has not been used, either in validity studies of the CANS or in the collection of the data used in this dissertation, further limiting the conclusions that can be drawn. The number of contacts within 30 days is not tracked in SF-CSOC. Though the clinicians are encouraged to talk to as many stakeholders as possible in the youth's life, there is no way to know how many stakeholders a clinician spoke with before filling out the SF-CANS online. Similarly, though clinicians are encouraged to establish a collaborative partnership while filling out the SF-CANS with the youth and family, and the SF-CANS includes a place for a caregiver signature, there is no reliable way to know if clinicians are working collaboratively or whether or not the caregiver signed, let alone whether or not the caregiver felt coerced to sign, secretly felt that the clinician was condescending or rude, or had questions that were still unanswered. The number of submitted CANS assessments that are signed versus un-signed are not reliably tracked. If there is a blank space on any of these questions, it is impossible to differentiate between cases where a caregiver was merely unreachable during the time of the assessment, or

whether there was not a caregiver in the youth's life at all. Also, the percentage of youth referred for services, where they were referred from and to, the percentage who went to services, the type, quantity, and quality of interventions offered and used (if any), and drop-out rates and reasons are not reliably tracked in SF-CSOC. Nor do we know if clinician reports correlate with youth and caregiver reports.

The SF-CANS is also limited in the strengths items it assesses. The SF-CANS only has six youth strength items since three of the items (Optimism, Psychological Well-Being, and Integration into Community Life) were cut by system leadership. There are many youth strengths that are not included in the SF-CANS, and it is unclear why the strengths domain contains certain strength items and not others. Also, the strength items are vaguely worded in the SF CANS 5-18 Scoring Manual (2010), making it even more challenging to feel confident in the meaning of the strength items, even if significant results are found in statistical analysis. For example, the spiritual religious item (pg. 55) is confusing because the manual does not clarify whether the item can be interpreted as 1) the youth is involved in a religious community or 2) the youth identifies with a spiritual or religious practice and finds strength in this, or both of these. Researchers can only hypothesize as to what the clinician intended when interpreting this item. It is unclear whether or not the domains are unidimensional or multi-dimensional, which has implications for scoring. Also, since each community might be using a different version of the CANS, it is challenging to make comparisons across communities serving similar populations and using the CANS. Since the CANS has not been normed on a nationally representative dataset, it is challenging to make comparisons between SF-CSOC and other SOCs.

Limitations of Administrative Data.

The strengths and limitations of the CANS, and of the SF-CANS specifically, are indicative of larger issues related to the strengths and limitations of administrative data. Administrative data is defined in this dissertation (based on the definition offered by the United Kingdom's Administrative Data Liaison Service, http://www.adls.ac.uk/) as: data that is collected for non-research purposes, but can be used for research (e.g. Garrett et al., 2010). Administrative data is not as commonly used by researchers as other data sources (e.g. Garrett, E., Barnes, H., & Dibben, C., 2010), however, use of administrative data is becoming more frequent, particularly in health services research (e.g. Card, D., Chetty, R., Feldstein, M., & Saez, R., 2010). The CANS-SF dataset used in this dissertation is an example of administrative data because the primary purpose of the data collection in San Francisco is to provide information for system-level and clinical-level practice with clients. The data may be useful to researchers, but the use of the CANS data by researchers (such as for this dissertation) is secondary.

Public systems often have large amounts of data that are collected for billing or communications purposes, but otherwise is untouched (e.g. a hospital Emergency Room may store information on all patients who enter the emergency room, but this information is collected for internal billing and records purposes, not research purposes). This data often has large sample sizes and may cover a large timeframe, allowing for tracking over time that might not have been possible for researchers (who are often limited by grant and university funding). Population-level inference is often possible with administrative data, as it is typically collected at a "system population" level (e.g. a state might collect data on every youth who enters its child welfare system, which would allow for inferences about the state's child welfare population as a whole). Use of administrative data is becoming more common in health services research (e.g. Card, Chetty, Feldstein, & Saez, 2010). For example, Kaiser Permanente, a large health care providers

in the United States, has a Division of Research where researchers can request de-identified data on Kaiser members that was collected as part of a Kaiser member's typical health service provision.

Administrative data has limitations. It is usually based on self-report, resulting in reporting bias and documentation error. There are often inconsistencies in collection of the data, and data may be incomplete due to system changes that a controlled research study might have avoided (for example, three items were cut from San Francisco's CANS Strengths Domain by San Francisco System of Care high level administrators in an attempt to shorten the SF-CANS, whereas if the data were being collected solely for research purposes, the items would likely have been left in). Also, since administrative data is not collected with research as the primary goal, the reliability and validity of the data (which is of great concern to researchers) may not have been prioritized by the people collecting the administrative data. For example, the San Francisco CANS Scoring Manual (2010) defines a score of 1 on the Trauma Exposure items as "indicates a dimension where a single event trauma occurred or suspicion exists of trauma experiences" (pg. 83). Researchers are not able to differentiate between whether a score of "1" indicates a "single event trauma" or a "suspicion of trauma experiences," though the clinician might have had a particular interpretation in mind.

In Summary

This dissertation introduction is intended as an orientation to historical, policy, theoretical, and measurement issues relevant to this dissertation. Topics include the framing and prevalence of emotional and behavioral problems in the United States, the System of Care concept, behavioral theory and related empirical literature, the Child and Adolescent Needs and Strengths (CANS) tool, San Francisco's Children's System of Care (SF-CSOC), the San Francisco CANS, and the benefits and limitations of administrative data. Behavioral theory and related research suggest a directional relationship between strengths and psychiatric symptom severity, whereby an increase in strengths decreases symptom severity. However, the administrative data used here do not allow determination of directionality. Three papers, using the SF-CANS dataset, follow the introduction. The dissertation conclusion discusses overall themes, limitations of this work, and implications for social workers and others.

PAPER 1

Strengths of Youth in San Francisco's System of Care Concept: Measurement Questions and Implications for Analysis

Sarah Accomazzo

School of Social Welfare
University of California, Berkeley

Introduction

The aim of this paper is to explore modeling options for the strengths domain of the CANS assessment used in San Francisco's Children's System of Care (SF-CANS) during June 2011 – August 2013. It is possible that alternative modeling strategies of youth strengths may lead to a different understanding of youth strengths on a child or population level. Because strengths have been understudied in the System of Care (SOC) literature when compared to problems and deficits, the researcher focused on the strengths domain. (It could be useful to explore modeling options for other domains, such as the psychiatric symptom severity domain, but it out of the scope of the current paper.). Figure One suggests three different modeling options for the San Francisco CANS strengths domain (composed of six items: family/caregiver relationship strength, peer and non-family/caregiver interpersonal relationships strength, school system and educational plan strength, extracurricular activities and talents strength, spiritual/religious beliefs and/or involvement strength, and relationship permanence strength) that are explored: a) Individual Items, b) Sub-Scales, and c) Overall Composite Score.

Figure One

Modeling Options for the Strengths Domain
(listed from least data reduction to most data reduction)

	Name	Description	Background
1	Individual Items	Six items treated as unique variables.	This is recommended in a CANS manual (Lyons, 2009, pg. 104). This has been utilized in various papers using the CANS (e.g. Sieracki et al., 2009).
2	Sub-scales (conceptually derived)	Two sub-scales: a) Family / Caregiver Strength factor (including two items) b) Other Strengths factor (including four items).	A factor analysis of the CANS as a whole or unique domains of the CANS has not been reported (e.g. Sieracki et al., 2009).
3	Overall Composite Score	Sum of the six individual items.	This is recommended in a CANS manual (Lyons, 2009, pg. 104). This method has been utilized in various papers for the strengths domain (e.g. Griffin et al., 2009; Whitson et al., 2012).

Three research questions are explored:
1) What is the prevalence of youth strengths, using the individual items (#1), two subscales, (#2), and the overall composite score (#3)?
2) Is Youth Strengths a unidimensional construct or is it better indicated by the two subscales of strengths? (What is the factor structure revealed in a CFA of the Strengths domain, if any?)
3) What are the demographic differences (on gender, age, and ethnicity) in strengths, if any, for the individual items (#1), subscales (#2), and the overall composite score (#3)?

Prevalence of Emotional and Behavioral Problems in the United States.

Emotional and behavioral problems in childhood are common. It is estimated that one-third to one-half of youth under in America will meet commonly accepted criteria for an emotional or behavioral problem (as defined in the Diagnostic Statistical Manual-IV-TR) at least once before the age of 21 (e.g. Costello et al., 2003; O'Connell et al., 2009; Kessler, 2012). (The Diagnostic Statistical Manual (DSM; 2014) is the predominant classification system for emotional and behavioral problems in the United States). Related research suggests that approximately 10% of youth will experience severe problems (impairments that disrupt functioning and developmental tasks in one or more life domains) during any given year (e.g. Kessler et al., 2012; Merikangas et al., 2010). It is estimated that emotional and behavioral problems in childhood and adolescence cost the public upwards of $200 billion each year (O'Connell et al., 2009, pg. 231), and that approximately half of youth who meet criteria for an emotional and behavioral problem receive inadequate or no mental health treatment at all (Kessler et al., 2012; Merikangas et al., 2010; President's New Freedom Commission, 2003).

Systems of Care (SOCs) and a Focus on Strengths

Since 1984, with the adoption of Public Law 102-321, the major federal response to youth with emotional and behavioral problems has been the funding of Children's Systems of Care (SOC) (Stroul et al., 2010). Currently, 173 U.S. communities have been awarded federal funding to serve youth with severe emotional and behavioral problems by building and operating a System of Care (SOC) in their communities (CMHI 2012, pg. ii), and a national evaluation of SOCs has been conducted annually since 2002 (e.g. CMHI, 2011). The most recent definition, used in federal policy, funding, and legislation, defines a SOC as:

> A spectrum of effective, community-based services and supports for children and youth with or at risk for mental health or other challenges and their families, that is organized into a coordinated network, builds meaningful partnerships with families and youth, and addresses their cultural and linguistic needs, in order to help them to function better at home, in school, in the community, and throughout life (pg. 6).

The overall goal of SOCs is to improve child, youth, and family outcomes (Stroul et al., 2010) by incentivizing coordination of services adhering to three core values: 1) family-driven and youth-guided, 2) community-based, and 3) culturally and linguistically competent (TAP, 2014). Embedded in these values is the notion that services for children with emotional and behavioral problems should include a focus on strengths. Specifically, the ideal of family-driven and youth-guided care must use "the strengths and needs of the child and family" to determine "the types and mix of services and supports provided" (TAP, 2014), and SOC services should always "build off of child, youth, and family strengths" (TAP, 2014). Beyond these values, literature regarding SOCs emphasizes eight core principles, including the principle that services are "based on service plans that are individualized, strengths based, and evidence informed" (TAP, 2014). The definition of a strength and specific information on how to use a strength in a System of Care are not provided.

Similarly, other youth policy documents use the term "strength" but do not clarify its meaning or give examples of how to use strengths in practice. For example, the Comprehensive Community Mental Health Services for Children and their Families Program's Annual Report to Congress (2010, pg. ii) states that services provided in federally-funded Children's Systems of

Care (SOC) should "build on the individual strengths of participating children, youth, and families to address their service needs" (CMHI, 2011, pg. ii), but a specific definition of "strength" is not offered. In 1999, the Report of the Surgeon General's Conference on Children's Mental Health stated as an action step: "Increase accessible, culturally competent, scientifically-proven services that are sensitive to youth and family strengths and needs" (Goal 4). In 2007, the Center for Mental Health Services (a division of SAMHSA) produced a report on "the effectiveness of programs that use a strength-based family approach to promoting mental wellness and preventing mental health problems among at-risk children and youth" (pg. 1). The report to the US Congress stated that "benefits can be measured in positive, strength-based "quality of life" terms as well as in reductions of negative consequences and their costs" (pg. 31). Similarly, the Institute on Medicine's 2009 report stated that "Interventions designed to prevent MEB disorders and problems and those designed to promote mental, emotional, and behavioral health both frequently involve directly strengthening children's competencies and positive mental health or strengthening families, schools, or communities" (pg. 110).

Towards a Definition of a Strength in a System of Care.

Terminological consensus has not been reached as far as the definition of a "strength," either in the policy documents discussed above or in the broader research literature. Conceptually related research has been conducted by researchers in different disciplines using different terminology, including researchers in the areas of behavioral theory and applied behavioral analysis (e.g. Goldiamond, 1974; Madden, 2013; Staats, 2012), resilience (e.g. Masten, 2007), positive youth development (e.g. Catalano et al., 2004), developmental psychopathology (e.g. Luthar & Brown, 2007), social policy (e.g. Jenson et al., 2013), youth violence (e.g. Ozer et al., 2004), and public mental health services (e.g. Graves, 2007; Lyons, 2009). Related research emphasizes the role of positive youth attributes, skills, and environments in protecting against the negative effects of risks and/or promoting health. Though conceptually related, these literatures use different terminology (e.g. strengths, alternative available repertoires, new alternative repertoires, resilience, assets, resources, competencies, protective factors, promotive factors), resulting in some conceptual confusion. (For example, see Probst (2009) for a discussion of terminological confusion in the strengths-based practice literature; see Luthar (2000) and Fergus & Zimmerman (2006)'s critique of the resilience terminology).

This paper defines a youth strength as: a characteristic, attribute, repertoire, ability, thought, skill, behavior, or resource, located at the individual, interpersonal, or environmental level, that may reinforce and/or construct desired behaviors, protect against the negative effects of risk, and/or promote health and well-being. In the context of mental health service delivery for children and adolescents with emotional and behavioral problems, strengths may be discussed as 1) characteristics for assessment that describe the youth's current repertoire of strengths 2) characteristics for intervention that may help a child or adolescent develop and thrive, or 3) service delivery outcomes. Examples of strengths include (but are not limited to): insight into oneself, resources and relationships that are available in the youth and family's social network and community, talents, cultural identity and practices, spirituality, high levels of commitment to school, opportunities for employment in the community.

Measuring Youth Strengths in a System of Care: A focus on the CANS and San Francisco's Children's System of Care

A focus on youth strengths in a System of Care (SOC) requires reliable and valid tools that measure strengths in addition to client problems. Assessments that identify current strengths and/ or potential for construction of future strengths will be termed "strengths-based assessment measures" for the purpose of this paper. Related assessment measures include Goldiamond's Constructional Questionnaire (Schwartz & Goldiamond, 1975, pgs. 79-81), the Family Assessment Form (FAF; McCroskey & Keezan, 1997); the Behavioral and Emotional Rating Scale (BERS) (Buckley et al., 2006), and the Child and Adolescent Needs and Strengths Assessment (CANS, Lyons 2009), and the Devereaux Student Strengths Assessment (DESSA; Lebuffe, Shapiro, & Naglieri, 2009).

Different strengths-based measurement tools were designed to focus on different areas, and they all have their own strengths and limitations. Strengths-based assessment measures might differ on: 1) the purpose of the assessment (e.g. clinical decision-making, pure research, versus a hybrid of these), 2) characteristics of the population being assessed (e.g. age, ethnicity, gender, socio-economic status, sexual minority status) 3) setting in which the tool is used (mental health clinics, schools, juvenile justice settings, Systems of Care), and 4) theoretical underpinnings. Other factors include length of the tool, psychometric properties, cultural sensitivity, language translations, options for types and numbers of raters (e.g. clinician, parents, youth, teacher, family member) and/or options for multiple raters (e.g. does the scale offer a parent and a teacher version in order to triangulate responses).

The CANS. The Children and Adolescent Needs and Strengths (CANS) is one of the many tools that qualifies as a strengths-based assessment according to the definition above as it was designed to measure the presence of both strengths and problems. The CANS is a free tool, copyrighted by a non-profit, the Praed Foundation (Praed Foundation, 2014, para. 1) and designed, by Dr. Jon Lyons, for use in public systems with youth with severe emotional and behavioral issues (Lyons, 2009). According to the Jon Lyons' biography on the Chapin Hall policy center website, versions of the CANS are currently in use in fifty states, there are currently statewide applications in 36 states, and versions of the CANS are used on every continent except Antarctica (http://www.chapinhall.org/experts/john-lyons). Every community has the opportunity to develop their own version of CANS by constructing a tailored CANS from a bank of hundreds of items. Every clinician who administers the CANS, regardless of location or community, is expected to complete a CANS training and demonstrate an inter-rater reliability of at least .70 on CANS case vignettes (compared to expert ratings) and is reported to be re-tested annually (the percentage who are actually re-tested annually is unknown) (Praed Foundation, 2014).

The CANS was designed as a "total outcome management tool," (Lyons, 2004), with a basis in Communimetrics Theory (Lyons, 2009). While many measurement tools are designed specifically for research purposes (e.g. the BERS-2, described above), the goal of a total outcome management tool "is not simply to measure, but to actively manage outcomes by providing 'clinically actionable' information that can be used to alter the course of treatment for an individual patient" (Toche-Manley et al., pg. 52). The CANS was intended to be primarily useful to public system clinicians, staff, and clients, with a secondary benefit of providing data for mental health services researchers.

The CANS was intended to improve service delivery in public systems by increasing communication among system stakeholders, aligning more easily with medical charts and billing

databases in public systems than research-focused tools, and capturing a wider breadth of data than that typically found in medical charts and billing databases (Lyons, 2009). For example, a study examined the rate of improvement for youth placed in a residential treatment center (RTC) (Chor et al., 2012). Youth in the sample (n=544) had received an initial CANS assessment, after which a multi-disciplinary team met to decide on level of care (e.g. assignment to an RTC or to other levels of care) based on the information reported in the CANS assessment. The research team went back to the case files and used a CANS decision making algorithm to assess whether or not the CANS algorithm agreed with the multi-disciplinary team decision. Youth who were placed in the RTC by the multidisciplinary team, but for whom the CANS algorithm indicated a different level of care, showed slower rate of improvement in symptoms over a year (mean difference between the groups: .65 points/year, $p<.01$) than youth for whom the CANS algorithm was consistent with the multidisciplinary team's decision (Chor et al., 2012).

While versions of the CANS have performed adequately on traditional psychometric standards (see Methods section), the CANS domains have not been explored for construct validity to the knowledge of this researcher. It is unclear whether or not the strengths section of the CANS is unidimensional or multi-dimensional. Understanding the dimensionality of the strengths domain has utility for future analyses. For example, if the strengths section is unidimensional, this would imply that using a composite score for the strengths section in future analyses might be useful. If the strengths section is shown to be multi-dimensional, this would imply that using multi-dimensional scoring methods, either the individual items themselves or factor scores, might be a useful way to model the strengths items in future analyses. Also, modeling the strengths section as unidimensional or multidimensional has implications for reporting prevalence of strengths by gender, age, and ethnicity, and for looking at group differences on strengths between gender, age, and ethnicity. A decision to model the strengths items in different ways could result in different interpretations of prevalence or group demographic differences. It is useful to explore this issue empirically before continuing with future analyses.

Method

Study Design and Measures.

The SF-CANS. In early 2008, San Francisco's Children's System of Care adopted the Child and Adolescent Needs and Strengths (SF-CANS) instrument as its primary assessment tool for all youth programs. Consistent with the tool's basis in Communimetrics theory (Lyons, 2009), which emphasizes item-level reliability and validity, different stakeholder groups have used different configurations of strength items in their own settings, and have added to the bank of strengths items used as part of the CANS. San Francisco's Children's System of Care went through a group process, involving consultation with subject experts and family members, in order to adopt the items for the SF-CANS.

The SF-CANS includes fifteen domains with more than 100 items. Fourteen of the domains were designed to identify type and acuity of clinical and psychosocial problems (e.g. psychiatric symptom severity, risk behaviors) and one domain was designed to identify youth strengths (SF CANS 5-18 Scoring Manual, 2010). For each item, assessors assign a rating using a four-point likert scale. The problem domains are designed to assess severity of problems (instead of counting frequency of problems), while the Strengths domain is designed to track the presence of strengths that could be used in a strengths-based plan. After a case is opened,

clinicians have thirty days to submit a CANS initial assessment online (the number of assessments submitted within this timeframe compared to number submitted late is not reliably tracked). During these thirty days, the clinicians are expected to identify and contact multiple stakeholders in the youth's life (e.g. parents, caretakers, other family, teachers, caseworkers, etc.) to gather as much information as possible.

The number of contacts within thirty days is not tracked in SF-CSOC, but could vary from one contact to ten or more (as estimated by a senior SF-CSOC administrator). Though clinicians are encouraged to talk to as many stakeholders as possible in the youth's life, there is no way to know how many stakeholders a clinician spoke with before filling out the SF-CANS online and whether the contacts were in person or by telephone. Though clinicians are encouraged to establish a collaborative partnership while filling out the SF-CANS with the youth and family, and the SF-CANS includes a place for a caregiver signature, there is no reliable way to know if clinicians are working collaboratively or whether or not the caregiver signed, let alone whether or not the caregiver felt coerced to sign, secretly felt that the clinician was condescending or rude, or had questions that were still unanswered, even if they signed. The number of submitted CANS assessments that are signed versus un-signed are not reliably tracked. If there is a blank space on any of these questions, it is impossible to differentiate between cases where a caregiver was merely unreachable during the time of the assessment, or whether there was not a caregiver in the youth's life at all.

An initial assessment is estimated to take around two hours to enter into the online database, but total amounts of time vary and are untracked in this system. Re-assessments must be completed at least every six months while the youth's case is marked "open" in the San Francisco Children's System of Care (SF-CSOC) database (actual number of assessments completed and number of contacts prior to re-assessment is unknown) and when the youth terminates contact with SF-CSOC. Reasons for termination are not reported, but could include dropping out, moving out of area, or completing a course of therapeutic intervention successfully, partially successfully, or unsuccessfully. The percentage of youth referred for services, where they were referred from, the percentage who went to services, and the type, quantity, and quality of interventions (if any) are not reliably tracked in SF-CSOC. Since August 1st, 2010, CANS data have been stored in an online, HIPAA-compliant database. De-identified data were extracted for this study by a data manager and stored by the researcher according to procedures approved by the institutional review company, Ethical and Independent Review Services.

Reliability and Validity of the CANS. All clinicians who use a version of the CANS are expected to go through a day of standard training (Lyons, 2009). At the end of the training, each clinician is required to take an exam where they rate a case vignette and must achieve an intra-class correlation (compared to expert ratings) above .70. All clinicians should be re-tested yearly (though this is not tracked nationally or internationally to the knowledge of the researcher) (Lyons, 2009). Versions of the CANS have been shown to be reliable at the item level. For example, one study reported an intra-class correlation (comparing researchers and clinicians) of .85 in sixty randomly drawn case files. (Anderson et al., 2003). Acceptable item-total correlations was demonstrated in a study with a sample of 6010 youth in a public system, for the strengths domain (Cronbach's alpha = .70) and the psychiatric symptom severity domain (Cronbach's alpha = .71) (Lyons, 2009). In San Francisco, the system-wide average inter-rater reliability of clinicians who administered the CANS during the timeframe of this study was .78.

CANS scores were shown to be correlated (r=.63) with another self-report measure of children's emotional and behavioral problems, the Child and Adolescent Functional Assessment Scale (e.g. Lyons et al., 2004). A version of the CANS accurately distinguished youth (n=150) who were placed in inpatient care and youth who were placed in community-based outpatient care youth (Anderson et al., 2001).

The Strengths Domain of the SF-CANS. This paper focuses on gaining a deeper understanding of the Strengths Domain. The Strengths Domain in the SF-CANS contains six items: family/caregiver relationship strength, peer and non-family/caregiver interpersonal relationships strength, school system and educational plan strength, extracurricular activities and talents strength, spiritual/religious beliefs and/or involvement strength, and relationship permanence strength. A detailed description of these items are given in the CANS 5-18 Scoring Manual (2010, pgs. 51-56) and have been included in Appendix A.

The CANS used in San Francisco's public system originally included three other youth strength items (optimism, psychological well-being, integration into community life). These items were cut (along with several other items from various domains) by high level administrators in early 2010 in response to clinicians in the system advocating for a shorter assessment. (Documentation as to why particular items were cut does not exist to the knowledge of this researcher.) From a research perspective, and for the purpose of this study, the cutting of these items is unfortunate because the optimism and psychological well-being items were the only two internal strengths items in the CANS. The remaining six items all track external strengths.

The SF-CANS Youth Strengths domain is designed to track the presence of accessible youth strengths in the past 30 days that could be used as the basis of a strengths-based plan. The items in this domain were originally developed from the Child and Adolescent Strengths Assessment (CASA; Lyons, Uziel-Miller, Reyes & Sokol, 2000), an early iteration of the CANS. The CASA has 30 items which were designed to measure strengths in six life domains: Family, School / Vocation, Psychology, Peers, Morality / Spirituality, Extracurricular Involvement. These 30 strengths items were later consolidated and updated to be consistent with the domains of functioning outlined in Stroul and Friedman's (1994) second monograph on Systems of Care philosophy and implementation (Stroul & Friedman, 1994). The revised set of strengths items were part of the original CANS –Mental Health tool (2003) and included nine items (Family, Interpersonal, Relationship Permanence, Educational, Vocational, Well-being, Spiritual / Religious, Talents / Interests, and Inclusion).

Sample.

The sampling frame included youth (aged 5-22 years) who received an initial CANS assessment between June 2011 and August 2013 and whose assessing clinician entered the initial assessment into an online database within the timeline required by the city, within 45 days of first of point contact (n=2,402). (Youth who entered into the system but whose clinician did not meet the 0-45 day deadline were excluded from the sample due to worry that the assessment was no longer capturing the youth's presentation upon entering the system). The final sample includes all youth who did not have any missing data on any of the variables included in the final analyses (n=2049). There are no statistically significant differences in the demographics between the sample used for this paper and the group who were dropped from the sampling frame due to missing data. Also, since there are no known reasons for why demographic data may be missing, and no patterns of missing data were spotted by the researcher, the missing data is assumed to be

Missing Completely at Random (MCAR; missing due to variable that are unrelated to the analyses), and thus not a concern for this analysis.

Youth were referred for an initial assessment from a variety of sources that are currently untracked in this public system, but which include schools, family resource centers, community centers, and juvenile justice programs. They were typically referred due to alleged emotional, behavioral, and environmental issues that were said to impair their functioning in their family, school, or community. Socio-economic status of children and families is not currently tracked in this SOC because 98% of client services are paid for by billing Medicaid, and so system leaders made a decision to not track socio-economic variables.

Demographics are presented in Table 1. This sample is about two-thirds male and ethnically diverse, as more than 90% of the youth in the sample identify as non-white. The mean age of youth in the current sample is 12.53 years (SD: 3.85, range: 5-22). Five to eighteen year olds represent about 99% of the sample, with 19-22 years olds representing 1% (this is likely because 19-22 year olds who are involved in the foster care system and choose to not to emancipate at age 18 will sometimes qualify to continue receiving services in San Francisco's System of Care).

Table 1

Demographics[a] of Youth Upon Entry to San Francisco's Children's System of Care (n=2049).

Characteristic	n	% [a]
Age		
5-8	426	21
9-12	528	26
13-15	501	24
16-18	569	28
19-22	25	1
Ethnicity		
African-American/Black	685	33
Asian	446	22
Latino	571	28
Multi-Ethnic	72	4
Pacific Islander/Native Hawaiian	44	2
White	163	8
Other	68	3
Gender		
Male	1258	61
Female	791	39

[a]Percentages may not add to 100 due to rounding.

Variables.

Strength Variables.

The Youth Strengths domain on the CANS assessment is designed to track the presence of accessible youth strengths in the past thirty days that could be used as the basis of a strengths-based plan. Each of the six items is rated on a 0-3 likert scale by clinicians. The Strengths items are scored as: 0 = efforts are needed in order to identify potential strengths for strength building efforts. 1 = strengths have been identified but they require significant strength building efforts before they can be effectively utilized as a focus of a strength-based plan 2 = strengths exist but require some strength building efforts in order for them to serve as a focus of a strength-based plan, and 3 = strengths exist that can be used as a centerpiece for a strength-based plan. This paper models youth strengths in three different ways (also see Figure 1):

Option 1 (Strengths domain modeled as individual items): The six strength items are listed below. Please see Appendix A for detailed descriptions of each item, as provided for clinicians in the CANS 5-18 Scoring Manual (2010, pgs. 51-56).

Famnew1. The Family/Caregiver Relationship Strength item is a continuous variable scored 0-3.
Interpnew1. The Peer and Non-Family/Caregiver Interpersonal Relationships Strength is a continuous variables scored 0-3.
Educnew1. The School System and Educational Plan Strength item is a continuous variable scored 0-3.
Talnew1. The Extracurricular Activities and Talents Strength item is a continuous variable scored 0-3.
Spirnew1. The Spiritual/Religious Beliefs and/or Involvement Strength item is a continuous variable scored 0-3.
Relnew1. The Relationship Permanence Strength item is a continuous variable scored 0-3.

Option 2 (Strengths domain modeled as sub-scales based on factor analysis):
FamilyRelpFactor is a continuous variable ranging from 0-6. It is a composite score created by adding up the family /caregiver relationship strength item and the relationship permanence item scores. This variable was used to report prevalence statistics (Table 2).

OtherFactor is a continuous variable ranging from 0-12. It is a composite score created by adding up the peer/non-caregiver item, the education item, the activities and talents item, and the spiritual/religious item. This variable was used to report prevalence statistics (Table 2).

FactorscoreStrength1 is a factor score (created using the regression method) for the FamilyRelpFactor. This variable was used when running t-tests and ANOVAs (Tables 5.1 and 6.1).

FactorScoreStrength2 is a factor score (created using the regression method) for the OtherFactor. This variable was used when running t-tests and ANOVAs (Tables 5.1 and 6.1).

Option 3 (Strengths domain modeled as an overall composite score):
Strength is a continuous variable ranging from 0-18. The six strength items (scored 0-3) were added up for each youth to create this composite score.

Demographic Variables:

Male is a categorical variable for gender, coded male = 1 and female = 0.

Ethnicity is a categorical variable for youth ethnicity, with 0 = white, 1 = African-American/Black, 2 = Asian, 3 = Latino/a, 4 = Multi-Ethnic, 5 = Pacific Islander/Native Hawaiian, 6 = Other.

Agedum is a categorical variable for youth age, coded 0 = 5-8 years, 1 = 9-12 years, 2 = 13-15 years, 3 = 16-18 years, and 4 =19-22 years.

Statistical Analysis.

The software package Stata13 was used for all analyses in this paper (see Appendix D for a list of Stata13 coding for the analyses in this paper). The analyses in this dissertation all use the strengths items in their ordinal variable form (scored 0-3). Another option would have been to recode each item into binary dummy variables, with a 2 or 3 recoded into a 1 and a 1 or 0 recoded into a 0. Scored in this way, a "1" suggests a strength that is currently available and could be clinically useful in a strengths-based treatment plan, while a 0 suggests a strength that could not currently be used in a strengths-based treatment plan. CANS creator Jon Lyons (Lyons, 2009, pg. 104) notes that this dichotomous scoring might be preferable in clinical settings where clinicians will be writing goals and treatment plans based on a CANS assessment and a straightforward, clinically interpretable scoring strategy is useful. However, this method does miss nuance in item level changes from 0 to 1, or 1 to 2, or 2 two three, and vice versa. From a researcher / statistical standpoint, less item reduction and more variation might be preferable, resulting in more overall information. Given this, for the purpose of this dissertation, the decision was made to leave the items in their continuous form. Regardless, every effort is made to specify clinical implications of any findings.

For research question one, frequencies, means, standard deviations, skew, kurtosis, and ranges were calculated for the variables presented in the analysis. The means reported for the subscales (modeling option #2) for this research question were based on the *FamRelp Factor* and *OtherFactor* variables (see the Variables section above), not the factor scores.

For research question two, a confirmatory factor analysis, positing two theoretically derived factors, was conducted. First, the Strength domain was modeled with two correlated factors (since it is theoretically reasonable that they might share some variance). To determine whether the two factor model adequately fit the data, a two-index strategy was used (Hu & Bentler, 1999). Specifically, the Comparative Fit Index (CFI) and Root Mean Square Error of Approximation (RMSEA) were examined in these analyses. There are many guidelines for assessing model fit (Harrington, 2009, pg. 53). According to Harrington (2009), Kline (2005)

posits that a CFI of greater than around .90 suggests an acceptable fit, and a CFI with a value greater than .95 indicates an excellent fit and a value from .90-.94 suggests a close fit; the RMSEA suggests an unacceptable fit when greater than .10, a reasonable fit when less than .08, and a close fit when less than .05 (pg. 53).

To determine whether a multi-dimensional factor structure is a better fit for the data than a unidimensional factor structure, the two factor model was compared to a single factor model. First, the chi-square goodness-of-fit tests, CFI, and RMSEA fit indices were compared between the two models, and then a likelihood ratio test was run comparing the two models. The likelihood ratio test compared two nested models at the 0.05 level. The null hypothesis for this test is that the restricted model (one factor model) is a better fit for the data than the full (two factor) model. A significant p-value suggests that there is evidence to reject the null hypothesis, and the full model (two factor model) is a better fit for the data. (The two factors used in this analysis were derived using a Principle Components Analysis (PCA) with a varimax rotation, and then a confirmatory approach used to test the two factor model. Next steps would include a detailed, theoretical exploration of the two factors and then a separate study confirming the factor structure with new validation data.)

Research question three asked whether there were gender, age, or ethnic differences in strengths in the sample, and whether measuring strengths in any of the three ways leads to alternative interpretations of differences among the groups. To address this question, three sets of analyses were completed. For all the research question three analyses, when attending to modeling option #2 (sub-scales), factor scores, using the regression method, were created and used.

To test gender differences, independent sample t tests were run, using the three strengths modeling options, to determine whether there were any significant differences between males and females on youth strengths upon entry to this Children's System of Care. Then, an effect size (Cohen's d) was calculated for each strength scoring option. Effect sizes are useful in measuring the magnitude of an effect, and thus providing a metric for deciding whether or not a significant difference is meaningful. The Cohen's d coefficient represents a standardized difference between two group means. Cohen (1988) suggested thresholds for determining whether the magnitude of an effect is small (about .2 or lower), medium (about .5), or large (about .8 or higher).

To test differences between age groups, one way ANOVA tests (F test, p-values) were conducted to learn whether there was a significant difference in strength variable means between at least two group pairs. An effect size (eta-squared, confidence interval) was calculated to suggest how much of the total variation in the strengths variable is explained by the age variable. Eta-squared is typically scaled with .01 as a small effect; .06 as a medium effect; and .14 as a large effect (http://www.ats.ucla.edu/stat/stata/faq/effectsize.htm). Then, if the one-way ANOVA provided evidence for significant differences between at least two age groups, post-hoc pairwise Bonferroni comparison tests (mean differences and p-values) were conducted to determine which groups had a significant mean difference. (Though there are several post-hoc analysis tests available after one-way ANOVAs, the Bonferroni tests are generally considered somewhat conservative, and so were chosen for this study.)

To test differences between ethnicity groups, one way ANOVA tests (F tests, p-values) were conducted to learn whether there was a significant difference in strength variable means between at least two group pairs. Eta squared effect sizes were calculated for these F tests as well.

Results

The statistical test assumptions for t-tests and ANOVAs were checked to make sure there were no major violations. Since this study has a large sample (n=2049), for the t-test and ANOVAs, the assumption of normality is not "necessary for valid estimation of the model parameters" (Rabe-Hesketh & Skrondal, 2012, pg. 17). However, as an additional check, boxplots for each of the three strengths modeling options were inspected (see Appendix B for the box plots for each of the three strengths scaling options). These boxplots suggest that, although the distributions are not perfectly normal, there are no severe violations.

For the ANOVA tests performed for research question three, the assumption of constant variance between groups was tested. According to the ratio for standard deviations check (Johnson & Whichern, 2002, pg. 290), the largest group standard deviation was compared to the smallest group standard deviation, and a ratio under 2 is considered to be a non-severe violation of constant variance. The age group means, standard deviations, and ratio tests are reported in Table 4.1. All the ratios are under 2, and so the assumption of constant variance is approximately met. The ethnicity group means, standard deviations, and ratio tests are reported in Table 5.1. All the ratios are under 2, and so the assumption of constant variance is approximately met. Research question one explored the prevalence of youth strengths, using the three alternative strengths modeling strategies, upon entry to San Francisco's Children's System of Care (SF-SOC). Prevalence rates are reported in Table 2.

Table 2

Prevalence of Youth Strengths Upon Entry to the Children's System of Care Modeled Three Different Ways

	Mean (SD)	Skew	Kurtosis	Range
Item Modeling Option	Continuous			
Individual Items				
Family/Caregiver Relationship Strength	1.91 (.80)	-0.12	2.12	0-3
Peer and Non-Family/Caregiver Interpersonal Relationships Strength	1.88 (.85)	-0.26	2.29	0-3
School System and Educational Plan Strength	1.77 (.90)	-0.24	2.25	0-3
Extracurricular Activities and Talents Strength	1.93 (.92)	-0.36	2.13	0-3
Spiritual/Religious Beliefs and/or Involvement Strength	1.79 (1.23)	-0.44	1.58	0-3
Relationship Permanence Strength	2.01 (.91)	-0.3	1.83	0-3
Sub-Scales				
Family and Relationship Strengths	3.92(1.41)	-.23	2.29	0-6
Other Strengths	7.37(2.70)	-.30	2.61	0-12
Overall Composite Score	11.29 (3.32)	-0.22	2.7	0-18

[a]Percentages may not add to 100 due to rounding.

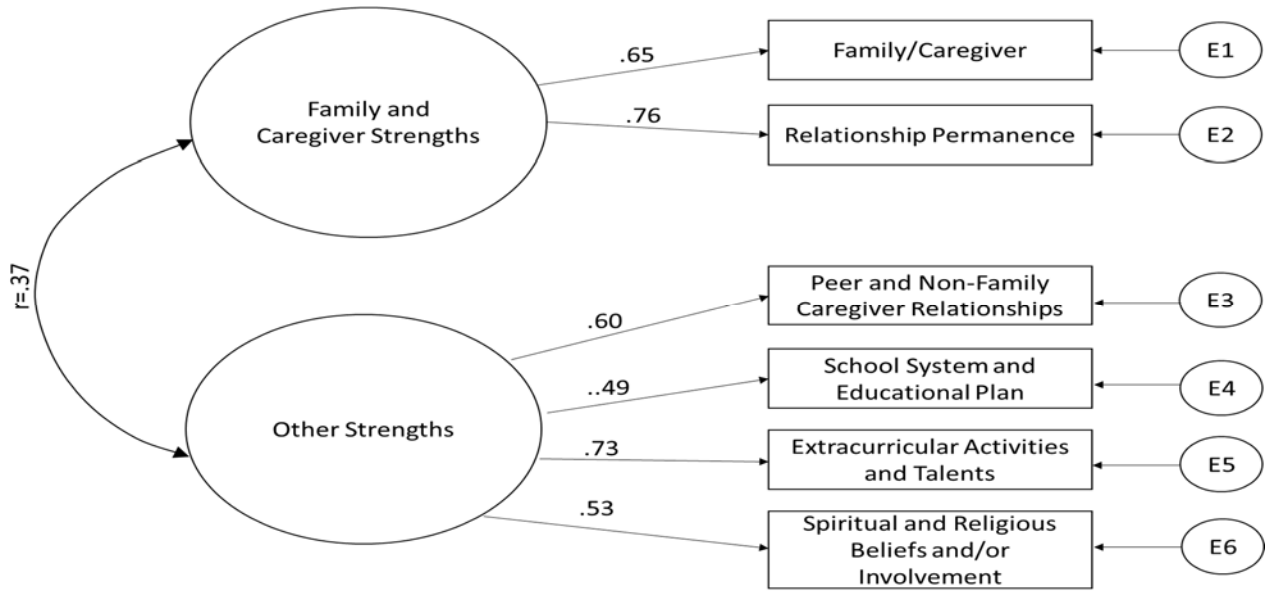
The standard deviations for all the modeling options are fairly large, suggesting variation in the data. For the individual strength items (modeling option 1), the clinician-reported mean strength scores are all close to 2. (A score of two on an individual item indicates that the clinician reported that the youth entered SF-CSOC with that particular strength available for

immediate use as part of a strengths-based treatment plan.) Thus, the individual items indicate, on average, that the youth in the sample appear to enter the SOC with many available strengths.

The sub-scale means and the overall composite score means are harder to interpret. For example, a mean score of 3.92 out of 6 points on the Family and Relationship Strength suggest that youth enter the SOC, on average, with family relationship strengths that can be used in a treatment plan. The overall composite score mean is the most challenging of the three modeling options to interpret as far as prevalence is concerned. To receive an average score of 11.92, a youth would have had to receive scores of 2 or 3 on at least 3 or 4 of the individual items, suggesting that the average youth has reasonable amounts of strengths. However, in a composite score, it is impossible to know which items made larger contribution (and thus are accessible to youth upon entry) and which items made little or no contribution (and thus could be the focus of strength-building). It is important to note, also, that the standard deviations for all the modeling options are fairly large, suggesting that there is considerable variation in the data. Regardless of how strengths are modeled, on average, strengths appear to be reasonably prevalent in San Francisco's Children's System of Care. However, the individual items appear to give the most information, particularly when clinical utility is considered.

For research question two, the path diagram for the two factor model is reported in Figure 2.

Figure 2.



The goodness-of-fit indices for both the one factor and two factor models, plus the likelihood ratio test comparing the two models, are presented in Table 3. First, the fit indices provide evidence that the two factor model adequately fits the data. The likelihood ratio test ($chi2=209.24$, $p<.001$) comparing the one factor model and the two factor model suggested that

Table 3							
Test of Unidimensionality of Youth Strengths							
	x^2	df	p	CFI	RMSEA	RMSEA lower bound	RMSEA upper bound
2-Factor Model	1494.61	15	<.001	0.90	0.10	0.08	0.11
1-Factor Model	1494.61	15	<.001	0.76	0.14	0.13	0.15
Likelihood Ratio Test: 2 Factor vs 1 Factor	209.24	1	<.001				

the two factor model fits the data better. For the two factor model (using the specifications described above), a CFI of .90 suggests a close fit, but the RMSEA of .1 is on the cut-off of an adequate fit. (The correlation table for all the strengths scoring options is reported in Appendix C.) The two factor model reported an estimated correlation between the two factors of .37, suggesting that the factors share only about 14% of the overall variation. Considered altogether, the fit indices appear to indicate that the two factor model is a better fit than the one factor model, but overall, the two factor model fit is only adequate (not excellent).

The interpretation of the two factors is also somewhat confusing. The Family and Caregivers Factor contains two items (Family Relationships and Relationship Permanence). It is possible that these two items loaded on the same factor (respectively r=.76 and r=.65) because the clinicians only recorded the presence of relationship permanence in family relationships. Or, this factor could suggest that families are the main relationship in youth's lives where relationship permanence exists. The Other Strengths factor (including the peer relationship, school system and education plan, extracurricular talents and interests, and the spiritual/religious strength items) may be a "catch all" factor merely reflecting the higher loadings of the Family and Caregivers Factor, particularly considering the low loading of the education item (r=.49) and the spiritual religious item (r=.53).

For research question three, the male and female means and standard deviations, t statistics (and p-values), and effect sizes (and confidence intervals) were reported for modeling options 1, 2, and 3 in Table 4. Overall, there do not appear to be meaningful differences between females and males on youth strengths upon entry to SF-CSOC, and strength modeling strategies do not appear to make a difference. This data show some significant differences between females and males, but the Cohen's d effect sizes for each of the significant differences suggest that the magnitude of the effect is negligible. Though females are reported as entering SF-CSOC with significantly higher scores on the peer and non-family/caregiver relationship strength item and school system and educational plan strength item, males enter with a higher score on the relationship permanence item and the Other Strengths factor. The Cohen's d effect sizes are all either in the small range or the confidence interval includes zero, suggesting that the differences, while significant, are not clinically meaningful.

Table 4

Gender Differences[a] *of Youth Strengths Upon Entry to the Children's System of Care*

Outcome Modeling Options	female mean (SD)	male mean (SD)	t (df)	pvalue	Effect Size (cohen's d)	CI for cohen's d
Individual Items						
Family/Caregiver Relationship Strength	1.87(.81)	1.93(.80)	-1.82 (2047)	0.07	-.07	-.16, .01
Peer and Non-Family/Caregiver Interpersonal Relationships Strength	1.93(.85)	1.84(.85)	2.35 (2047)	0.02	0.12	.02, .19
School System and Educational Plan Strength	1.87(.90)	1.70(.89)	4.21 (2047)	<.001	0.19	.10, .28
Extracurricular Activities and Talents Strength	1.89(.93)	1.94(.92)	-1.25 (2047)	0.21	-.05	-.14, .03
Spiritual/Religious Beliefs and/or Involvement Strength	1.73(1.22)	1.82(1.23)	-1.62 (2047)	0.12	-.07	-.16, .02
Relationship Permanence Strength	1.93 (.87)	2.06(.92)	-3.14 (2047)	0.002	-.14	-.23, .05
Sub-scales						
Family and Relationship Factor	.007(.40)	-.004(.39)	.65 (2047)	0.51	0.04	-.05, .12
Other Strengths Factor	-.02(.31)	.02(.32)	-2.74 (2047)	0.006	-.13	-.22, .04
Overall Composite Score	11.24(3.27)	11.32(3.36)	-.51 (2047)	0.61	0.003	-.09, .09

[a]There are 791 females and 1,258 males in the sample.

Next, for research question three, one way ANOVA tests, p-value, and effect sizes (with confidence intervals) were run to determine whether there were any significant differences between the five age categories (5-8, 9-12, 13-15, 16-18, and 19-22 years) on the three youth strength modeling options (see Table 5.1). The F tests and p-values suggest that there are significant differences on strengths between at least one pair of age groups for the family/caregiver relationship strength item, the school system and education plan strength, extracurricular activities and talents strength, the Family Relationship Factor, the Other Strengths Factor, and the Overall Composite Score. However, the eta squared effect sizes are all in the small range, suggesting that only a small amount of variation in the strength variables are explained by the age groups.

Table 5.1

Age Differences of Youth Strengths Upon Entry to the Children's System of Care

	5-8 Years Mean (SD)	9-12 Years Mean (SD)	13-15 Years Mean (SD)	16-18 Years Mean (SD)	18-22 Years Mean (SD)	Ratio test for constant variance	between groups F(df)	pvalue	Effect Size eta squared(CI)
Individual Items									
Family/Caregiver Relationship Strength	1.98(.79)	2.00(.79)	1.87(.78)	1.81(.82)	1.96(.73)	1.12	5.29 (4)	<.001	.01(.00, .02)
Peer and Non-Family/Caregiver Interpersonal Relationships Strength	1.88(.87)	1.83(.88)	1.88(.83)	1.93(.84)	1.76(.78)	1.13	1.12 (4)	0.35	-
School System and Educational Plan Strength	1.94(.88)	1.83(.89)	1.64(.88)	1.70(.92)	1.88(.83)	1.11	8.23 (4)	<.001	.02 (.00, .03)
Extracurricular Activities and Talents Strength	1.96(.97)	2.06(.88)	1.93(.89)	1.79(.93)	1.72(.98)	1.24	6.65 (4)	<.001	.01(.00, .02)
Spiritual/Religious Beliefs and/or Involvement Strength	1.72(1.32)	1.82(1.26)	1.87(1.18)	1.74(1.17)	1.92(1.26)	1.13	1.19 (4)	0.31	-
Relationship Permanence Strength	2.07(.94)	2.04(.91)	2.03(.90)	1.93(.89)	2.08(.76)	1.07	1.85 (4)	0.12	-
Sub-Scales									
Family and Relationship Factor	.02(.40)	.03(.41)	-.01(.37)	-.04(.38)	-.03(.38)	1.09	2.92(4)	0.02	.006(.00, .01)
Other Strengths Factor	7.49(2.86)	7.54(2.84)	7.31(2.56)	7.16(2.54)	7.28(2.70)	1.13	4.41(4)	0.001	.009(.00, .02)
Unidimensional (Composite Score)	11.55(3.33)	11.58(3.49)	11.20(3.19)	10.90(3.24)	11.32 (3.18)	1.10	3.75 (4)	0.005	.007(.0007, .01)

Though the effect sizes for the significant F tests were quite small, the Bonferroni post-hoc analyses identified differences between the pairs of groups listed in Table 5.2. Overall, it

appears that older youth generally enter San Francisco's Children's System of Care with more strengths than younger youth. The largest difference between a pair of age groups was found on the overall composite score between 9-12 year olds and 16-18 years olds, with 16-18 year olds scoring, on average, .68 points higher (p=.007) on the strengths composite score than 9-12 year olds.

Table 5.2

Bonferroni Post-Hoc Tests Showed Significant Differences Between the Groups[a]

	Mean difference (Group A minus Group B)	p-value
Individual Items		
Family / Caregiver Strength		
5-8years and 16-18 years	-.17	0.007
9-12 and 16-18 years	-.19	0.001
School System and Educational Plan Strength		
5-8 years and 9-12 years	-.30	<.001
9-12 years and 13-15 years	-.19	0.006
5-8 years and 16-18years	-.24	<.001
Extracurricular Activities and Talents Strength		
5-8 years and 16-18 years	-.18	0.03
9-12 years and 16-18 years	-.28	<.001
Sub-Scales		
Family and Relationship Strength Factor		
9-12 years and 16-18 years	-.07	0.03
Other Strengths Factor		
5-8 years and 16-18 years	-.07	0.004
9-12 years and 16-18 years	-.07	0.004
Overall Composite Score		
9-12 years and 16-18 years	-.68	0.007

[a] Group differences that were non-significant at the .95 level are not shown in the table.

Next, for research question three, group means (and standard deviations), one way ANOVA tests, p-values , and effect sizes (with confidence intervals) were run to determine whether there were any significant differences between the seven ethnicity categories (African-American, Asian, Latino, Multi-Ethnic, Pacific Islander/Hawaiian, White, Other) on the three youth strength modeling options. The results of the one-way ANOVA tests are reported in Table 6.1. The F tests and p-values suggest that there are significant differences on strengths between at least one pair of ethnicity groups for every strength variable except the family/caregiver relationship strength (F=1.28, df=6, p=.26). Though the eta squared effect sizes suggest that

group differences are generally small, the effect sizes move towards the medium effect range for the spiritual religious beliefs and/or involvement strength (ES: .03, CI: .02, .05) and the relationship permanence strength (ES: .04, CI: .02, .06).

Table 6.1

Ethnicity differences of youth strengths upon entry into the ChidIren's System of Care.

	African-American group mean	Asian group mean	Latino group mean	Multi-ethnic group mean	Pacific Islander/Hawaiian group mean	White group mean	Other group mean	Ratio Test for Constant Variance	between groups F(df)	pvalue	Effect Size eta squared(CI)
Individual Items											
Family/Caregiver Relationship Strength	1.85(.82)	1.94(.79)	1.91(.79)	2.07(.81)	1.93(.76)	1.96 (.80)	1.90(.79)	1.08	1.28(6)	0.26	-
Peer and Non-Family/Caregiver Interpersonal Relationships Strength	1.86(.85)	1.75(.87)	2.00(.81)	2.11(.85)	2.02(.93)	1.79(.90)	1.85(.83)	1.15	5.02(6)	<.001	.01(.00, .02)
School System and Educational Plan Strength	1.67(.90)	1.83(.87)	1.84(.93)	1.77(.86)	1.63(.84)	1.83(.86)	1.71(.92)	1.11	2.86(6)	0.01	.01(.00, .01)
Extracurricular Activities and Talents Strength	1.93(.89)	1.77(.98)	1.97(.92)	2.18(.81)	2.02(.85)	2.01(.94)	2.01(.91)	1.21	3.74(6)	0.001	.01(.00, .02)
Spiritual/Religious Beliefs and/or Involvement Strength	1.82(1.20)	1.43(1.30)	2.02(1.13)	1.68(1.23)	2.20(1.02)	1.67(1.28)	1.94(1.21)	1.27	11.29(6)	<.001	.03(.02, .05)
Relationship Permanence Strength	1.81(.87)	2.32(.86)	2.03(.91)	1.99(.96)	1.98(.90)	2.02(.92)	1.88(.94)	1.12	14.73(6)	<.001	.04(.02, .06)
Sub-Scales											
Family Relationships Factor	3.67(1.42)	4.26(1.32)	3.95(1.41)	4.06(1.32)	3.91(1.39)	3.98(1.43)	3.78(1.49)	1.13	4.71(6)	<.001	.01(.00, .02)
Other Strengths Factor	7.28(2.67)	6.79(2.81)	7.83(2.57)	7.74(2.41)	7.89(2.56)	7.31(2.77)	7.51(2.79)	1.17	8.30(6)	<.001	.02(.01, .04)
Overall Composite Score	10.95(3.33)	11.04(3.37)	11.78(3.22)	11.79(2.89)	11.80(2.99)	11.29(3.44)	11.29(3.60)	1.25	4.12(6)	<.001	.01(.00, .02)

Since significant differences were identified in some of the one-way ANOVAs above, Bonferroni follow-up tests were run to identify which groups were significantly different, as reported in Table 6.2. A few patterns emerge. First, the overall composite score suggests that African-American youth enter the SOC with an average composite strengths score that is nearly one point higher than Latino youth (p<.001). The two subscales provide additional information, suggesting that African American youth enter with higher scores on the Family / Caregiver Strength Factor (.08 points, p=.006) and the Other Strengths Factor (.08 points, p<.001), though the mean differences are quite small.

The individual strengths items provide the most detailed information and help contextualize the significant group mean differences between African American and Latino youth on the overall composite and the two Factors. African American and Latino youth enter the SOC with significant differences on mean scores on only two strength items: school system and education plan (mean difference: .16. p=.01) and relationship permanence (mean difference: .22, p<.001), suggesting that these items are driving the significant mean differences found in the overall composite score and the sub-scales. This provides support for using the item-level strengths.

Also, Asian youth were reported as entering with a slightly higher score than Latino youth on the overall composite score (mean difference: .02, p=.01) and the Family/Caregiver Strength Factor (mean difference: .11, p<.001). Only considering this information, it would appear that family relationships were driving the difference in strengths on the overall composite score. However, a closer look at the individual strength items suggests that while Asian youth were reported as entering the SOC with higher scores than Latino youth on the peer relationships (.25, p<.001) and the extracurricular activities items (.2, p=.01), Latino youth entered with a higher mean score on the relationship permanence item (-.28, <.001). Given this, it is unclear why Asian youth, and not Latino youth, have a higher score on the Family/Caregiver Strength Factor, when the relationship permanence strength item has a high loading score (.76) on this factor.

Table 6.2

Bonferroni Post-Hoc Tests Showed Significant Differences Between the Groups[a]

	Mean difference (Group A minus Group B)	p-value
Individual Items		
Peer and Non-Family/Caregiver Interpersonal Relationships Strength		
Asian and Latino	0.25	<.001
Asian and Multi-Ethnic	0.36	0.02
School System and Educational Plan Strength		
African-American and Asian	0.17	0.05
African-American and Latino	0.16	0.01
Extracurricular Activities and Talents Strength		
Asian and Latino	0.2	0.01
Asian and Multi-Ethnic	0.41	0.01
Spiritual/Religious Beliefs and/or Involvement Strength		
African-American and Asian	-.40	<.001
White and Latino	0.34	<.03
Asian and Pacific Islander	0.78	0.001
Asian and Other	0.51	0.03
Relationship Permanence Strength		
White and Asian	0.3	0.005
African-American and Asian	0.5	<.001
African-American and Latino	0.22	<.001
Asian and Latino	-.28	<.001
Asian and Other	-.44	0.004
Sub-Scales		
Family and Relationships Factor		
African-American and Latino	0.08	0.006
Asian and Latino	0.11	<.001
Asian and Multiethnic	-.33	0.07
Other Strengths Factor		
African-American and Asian	0.13	<.001
African-American and Latino	0.08	<.001
Overall Composite Score		
African-American and Latino	0.83	<.001
Asian and Latino	0.02	0.01

[a] Group differences that were non-significant at the .95 level are not shown in the table.

Discussion

The purpose of this paper was to 1) explore the prevalence of reported strengths (including the family/caregiver relationship strength, peer and non-family/caregiver interpersonal relationships strength, school system and educational plan strength, extracurricular activities and talents strength, spiritual/religious beliefs and/or involvement strength, and relationship permanence strength) in San Francisco's public System of Care, 2) explore the psychometric properties of the CANS' Strengths, and 3) explore group differences (in age, gender, and ethnicity) using three alternative modeling options. These analyses suggest that individual items (modeling option #1) may be the most useful way to model the strengths items. The individual item analyses were generally easier to interpret. For example, on average, youth were reported as entering SF-CSOC with a relationship permanence strength score of 2.01 (see Table 2). In the San Francisco CANS Scoring Manual (2010), a score of 2 is described as: "strength exists but require some strength building efforts in order for it to serve as a focus of a strength-based plan" (pg. 51). Thus, the reported average score of 2.01 could be interpreted as indicating that youth entered SF-CSOC, on average, with at least one permanent relationship available that could be used to increase desired behaviors, though efforts may be needed to build up this strength before it could be used to its full potential. It is unclear what was added by the factor scores (modeling option #2), particularly given that the two factor model fit was only adequate and the interpretation of the two factors was unclear. (One option might be for future analyses to use the Family and Caregiver factor as a proxy for family-centered relationships, and include this factor score in future analyses, but use the other four items individually since the Other Strengths factor appeared to be a "catch-all" factor). The overall composite score (modeling option #3) was the least useful because it masked the contribution of each item to the composite, limiting its clinical relevance (for example, it is impossible to know whether a total composite score of 6 was the result of two individual items with a score of 3 and four items with a score of 0, or six individual items each with a score of 1).

Research question one concerned the prevalence of youth strengths, using the three alternative strengths modeling strategies, upon entry to San Francisco's Children's System of Care (SF-SOC). Regardless of how strengths are modeled in the three options used, strengths appear to reasonably prevalent in SF-CSOC. The average youth in the sample was reported as entering SF-CSOC with a score of nearly 2 points on each strength item, suggesting these strengths were available at entry for use in a strengths-based treatment plan (as defined by the SF CANS User Manual, see Appendix A). However, the large standard deviations imply that there is variation on strengths items, suggesting that there may be some youth with extremely high or extremely low strengths scores. Further analyses should examine the data for extreme groups and look for differences between groups.

Question two explored the factor structure of the SF-CANS Strengths domain. A previous principle components analysis (PCA) with a varimax rotation had suggested two factors. A Confirmatory Factor Analysis (CFA) was used to see if a two factor structure was a better fit for the data than a one factor structure. A two factor model was a better fit than the one factor model, though the two factor model fit was only adequate (see Figure 2 and Table 3). The meaning of the factors are difficult to interpret. Further theoretical clarification and more psychometric exploration is needed.

Question three examined the demographic differences (on gender, age, and ethnicity) in strengths, using the three modeling options (see Tables 4-6.2). Regardless of modeling option,

males and females appear to enter SF-CSOC with similar levels of strengths. Overall, older youth appeared to enter SF-CSOC with more strengths than younger youth, though the differences between groups and effect sizes were fairly small. Some differences between ethnic groups were found on youth strengths at entry to SF-CSOC. African American youth appeared to enter SF-CSOC with higher strengths than Latino youth. Asian youth appeared to enter with higher strengths than Latino youth. However, the differences between groups and the effect sizes were small. A closer examination of the theoretical and empirical literature on gender, age, and ethnic group differences among youth receiving public mental health services may be useful to provide context for these analyses and to further interpretation.

Limitations.

This study has several limitations, many of which are related to the use of administrative data (data collected for non-research purposes, but used for research purposes, Garrett et al., 2010). The San Francisco CANS data relies solely on clinician self-report, meaning that the data is always at risk of clinician bias and documentation error. There is no youth report or parent report to act as a check on the clinician's report. Observation (such as videotaping assessments) has not been used, either in validity studies of the CANS or in the collection of the data used in this dissertation, further limiting the conclusions that can be drawn. Also, the number of contacts within 30 days is not tracked in SF-CSOC. Though clinicians are encouraged to talk to as many stakeholders as possible in the youth's life, there is no way to know how many stakeholders a clinician spoke with, either in person or by telephone, before filling out the SF-CANS online. Also, the SF-CANS User Manual is long and it is possible that clinicians might not have the time or motivation to reference the manual every time they are scoring an assessment, thus resulting in inaccurate scoring. The percentage of youth referred for services, where they were referred from, the percentage who went to services, and the type, quantity, and quality of interventions (if any) are not reliably tracked.

Similarly, though clinicians are encouraged to establish a collaborative partnership while filling out the SF-CANS with the youth and family, and the SF-CANS includes a place for a caregiver signature, there is no reliable way to know if clinicians are working collaboratively or whether or not the caregiver signed, let alone whether or not the caregiver felt coerced to sign, secretly felt that the clinician was condescending or rude, or had questions that were still unanswered, even if they signed. If there is a blank space on any of these questions, it is impossible to differentiate between cases where a caregiver was merely unreachable during the time of the assessment, or whether there was not a caregiver in the youth's life at all.

The SF-CANS is limited in the strengths items it assesses. The SF-CANS only includes six youth strength items. (Three of the items - Optimism, Psychological Well-Being, and Integration into Community Life - were cut by system leadership.) Thus, there are many youth strengths that are not included in the SF-CANS, and it is unclear why the strengths domain contains certain strength items and not others. Also, though the SF-CANS may be considered a strengths-based assessment, using the broad definition above, most of the hundreds of possible items are items that measure problems, symptoms, and deficits. There is only one Strengths domain included in a version of the CANS, with eight or ten strength items included at most. Some strengths empirically demonstrated to be related to improved functioning (e.g. self-efficacy, quality of neighborhood or community resources) are not included. The addition of other strengths items would provide a more nuanced picture of the strengths profiles of youth entering SF-CSOC.

Also, since administrative data is not collected with research as the primary goal, the reliability and validity of the data (which is of great concern to researchers) may not have been prioritized by the people collecting the administrative data. For example, the spiritual religious item (pg. 55) is confusing because the manual does not clarify whether the item can be interpreted as 1) the youth is involved in a religious community or 2) the youth identifies with a spiritual or religious practice and finds strength in this, or both of these. Researchers can only hypothesize as to what the clinician intended when interpreting this item. Future psychometric testing would be helpful here in order to establish the validity of the SF-CANS items. Since every community has its own version, perhaps every community should consider conducting its own validity study in order to truly understand what the CANS items mean, and San Francisco could lead the way in this. Qualitative work might also be helpful here, for example interviews or focus groups with clinicians who use the SF-CANS could shed light on how clinicians understand, use, and interpret the Strengths domain items, the referral process and how many took up services, reasons for termination, what interventions are provided, if any, and other factors that are not currently reliably tracked in SF-CSOC.

Also, a closer examination into the theoretical and empirical literature on gender, age, and ethnic group differences among youth receiving public mental health services is warranted to provide context for these analyses and to further interpretation. This would be helpful in contextualizing group differences on all the CANS domains. However, since strengths have generally received less attention than emotional and behavioral problems in the mental health services literature at large, further theoretical and empirical work on gender, age, and ethnic group differences on strengths would further the work of this paper considerably.

PAPER 2

Strengths and Psychiatric Symptom Severity of Youth Entering San Francisco's Public System of Care

Sarah Accomazzo

School of Social Welfare
University of California, Berkeley

Introduction

This paper describes the profiles of youth entering San Francisco's Children's System of Care (SF-CSOC), with particular attention to levels of strengths at initial assessment, including an examination of the contribution of clinician and agency to variation in psychiatric symptom severity. Behavioral theory and related research in experimental and applied settings is suggested as a theoretical framework for strengths approaches in Systems of Care settings. Over the past thirty years, the System of Care (SOC) approach has been the cornerstone of the United States government's attempts to respond to children and adolescents ("youth") with severe emotional and behavioral problems (problems that are expressed through emotions or behaviors that disrupt life functioning or developmental tasks) (e.g. CMHI, 2011). Intended as a guiding philosophy and values framework for mental health and substance abuse service provision for youth with severe emotional and behavioral issues, the System of Care approach has been implemented in 173 communities throughout the United States and its territories since 1984, with a total of approximately two billion dollars in federal funding. System of Care (SOC) policy emphasizes strengths in addition to needs/problems/symptoms. Specifically, the most recent updated definition of a SOC offered by the federal government (Stroul et al., 2010), states that SOCs should be constructed based on three core values, the first of which is: "Family driven and youth guided [care], with the strengths and needs of the child and family determining the types and mix of services and supports provided" (pg. 6). However, in Systems of Care research, strengths are generally given less attention than emotional and behavioral problems, even though federal policy emphasizes strengths.

Theoretical Conceptualizations of Strengths in Systems of Care (SOC): The Under-Valued Contribution of Behavioral Theory and Related Research.

Perhaps underlying the lack of focus on strengths is the fact that theoretical frameworks for focusing on strengths in addition to needs, problems, or symptoms, are not articulated in the federal System of Care (SOC) policy documents. (For example, Stroul, Blau, & Friedman's (2010) policy brief mentioned strengths but did not define or theorize on strengths.) Behavioral theory and related research in applied and experimental contexts provide a rational for a focus on strengths (e.g. Goldiamond, 1974; Madden, 2013; Staats, 2012). With a few exceptions, these approaches are typically misunderstood and/or ignored (Gambrill, 2013; Thyer, 2005) and are rarely described in strengths literature outside of the behavioral literature. (For example, both McCammon (2012) and Saleebey (2009) did not mention behavioral theory and related research in their discussion of strengths). However, these provide theoretical and empirical support based on decades of research in both experimental and applied settings for decreasing undesirable behaviors by increasing desirable ones. (Approaches that emphasize strengths also draw on resilience theories and frameworks, health promotion approaches, positive psychology approaches, social capital theories, empowerment theories, and the recovery and peer support literatures, among others.)

Empirical research in applied and experimental settings related to behavioral theory shows that behavior is acquired or altered via learning processes (Fisher, et al., 2011; Madden, 2013; Staats, 2012). Behavior is understood in the context of environmental contingencies, defined as "the relationship of behavior to its consequences and the relation of the behavior to the antecedent conditions under which this behavior occurs" (Schwartz & Goldiamond, 1975, pg. 17). Certain consequences are contingent on certain behaviors, given a particular environmental

context. Behavioral interventions focus on strengths, drawing on core principles of behavioral theory and related research (e.g. see literature on Applied Behavioral Analysis). Applied Behavioral Analysis (ABA) is "a scientific approach to the study of behavior paying careful attention to the social validity of concerns addressed and related outcomes. This applies to clients as well as to their significant others, those who influence clients and who may be affected by interventions" (Gambrill, 2013a, pg. 121). Empirical studies within ABA apply behavioral theory principles to develop, implement, and test clinical interventions with a goal of client behavior change.

Empirical studies within Applied Behavioral Analysis (ABA) emphasize that behavior is functional, or serves a purpose for the individual (Layng, 2009). Maladaptive behaviors are understood through the behavior's function - the environmental contingencies affect behavior (Layng, 2009). If behaviors are viewed as functional, even distressful or seemingly bizarre behaviors can be understood as a creative, though perhaps costly, responses to difficult life circumstances (Gambrill, 2013b, pg. 132). Decades of research provide evidence for the effectiveness of interventions based on this conceptualization of behavior (e.g. Positive Behavior Supports (Horner, Sugai, & Anderson, 2010), early and intensive intervention for children with autism spectrum disorder (Smith, 2010), and behavioral treatment for enuresis (Houts, 2006). (See also Madden, 2013.)

Behavioral Theory's "Constructional Approach" and Implications for the Relationship between Strengths and Psychiatric Symptom Severity. Behavioral theory's constructional approach provides theoretical and empirical grounds for focusing on strengths. In describing the difference between a pathological approach and a constructional approach, Goldiamond (Schwartz & Goldiamond, 1975) wrote:

> One of the aims of the [constructional] interviewer is to give the client confidence that he can assume control over the solution to his own problems. What the client may interpret as failure and deficit on his part the therapist may actually interpret as signs of success and strengths…therapists…have noted how difficult it often is to get patients to indicate their strengths, or what is good about them. This difficulty is not necessarily a sign of malfunctioning. Rather it suggests that the patient is in tune with a culture that is heavily oriented pathologically and has applied this orientation to himself… (pg. 78-79, brackets added).

This approach does not ignore the existence of problems, it shifts the focus towards identifying related environmental contingencies and constructing positive alternative repertoires as the means of reducing problems. A repertoire is defined as a strength, skill, asset, and environmental or relational contingency (Goldiamond, 1974, pg. 80). In the constructional approach, behaviors are changed through an increase in "strengths" (alternative available repertoires) and/or the construction of "strengths" (new positive alternative repertoires).

For example, Goldiamond (1974) used the term "strength" throughout his discussion of the constructional approach and its applications to reducing emotional and behavioral problems. Goldiamond (1974) describes that behavioral theory's constructional approach emphasizes self-determination of the client (pgs. 73-74), empowering the client to solve his/her own problems (pg. 77), the clinician's attention to what is going right in addition to what is going wrong (pg. 79-81), collaboration with the client (pg. 73), the role of context and the environment (pg. 69), the idea that all behavior typically serves a function even if it is not immediately apparent to

either the clinician, the client, or family (pg. 80), and turning away from a focus on deficits and pathology (pg. 78-79). Clients' strengths (e.g. available alternative repertoires) are drawn on to increase desirable behaviors and reduce maladaptive behavior and symptoms. New strengths (e.g. new positive alternative repertoires) can be constructed to decrease maladaptive behavior and symptoms.

Lack of Attention to Strengths in Systems of Care Research.

Many studies with a System of Care sample do not measure strengths at all (e.g. Snyder et al., 2012, and Walrath et al., 2006, do not report on strengths). Studies that do include strengths have several limitations. Some SOC studies that include strengths report problem and need prevalence (such as prevalence of psychiatric symptoms, risk behaviors, trauma experiences), but do not report on strengths prevalence (e.g. Griffin et al., 2009). Some SOC studies may report demographic characteristics of the sample but do not report on prevalence of either youth emotional and behavioral problems or strengths (e.g. Anderson et al., 2008; Whitson et al., 2010). Another issue in the SOC literature is that many studies have not accounted for clustering effects (or violations, due to the effects of unmeasured, shared variables, of the assumption that all subjects are independent). Youth, clinician, or agency are all potential clustering variables in Systems of Care data (e.g. some youth who received services from the same clinician may be more similar to each other than to youth who received services from a different clinician). For example, Snyder et al (2012) did not account for clustering effects in their analysis of a sample from a System of Care.

The Current Study.

Though federal youth mental health policy emphasizes the role of youth strengths in providing optimal care for clients (e.g. Stroul et al., 2010), strengths are often given less attention than emotional and behavioral problems (e.g. McCammon, 2012). Theoretical conceptualizations for including a focus on strengths in Systems of Care (SOCs) are rarely articulated in these policy documents. Given the gaps in Systems of Care literature regarding strengths, the following paper examines the strengths and needs profiles of youth upon entry (at initial assessment) to San Francisco's Children's System of Care (SF-CSOC). Behavioral theory and related research both in applied and experimental settings provide a framework and empirical underpinnings for conceptualizing the role of strengths in Systems of Care, and particularly for the value of increasing strengths in order to decrease symptoms. The following research questions are explored: 1) What are the demographics (age, gender, and ethnicity) of youth entering SF-CSOC? What is the prevalence of youth strengths variables, the psychiatric presentation composite, and the severity of lifetime traumatic experiences composite for youth entering SF-CSOC? 2) Are higher scores on strength items upon entry to SF-CSOC associated with lower severity of psychiatric symptoms, adjusting for child characteristics (age, gender, ethnicity, and reported severity of lifetime traumatic experiences)? 3) What is the contribution of agency and clinician to variation in psychiatric symptom severity, adjusting for child level characteristics?

Method

Study Design and Measures.

In early 2008, San Francisco's Children's System of Care (SF-CSOC) adopted the Child and Adolescent Needs and Strengths (SF-CANS) instrument as its primary assessment tool for

all youth programs. Consistent with the tool's basis in Communimetrics theory (Lyons, 2009), which emphasizes item-level reliability and validity, different stakeholder groups have used different configurations of strength items in their own settings, and have added to the bank of strengths items used as part of the CANS. SF-CSOC went through a group process, involving consultation with subject experts and family members, in order to adopt the items for the SF-CANS.

The SF-CANS includes fifteen domains with more than 100 items. Fourteen of the domains were designed to identify type and acuity of clinical and psychosocial problems (e.g. psychiatric symptom severity, risk behaviors) and one domain was designed to identify youth strengths (SF CANS 5-18 Scoring Manual, 2010). For each item, assessors assign a rating using a four-point likert scale. The fourteen problem domains are designed to assess severity of problems (instead of counting frequency of problems), while the Strengths domain is designed to track the presence of strengths that could be used in a strengths-based plan. After a case is opened, clinicians have thirty days to submit a CANS initial assessment online. During these thirty days, the clinicians are expected to identify and contact multiple stakeholders in the youth's life (e.g. parents, caretakers, other family, teachers, caseworkers, etc.) to gather as much information as possible. The number of contacts within thirty days is not tracked in SF-CSOC, but could vary from one contact to ten or more (as estimated by a senior SF-CSOC administrator). Though clinicians are encouraged to talk to as many stakeholders as possible in the youth's life, there is no way to know how many stakeholders a clinician spoke with before filling out the SF-CANS online and whether the contacts were in person or by telephone. Clinicians are encouraged to establish a collaborative partnership while filling out the SF-CANS with the youth and family, and the SF-CANS includes a place for a caregiver signature, but there is no reliable way to know if clinicians are working collaboratively or whether or not the caregiver signed, let alone whether or not the caregiver felt coerced to sign, secretly felt that the clinician was condescending or rude, or had questions that were still unanswered. The number of submitted CANS assessments that are signed versus un-signed are not reliably tracked. Also, if there is a blank space on any of these questions, it is impossible to differentiate between cases where a caregiver was merely unreachable during the time of the assessment, or whether there was not a caregiver in the youth's life at all.

An initial assessment is estimated to take around two hours to enter into the online database, but total amounts of time may vary and are untracked in this system. Re-assessments must be completed at least every six months while the youth's case is marked "open" in the San Francisco Children's System of Care (SF-CSOC) database (actual number of assessments completed and number of contacts prior to re-assessment is unknown) and when the youth terminates contact with SF-CSOC. Reasons for termination are not reliably tracked, but could include dropping out, moving out of area, or completing a course of therapeutic intervention successfully, partially successfully, or unsuccessfully. The percentage of youth referred for services, where they were referred from, the percentage who went to services, and the type, quantity, and quality of interventions (if any) are not reliably tracked. Since August 1st, 2010, CANS data have been stored in an online, HIPAA-compliant database. De-identified data were extracted for this study by a data manager and stored by the researcher according to procedures approved by the institutional review company, Ethical and Independent Review Services.

Reliability and Validity of the CANS. The CANS has been shown to perform reasonably well on some traditional psychometric standards. Versions of the CANS have been shown to be reliable at the item level, for example, one study found an intra-class correlation of .85 between

researchers and clinicians in sixty randomly drawn case files. (Anderson et al., 2003). Acceptable item-total correlations were demonstrated in a study with a sample of 6010 youth in a public system, for the strengths domain (Cronbach's alpha = .70) and the psychiatric symptom severity domain (Cronbach's alpha = .71) (Lyons, 2009). Also, versions of the CANS have demonstrated convergent validity. For example, CANS scores were shown to be correlated (r=.63) with another self-report measure of children's emotional and behavioral problems, the Child and Adolescent Functional Assessment Scale (e.g. Lyons et al., 2004). A version of the CANS accurately distinguished youth (n=150) who were placed in inpatient care and youth who were placed in community-based outpatient care youth (Anderson et al., 2001).

Sample.

The sampling frame for this study included youth aged 5-22 who received an initial CANS assessment between June 2011 and August 2013 and whose assessing clinician entered the initial assessment into an online database within the timeline required by the city, within 45 days of first of point contact (n=2,402). (Youth who entered into the system but whose clinician did not meet the 0-45 day deadline were excluded from the sample due to worry that the assessment was no longer capturing the youth's presentation upon entering the system). The final sample includes all youth who did not have any missing data on any of the variables included in the final analyses (n=2049). There are no statistically significant differences in the demographics between the sample used for this paper and the group who were dropped from the sampling frame due to missing data. Also, since there are no known reasons for why demographic data may be missing, and no patterns of missing data were spotted by the researcher, the missing data is assumed to be Missing Completely at Random (MCAR; missing due to variable that are unrelated to the analyses), and thus not a concern for this analysis.

Youth were referred for an initial assessment (from a variety of sources that are currently untracked in this public system, but including family members, school teachers or counselors, child welfare worker, probation officers, and others) due to alleged emotional, behavioral, and environmental issues that were said to impair their functioning in their family, school, or community. Socio-economic status of children and families is not currently tracked in this SOC because 98% of client services are paid for by billing Medicaid.

Youth in this sample (n=2049) had an average age of 12.5 (SD: 3.85). Table 1 shows the other demographics of the sample youth. Forty-three agencies and 465 clinicians are included in this sample. On average, there are 48 youth per agency (SD: 49.52, Range: 1-186), 4 youth per clinician (SD: 6.31, Range: 1-58), and 10 clinicians per agency (SD: 11.0, Range: 0-52). Though the standard deviations and ranges are quite large for each of these averages, multi-level modeling can adequately handle clusters of different sizes (see above).

Table 1

Youth Demographics Upon Entry to San Francisco's Children's System of Care (n=2049).

Characteristic	n	%
Sex		
Male	1258	61
Female	791	39
Ethnicity		
African-American/Black	685	33
Asian	446	22
Latino	571	28
Multi-Ethnic	72	4
Pacific Islander/Native Hawaiian	44	2
White	163	8
Other	68	3

Variables.

Dependent Variable.

Psychiatric Symptom Severity (Psych). The "Psychiatric Symptoms" domain is designed to track psychiatric symptom severity in the last thirty days. This domain consists of 9 items: psychosis, impulse/hyper, depression, anxiety, oppositional, conduct, substance use, somatization, and anger control. Each item receives a score on a four point likert scale (0 = no evidence of any needs, 1 = a dimension that requires monitoring watchful waiting, or preventive activities, 2= a dimension that requires action to ensure that this identified need is addressed/child meets criteria for a DSM diagnosis in this category, 3 = a dimension that requires immediate or intensive action). This variable was created by adding up the 0-3 score for the nine items in the domain. This continuous variable ranges from 0-27.

Level One Covariates (youth = i).

Youth. Youth is a unique identification number. Youth is the level one variable, and so youth are considered nested in clinicians nested in agency.

The following six items are part of the Youth Strengths domain. The Youth Strengths domain is designed to track the presence of accessible youth strengths in the past thirty days that could be used as the basis of a strengths-based plan. Each item is rated on a 0-3 likert scale and was reverse coded to assist in interpretation (with reverse coding: 3 = strengths exist that can be used as a centerpiece for a strength-based plan, 2 = strengths exist but require some strength building efforts in order for them to serve as a focus of a strength-based plan, 1 = strengths have been identified but they require significant strength building efforts before they can be effectively utilized in as a focus of a strength-based plan, 0 = efforts are needed in order to identify potential strengths for strength building efforts).

Family/Caregiver Relationship Strength (famnew1).

Peer and Non-Family/Caregiver Interpersonal Relationships Strength (interpnew1).

School System and Educational Plan Strength (educnew1).

Extracurricular Activities and Talents Strength (talnew1).

Spiritual/Religious Beliefs and/or Involvement Strength (spirnew1).

Relationship Permanence Strength (relnew1).

Male is a categorical variable for gender, coded male = 1 and female = 0.

Ethnicity. For the seven ethnic groups (African-American/Black, Asian, Latino/a, Multi-ethnic, Native Hawaiian or Other Pacific Islander, White, and Other), white was used as the reference category and dummy variables were constructed for the other categories: African-American/Black (AfAm = 1, all others = 0), Latino (Latino=1, all others = 0), Asian (Asian=1, all others = 0), Multi-Ethic (Multi-Ethnic = 1, all others = 0), Pacific-Islander (pacisl =1, all others = 0), and Other (Other=1, all others = 0).

Age_cent. Age in years at initial assessment is a continuous variable ranging from 5-22, mean-centered in order to aid in interpretation.

Lifetime Trauma Experience Severity at Initial Assessment (Trauma). This variable is designed to track lifetime exposure to trauma. This continuous composite variable was created by adding together the score for each of the twelve trauma types items on the CANS initial assessment (physical abuse, sexual abuse, emotional abuse, neglect, medical trauma, witness family violence, witness community violence, school violence, natural/man-made disaster, traumatic grief/separation, war affected, terrorism affected) rated on a 0-3 likert scale, with 0 = a dimension where there is no evidence of any trauma of this type, 1 = a dimension where a single event trauma occurred or suspicion exists of trauma experiences, 2 = a dimension on which the child has experienced multiple traumas, and 3 = a dimension which describes repeated and severe trauma with medical and physical consequences. The range of this variable is 0-36.

Level Two Covariates (clinician = j).

Clinician. Each clinician has a clinician identification number (CIN) assigned to him or her. Clinician is the level two variable, so youth is nested in clinician.

Level Three Covariates (agency = k).

Agency. Each agency has a unique agency identifier assigned to it. Agency is the level three variable, so youth is nested in clinician is nested in agency.

Statistical Analysis.

The software package Stata13 was used for all analyses. Stata is particularly strong at handling missing data and unbalanced designs (due to options using a variety of maximum likelihood-based estimates to calculate coefficients, offering flexible ways to handle variance and covariance structures, and the ability to calculate robust standard errors) (Statacorps, 2013).

For research question 1, frequencies, means, standard deviations, and ranges were calculated for the variables presented in the analysis. The following random-intercept model was used to answer research questions two and three (please see Appendix E for a detailed discussion of the interpretation of this model):

$$
\begin{aligned}
Psych_{ijk} = \; & \beta_1 + \beta_2 Famnew1_{ijk} + \beta_3 Interpnew1_{ijk} + \beta_4 Educnew1_{ijk} \\
& + \beta_5 Talnew1_{ijk} + \beta_6 Spirnew1_{ijk} + \beta_7 Relnew1_{ijk} \\
& + \beta_8 Trauma_{ijk} + B_9 Male_{ijk} + \beta_{10} Age_{ijk} + \beta_{11} Black_{ijk} \\
& + \beta_{12} Asian_{ijk} + \beta_{13} Latino_{ijk} + \beta_{14} Multi_{ijk} + \beta_{15} Pacisl_{ijk} \\
& + \beta_{16} Other_{ijk} + \zeta_{jk}^{(2)} + \zeta_k^{(3)} \\
& + \epsilon_{ijk}, \quad \zeta_{jk}^{(2)} \sim N\left(0, \psi^{(2)}\right), \quad \zeta_k^{(3)} \sim N\left(0, \psi^{(3)}\right), \\
& \epsilon_{ijk} | \zeta_j \sim N(0, \theta)
\end{aligned}
$$

Four models were evaluated to answer research questions two and three (see Appendix F for coding for the analyses in this paper). First, a model (model 1) with no covariates was constructed to determine the percentage of variance that is between agency, within agency (different clinician and youth) and within agency, within clinician (different youth). Then, a model (model 2) with the six strengths covariates added (X_2 $to X_7$) was assessed to determine the contribution to variation in psychiatric symptom severity by the strengths items. Then, the trauma covariate was added (model 3) to assess the contribution to variation in psychiatric symptom severity of the strengths items, adjusting for the lifetime experience of trauma item. Then, the remaining child demographics variables (X_9 to X_{16}) was added (model 4) to assess the contribution to variation in the psychiatric symptom severity of the strengths items, adjusting for trauma and child demographics. Likelihood ratio t-tests were conducted to compare the fit of model 1 vs 2, 2 vs. 3, and 3 vs. 4. The likelihood ratio tests were evaluated at the .05 significance level and test the null hypothesis that the restricted models fit the data as well as the full model against the alternative hypothesis that the full model fits the data better than the restricted model. Then, the estimated intra-class correlations were calculated to further answer research question three.

Multi-level modeling (MLM) is a useful tool for statistical analysis of clustered data when compared to more traditional ordinary least squared regression (OLS) approaches. Most importantly, in a cross-sectional dataset that includes clustered data, the independence assumption, a required assumption for OLS, is violated since items within clusters may be more similar to each other due to similar unmeasured individual characteristics, influenced by environment and captured through clustering, than items between clusters (Rabe-Hesketh & Skrondal, 2012). Thus, in the current dataset, MLM is useful because it allows for strategies to account for clustering instead of just ignoring clustering, which risks calculating standard errors that are inaccurately small and increasing the probably of a type one error (Kwok et al., 2008). (One strategy by which MLM accounts for clustering is that, instead of just using one error for each level one subject as OLS would do, in MLM the error term is divided into separate residual

components, one for each level of clustering). MLM is also useful in accounting for unbalanced clusters, where clusters may be quite different in size. MLM handles this, in models when random effects are predicted, by using an empirical Bayes estimation procedure in which information from each level unit is weighted based on the sample size of the cluster, so the final parameter accounts for different sizes across the clustering levels (Kahn, 2011, pg. 260).

Results

Research question one explored the prevalence of youth strengths variables, the psychiatric presentation composite, and the severity of lifetime traumatic experiences composite for youth entering San Francisco's Children's System of Care (SF-CSOC). Table 2 shows the prevalence of youth strength items (with each item's number and percentage of total sample for each 0-3 score) and the means, standard deviations, and ranges of the psychiatric symptoms and the exposure of lifetime traumatic experience composite.

Table 2

Prevalence of Youth Strengths, Psychiatric Symptom Severity, and Lifetime Trauma Experience Upon Entry to San Francisco's Children's System of Care (n=2049).

	n (% of total sample)				M(SD)	Range
Item	0	1	2	3		
Family/Caregiver Relationship Strength	48 (2)	618 (30)	855 (41)	528 (26)	1.91 (.80)	0-3
Peer and Non-Family/Caregiver Interpersonal Relationships Strength	103 (5)	571 (28)	844 (41)	531 (26)	1.88 (.85)	0-3
School System and Educational Plan Strength	176 (9)	592 (29)	809 (40)	472 (23)	1.77 (.90)	0-3
Extracurricular Activities and Talents Strength	136 (7)	544 (27)	702 (34)	667 (33)	1.93 (.92)	0-3
Spiritual/Religious Beliefs and/or Involvement Strength	539 (26)	195 (10)	475 (23)	840 (41)	1.79 (1.23)	0-3
Relationship Permanence Strength	70 (3)	621 (30)	572 (28)	786 (38)	2.01 (.91)	0-3
Psychiatric Symptom Severity Composite (Youth)	-	-	-	-	6.65 (3.10)	0-21
Psychiatric Symptom Severity Composite (Clinician)	-	-	-	-	6.91 (3.11)	0-21
Psychiatric Symptom Severity Composite (Agency)	-	-	-	-	7.02 (3.45)	19-Feb
Lifetime Trauma Experience Composite (Youth)	-	-	-	-	4.39 (4.08)	0-26

Table 3 reports the means and standard deviations for continuous variables included in the analysis. Because the data for this analysis are clustered at three levels (i=youth, j=clinician, and k=agency), two sets of standard deviations are provided, the first using clinician as the

clustering variable (reporting the overall standard deviation and then how much the response variable psychiatric symptom severity varies between clinicians and within clinicians), and the second using agency as the clustering variable (reporting the overall standard deviation of the variable and then how much the variable varies between agencies and within agencies). Table 2 suggests that at the clinician level, psychiatric symptom severity, lifetime trauma experience, peer, extracurricular activities, spiritual/religious, and relationship permanence vary more between clinicians than within clinicians, while school system supports varies more within clinicians, and family/caregiver varies equally between and within. At the agency level, all the variables vary more within agency than between agency. Since there is variation both between and within agencies and clinicians on all the variables, this suggests that multi-level modeling is appropriate with this dataset in order to account for the variation due to clustering in the final analyses.

Table 3

Mean and Standard Deviations of Continuous Variables at Entry to San Francisco's Children's System of Care (n=2049).

Clinician Level Clustering					
Variable		Mean	SD	Min	Max
Psychiatric Symptom Severity Composite	overall	6.65	3.10	0.00	21.00
	between		2.66	1.00	21.00
	within		2.19	-2.13	18.73
Lifetime Trauma Experience Composite	overall	4.39	4.08	0.00	26.00
	between		3.88	0.00	22.00
	within		2.65	-5.81	24.24
Family/Caregiver Relationship Strength	overall	1.91	0.80	0.00	3.00
	between		0.63	0.00	3.00
	within		0.63	-0.45	3.74
Peer and Non-Family/Caregiver Strength	overall	1.88	0.85	0.00	3.00
	between		0.69	0.00	3.00
	within		0.67	-0.32	3.74
School System and Educational Plan Strength	overall	1.77	0.90	0.00	3.00
	between		0.69	0.00	3.00
	within		0.71	-0.61	3.86
Extracurricular Activities and Talents Strength	overall	1.93	0.92	0.00	3.00
	between		0.75	0.00	3.00
	within		0.67	-0.64	4.36
Spiritual/Religious Beliefs / Involvement Strength	overall	1.79	1.23	0.00	3.00
	between		1.02	0.00	3.00
	within		0.80	-1.02	4.54
Relationship Permanence Strength	overall	2.01	0.91	0.00	3.00
	between		0.77	0.00	3.00
	within		0.63	-0.37	4.01

Agency Level Clustering					
Variable		Mean	SD	Min	Max
Psychiatric Symptom Severity Composite	overall	6.65	3.10	0.00	21.00
	between		2.02	3.00	12.83
	within		2.69	-2.28	18.90
Lifetime Trauma Experience Composite	overall	4.39	4.08	0.00	26.00
	between		2.65	0.00	11.00
	within		3.39	-4.36	27.19
Family/Caregiver Relationship Strength	overall	1.91	0.80	0.00	3.00
	between		0.34	1.17	3.00
	within		0.77	-0.36	3.56
Peer and Non-Family/Caregiver Interpersonal Relationships Strength	overall	1.88	0.85	0.00	3.00
	between		0.43	0.95	3.00
	within		0.81	-0.42	3.93
School System and Educational Plan Strength	overall	1.77	0.90	0.00	3.00
	between		0.45	0.68	3.00
	within		0.86	-0.56	4.09
Extracurricular Activities and Talents Strength	overall	1.93	0.92	0.00	3.00
	between		0.50	0.00	3.00
	within		0.88	-0.62	4.13
Spiritual/Religious Beliefs and/or Involvement Strength	overall	1.79	1.23	0.00	3.00
	between		0.65	0.00	3.00
	within		1.14	-0.87	4.00
Relationship Permanence Strength	overall	2.01	0.91	0.00	3.00
	between		0.55	1.00	3.00
	within		0.82	-0.48	3.89

The residuals of the response variable were checked to make sure that the multi-level modeling assumptions were not violated. Empirical Bayes estimators of the residuals were used instead of maximum likelihood estimators because empirical Bayes procedures use the data from all clusters to estimate the residuals, thus tending to produce more accurate estimates, particularly if cluster size happens to be small (in this data, cluster sizes vary considerably). Figure 1 suggests that the empirical Bayes predictions of the agency-level residuals of the response variable, though generally meeting the assumption of normal distribution, are a bit skewed to the right. Figure 2 suggests that the empirical Bayes predictions of the clinician-level residuals of the response variable, though generally meeting the assumption of normal distribution, are skewed to the right. Figure 3 suggests that there are some outliers, bringing the normality assumption into question. In order to address any potential violation of normality, robust standard errors (which are more conservative, or larger, than non-robust errors and so help

account for potential violations) using the sandwich estimator (Rabe-Hesketh & Skrondal, 2013) were used to run the models.

Figure 1

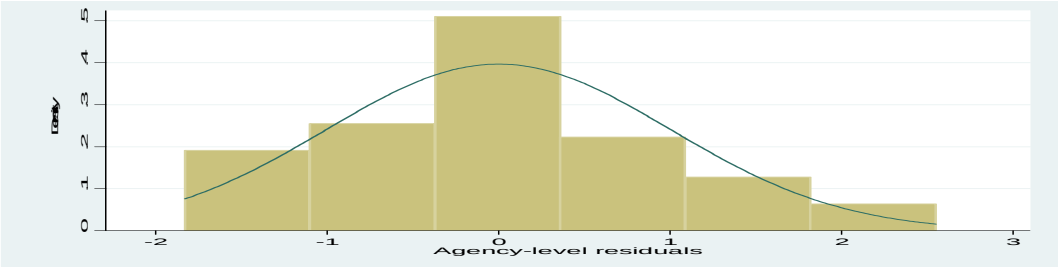


Figure 2

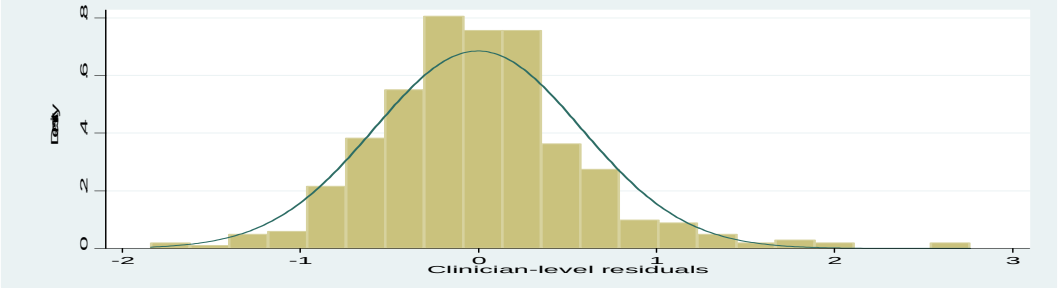
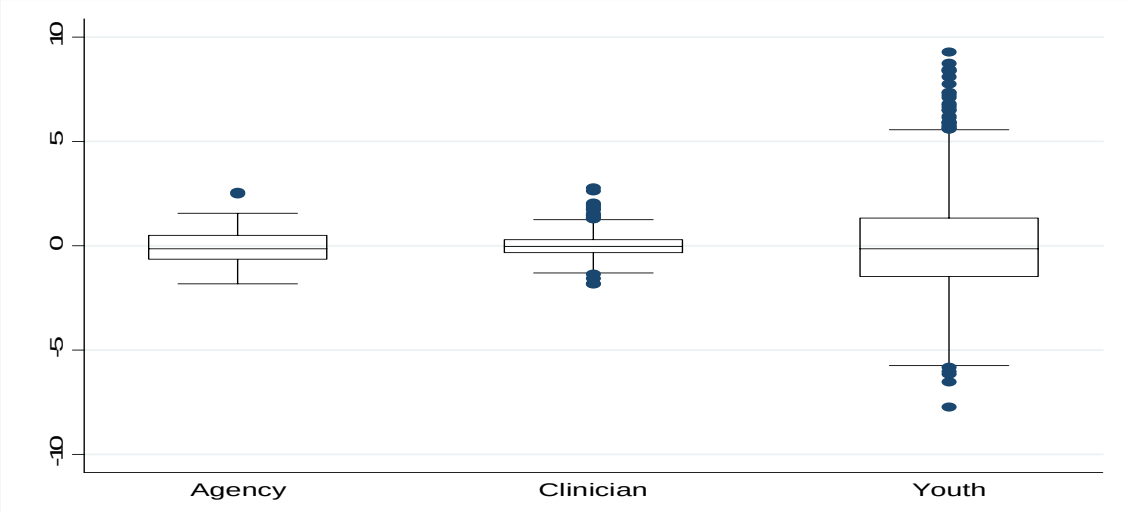


Figure 3
Box Plots of Empirical Bayes Predictions For Random Intercepts at the Agency, Clinician, and Youth Level



Research questions 2 and 3 were addressed by running four random intercept models and calculating intra-class correlations based on the final model. Likelihood ratio tests confirmed that for each model, the between agency variance was significantly different from zero, supporting the use of multi-level modeling estimation techniques in this data (all p-values were <.001 as reported in Table 5). Likelihood ratio tests also confirmed that model 4 was a better fit for the data than the other models (see Table 4).

Table 4
Likelihood Ratio Tests Comparing Nested Models

	$\chi 2$	df	p-value
Model 1 vs. 2	225.7	6	<.001
Model 2 vs. 3	116.74	1	<.001
Model 3 vs. 4	50.13	8	<.001

The analyses reported in model 4 suggested that youth who entered San Francisco's System of Care (SF-CSOC) with a higher score on the family/caregiver, the peer and non-family/caregiver, or the school system/educational plan strengths item were estimated to enter with significantly lower symptoms by approximately half a point (out of a total of 27 points), adjusting for the other covariates. In other words, higher levels of strengths were associated with lower levels of symptoms at entry to SF-CSOC. Males were estimated to enter SF-CSOC with an average psychiatric severity score half a point higher than females (out of 27 total points), adjusting for the other covariates. Older youth tended to enter SF-CSOC with a slightly higher average psychiatric severity score than younger children, but the coefficient was so small (.05) that the interpretation is clinically meaningless. Youth who entered SF-CSOC with higher levels of lifetime exposure to trauma were estimated to enter with lower levels of psychiatric symptom severity. Specifically, as trauma scores upon entry to SF-CSOC increased by a point, youth were estimated to have a significant increase in mean psychiatric symptom severity score of .18 points. Asian youth were estimated to enter SF-CSOC with an average psychiatric severity score that was significantly lower than white youth by about a point (out of 27 points), adjusting for the other covariates. Other ethnic group differences were not found in this analysis. (See Table 5.)

Table 5

Multi-level Models of Psychiatric Symptom Severity Upon Entry into San Francisco's Children's System of Care (n=2049).

	Model 1		Model 2		Model 3		Model 4	
	Beta (Robust SE)	*CI*	*Beta (Robust SE)*	*CI*	*Beta (Robust SE)*	*CI*	*Beta (Robust SE)*	*CI*
Fixed Part								
Intercept	6.86(.29)***	6.29, 7.44	9.84(.43)***	9.00, 10.69	8.43(.40)***	7.65, 9.21	8.28(.42)***	7.45, 9.11
Family/Caregiver Relationship Strength			-.44(.09)***	-.55, -.33	-.33(.06)***	-.45, '-.23	-.34(.06)***	-.45, -.23
Peer and Non-Family/Caregiver Strength			-.44(.09)***	-.61, -.27	-.47(.08)***	-.62, -.30	-.50(.07)***	-.65, -.34
School System and Educational Plan Strength			-.54(.07)***	-.67, -.40	-.55(.07)***	-.68, -.42	-.48(.07)***	-.62, -.34
Extracurricular Activities and Talents Strength			.06(.07)	-.09, .20	.07(.07)	-.07, .20	.05(.07)	-.08, .18
Spiritual/Religious Beliefs/Involvement Strength			-.003(.06)	-.11, .12	.007(.06)	-.10, .11	-.02(.05)	-.13, .09
Relationship Permanence Strength			-.25(.06)***	-.37, -.14	-.05(.06)	-.17, .07	-.04(.06)	-.16, .09
Lifetime Trauma Severity					.19(.02)***	.14, .23	.18(.02)***	.13, .22
Male							.48(.12)***	.25, .71
Age_Centered							.06(.02)**	.02, .09
Ethnicity[a]								
Black							-.07(.27)	-.60, .45
Asian							-.83(.23)***	-1.29, -.36
Latino							.16(.29)	-.41, .73
Multi							.02(.28)	-.52, .57
Pacisl							-.42(.41)	-1.21, .38
Other							.08(.46)	-.82, .99
Random Part								
Between Agencies standard deviation	1.65(.24)***	1.24, 2.21	1.49(.20)***	1.14, 1.95	1.28(.20)***	.95, 1.74	1.19(.20)***	.85, 1.67
Between Clinicians standard deviation	1.23(.11)	1.03, 1.45	1.18(.10)	.10, 1.40	1.06(.10)	.87, 1.29	1.02(.11)	.83, 1.25
Between Clients standard deviation	2.48(.06)	2.36, 2.61	2.34(.05)	2.24, 2.46	2.30(.05)	2.21, 2.40	2.28(.05)	2.19, 2.38
Note								
[a]Reference category = White/Caucasian								
*p<.05 **p < .01. ***p < .001.								

Two estimated intra-class correlations (ICCs) were calculated. The ICC is interpreted as the proportion of the total residual variance shared among units in the same cluster. The level-three intra-class correlation (ICC^3) is estimated at .26. This can be interpreted in two ways: 1) 26% of the residual variance in psychiatric symptom severity (that is not explained by covariates) is due to unmeasured agency variables, or 2) the estimated residual correlation on psychiatric symptom severity between clients of two different clinicians for the same agency is .26. The level-2 intra-class correlation (ICC^2) is estimated at .41. This can be interpreted in two ways: 1) 41% of the residual variation in psychiatric symptom severity for clinicians in the same agency is due to unmeasured clinicians variables, or 2) the estimated residual correlation on psychiatric symptoms between two different youth for the same clinician in the same agency is .41. Because of the use of robust standard errors (which are more conservative, or larger, than non-robust standard errors to account for potential violations of the normality assumptions), they are a bit conservative.

Discussion

This paper explored cross-sectional, administrative data collected by clinicians from youth (aged 5-22) with emotional and behavioral problems as they entered San Francisco's Children's System of Care (SF-CSOC). First, these analyses suggest that youth entered SF-CSOC with many accessible strengths (a strength that could currently be the focus on a strengths-based treatment plan) (see Table 2). Youth entered with an average of nearly four accessible strengths, indicating that even youth with high psychiatric symptom severity may enter with strengths

available for use during the course of an intervention. Approximately twenty percent of youth entered with two or fewer accessible strengths, with about 8% entering with one or zero accessible strengths. In the future, it might be useful to examine a "low strengths" or a "high strengths" client group in order to ascertain if either of these group entered with a different risk profile or behaved differently over a course of intervention, as this might have clinical implications.

Also, youth who entered SF-CSOC with a score that was a point higher on either the family/caregiver strength, peer and non-caregiver strength, or educational system strength item, were estimated to enter SF-CSOC with a lower average psychiatric severity score by approximately half a point, adjusting for the other covariates (see Table 5). In other words, higher levels of these three strengths were associated with lower levels of psychiatric symptom severity upon entry to SF-CSOC. However, the other strength items, including the extra-curricular talents and interest, spiritual/religious, and relationship permanence items, were not significant at the .05 level in this data (see Table 5). It is unclear why some strengths items were associated with lower levels of psychiatric symptom severity while others were not. One possibility is that the extracurricular talents and interests item and the spiritual/religious item were both worded somewhat confusingly in the SF CANS User Manual (2010), and so lack of clarity in scoring of these items might impact the analysis. A more detailed examination of the strengths literature is warranted to contextualize this finding.

This data also suggests that there is more variation in psychiatric symptom severity between clients than between clinicians or agencies, and more variation between agencies than clinicians. This suggests that there are a number of youth characteristics which are related to psychiatric severity but that are not included in these analyses. Future research should include other youth variables as supported by the literature. The fact that there is more variation between agencies than between clinicians in this data may be explained by the fact that some agencies serve youth with higher levels of psychiatric distress than do other agencies. For example, some agencies provide mostly outpatient services, where youth likely have lower psychiatric severity, while some agencies provide mostly inpatient services and serve youth with higher levels of psychiatric severity. However, the fact that there is meaningful variation at all three levels confirms that multi-level modeling is a useful analysis strategy to account for lack of independence due to clustering in the data.

This study is limited by the cross-sectional nature of the data. Only associations at entry to the system can be examined. Since this data suggests that higher levels of three strengths (family/caregiver strength, peer and non-caregiver strength, or educational system strength item) are related to lower levels of psychiatric symptom severity at entry to San Francisco's Children's System of Care (SF-CSOC), further analyses using a longitudinal dataset are needed to explore whether this relationship holds over time. Future analyses, using longitudinal data, should examine whether change in the family/caregiver strength, peer and non-caregiver strength, or educational system strength items are associated with change in psychiatric symptom severity over time. Any findings will be limited, however, by questions of directionality (as discussed in more detail below).

This paper has several additional limitations, many of which are related to the reliance on administrative data (data collected for non-research purposes but which may be used for research). It is not possible to establish directionality in this dissertation because administrative data were used. Time order cannot be established; data were collected through clinician self-reports, and the dataset is missing many variables that might impact both strengths and

psychiatric symptoms such as medication and interventions offered and "taken up." Behavioral theory and related research suggest that increase in strengths decreases symptoms (e.g. Madden, 2013). Future analyses might explore directionality further, for example using a cross-lagged panel model (e.g. Finkel, 1995) or propensity score matching techniques (e.g. Rosenbaum & Rubin, 1983).

Also, the SF-CANS was designed as a clinician self-report tool, subject to reporting bias and documentation error. The SF-CANS does not include youth or parent report to act as a check on the clinician's report. There is no way to track number of contacts, the number of hours the clinician spends on contacting stakeholders before filling out the CANS, and the process the clinician went through to obtain the signature of the primary caregiver. Observational methods of data collection (such as videotaping assessments) have not been used either in validity studies of the CANS or in the collection of data used in this dissertation. The percentage of youth referred for initial assessment, where they were referred from, the percentage who took up services, and future interventions are not reliably tracked. Also, the SF-CANS User Manual (2010) is long and it is possible that clinicians might not have the time or motivation to consult the manual every time they are scoring an assessment, thus resulting in inaccurate scoring.

Also, since administrative data is not collected with research as the primary goal, the reliability and validity of the data (which is of great concern to researchers) may not have been prioritized by the people collecting the data. For example, the spiritual religious item is confusing because the manual does not clarify whether the item can be interpreted as 1) the youth is involved in a religious community or 2) the youth identifies with a spiritual or religious practice and finds strength in this, or both of these. Researchers can only hypothesize as to what the clinician intended when interpreting this item. Future psychometric testing would be helpful here in order to explore the validity of the SF-CANS items. Even though some validity issues have been examined using some versions of the CANS, since every community has its own version, perhaps every community should consider conducting its own validity study in order to understand how the CANS items are interpreted. San Francisco could lead the way in this. Qualitative work might also be helpful here, for example interviews or focus groups with clinicians who use the SF-CANS could shed light on how clinicians understand, use, and interpret the Strengths domain items, the referral process and how many youth took up services, reasons for termination, what interventions are provided, if any, and other factors that are not currently reliably tracked in SF-CSOC.

Also, though the SF-CANS may be considered a strengths-based assessment, using the broad definition suggested, most of the hundreds of possible items measure problems, symptoms, and deficits. There is only one Strengths domain included in a version of the CANS, with eight or ten strength items included at most. Some strengths claimed or found to be related to improved functioning (e.g. self-efficacy, quality of neighborhood or community resources) are not included. The addition of other strengths items would provide a more nuanced picture of the strengths profiles of youth entering SF-CSOC.

PAPER 3

Youth Strengths, Psychiatric Symptom Severity, and Clinician/Client Effects Six and Twelve Months after Initial Assessment in San Francisco's Public System of Care

Sarah Accomazzo

School of Social Welfare
University of California, Berkeley

Introduction

This paper explores strengths and psychiatric symptoms for children and adolescents ("youth") with severe emotional and behavioral problems who are in contact with a public System of Care (SOC) over twelve months. Related research suggests that approximately 10% of youth in the United States will qualify as having "severe" emotional and behavioral problems (resulting in severe impairments or disruptions in one or more life functioning domains or developmental tasks) during any given year (Kessler et al., 2012; Merikangas et al., 2010). (The Diagnostic Statistical Manual (DSM; 2014) is the predominant classification system for emotional and behavioral problems in the United States.) The federal government has faced criticism for its lack of a comprehensive legislative strategy to address severe emotional and behavioral problems among youth as a major public health threat (e.g. Knitzer, 1982; Lourie & Hernandez, 2003). Though service rates have overall increased in the past ten years, it is estimated that approximately half of youth who meet criteria for an emotional or behavioral problem receive inadequate or no mental health intervention at all (Kessler et al., 2012; Merikangas et al., 2010; President's New Freedom Commission, 2003). It has been estimated that 28%-85% of youth drop out "prematurely" (or before the main problem has been reported resolved by youth, caregiver, or clinician) (e.g. Garcia & Weisz, 2002).

Over the past thirty years, the cornerstone of the federal government's attempts to respond to youth with emotional and behavioral problems has been the development of the System of Care (SOC) approach (e.g. CMHI, 2011). Intended as a guiding philosophy and values framework for mental health and substance abuse service provision for youth with severe emotional and behavioral issues, the SOC approach has been implemented in 173 communities throughout the United States and its territories since 1984, with a total of approximately two billion dollars in federal funding. A focus on strengths has been identified as one way to improve service delivery in Systems of Care (e.g. McCammon, 2012). The most recent System of Care (SOC) federal policy brief recommends that SOCs should "provide individualized services in accordance with the unique potentials and needs of each child and family, guided by a strengths-based, wraparound service planning process" (pg. 5). The first Core Value of the System of Care (SOC) Concept (Stroul et al., 2010), states that SOCs should provide: "family-driven and youth-guided care, with the strengths and needs of the child and family determining the types and mix of services and supports provided" (pg. 5). Other federal youth mental health policy documents use the term "strength" (e.g. Report of the Surgeon General's Conference on Children's Mental Health, 2000). Variations on the "strengths-based perspective in social work" (e.g. Saleebey, 2009) are taught in many social work education programs.

One underlying implication of a "focus on strengths" in a System of Care is that this will improve outcomes for youth. Behavioral theory, and related research in experimental and applied settings provide both theoretical and empirical support for the contribution of client strengths (e.g. alternative positive repertoires) in reducing emotional and behavioral problems (e.g. Fischer et al., 2011; Goldiamond, 1974; Madden, 2013; Staats, 2012). However, in studies using Systems of Care samples, strengths have received less attention than problems and needs, as evidenced by the fact that many do not report on strengths at all (e.g. Bickman, 1999; Snyder et al., 2012; Trask & Garland, 2012).

The Effectiveness of the System of Care Approach in Reducing Youth Emotional and Behavioral Problems.

Since a meta-analysis and/or systematic review does not currently exist for this body of literature (to the knowledge of this author), the following section reviews several studies that used a System of Care sample and explored the role of strengths to some extent, but is not comprehensive. This review only includes studies that identify the sample as a "System of Care." Wraparound literature was not included (as Wraparound interventions are not exclusively implemented in System of Care settings). The challenges in conducting research in Systems of Care, and specifically in examining the role of strengths in Systems of Care are discussed.

The effectiveness of the Systems of Care (SOCs) approach in reducing mental health problems compared with other mental health service approaches is unclear, based on the studies reviewed in this paper. An early comparison study (a study that makes at least some attempt to include a comparison group) of a System of Care approach (n=984) compared three military areas in Fort Bragg, North Carolina. One of these had replaced their existing children's health care framework with a System of Care approach (Bickman, 1996). The System of Care group reported increased use of services, less restrictive levels of care, and higher client satisfaction with services. However, the SOC and comparison groups reported the same rate of overall improvement on psychiatric symptoms, and the SOC reported higher costs than the non-SOC sites. In a five year follow up study, consistent with previous findings, no differences on overall improvement were seen between the groups (e.g. Bickman et al., 2000). The Fort Bragg study was critiqued for lack of random assignment and failure to document type and effectiveness of services provided during the course of the study (Bickman, 1996). A follow-up study using random assignment of children and families (n=350) to either a pre-existing System of Care (SOC) or to a control group who sought services on their own was conducted (Bickman, 1999). This study found that youth and families received more case management and home visits in the SOC, but replicated the finding that there were no significant differences between the two groups in psychiatric symptom reduction over time (Bickman, 1999). Alternatives for evaluating System of Care effectiveness have been proposed, including a focus on 1) the type and quality of services (e.g. Salzer & Bickman, 1999), 2) the implementation of, and fidelity to, SOC principles and values (e.g. Stephens et al., 2004), 3) understanding and integrating common elements (e.g. Garland, Bickman, & Chorpita, 2010) and common factors (e.g. Wampold, 2010), and 4) compiling evidence of effectiveness through site-specific evaluations on a range of outcomes (e.g. Foster et al., 2007).

Existing site-specific studies (exploring outcomes within one SOC instead of comparing SOCs) have generally reported an improvement in emotional and behavioral problems for youth in contact over time with a System of Care (SOC), though these are often based on pre-posttest designs without comparison groups. (None of the studies reviewed here were randomized controlled trials.) For example, using a pre-post design, the most recent national evaluation of sixty-eight federally funded SOC communities reported that among youth who entered a System of Care from 2005-2006 and were tracked over 36 months (n=3,004), 27% experienced a clinically significant reduction in their overall level of behavioral and emotional problems (as reported by caregivers on the Child Behavioral Checklist, Achenbach & Rescorla, 2000; 2001) within six months after beginning intervention in a System of Care (SOC) (CMHI, 2011). Based on an initial assessment and a follow up assessment every six months (without comparison groups), nearly one-third of youth reported a clinically significant decrease in emotional and behavioral problems by twelve months, nearly a half reported a reduction by 24 months, and

40% had a clinically significant reduction in psychiatric symptoms by 30 months (pg. 12). The percentage captured in follow-up and the percentage who received interventions is not reported. Manteuffel et al. (2002) found that, in youth who entered Systems of Care in 1995 and 1996, nearly half of young people reported clinically significant improvement by 24 months after entry to the SOC (n= 2,580). Anderson et al. (2008) found that youth impairment and functioning scores, as reported by clinicians, moved from a clinical level to a sub-clinical level approximately 36 months after entry to the SOC (n=351). Sieracki et al. (2008) found that youth improved one point (out of a total of 27 points) over ten months of SOC involvement (n=563).

Complications and Limitations of Evaluating the Effectiveness of the System of Care (SOC) Approach. There are several complications in evaluating the effectiveness of a System of Care (SOC) approach. It is challenging to count or estimate the total number of studies that evaluate a SOC approach. Studies tend to be inconsistent in their terminology, making it hard to know if the samples are comparable. One study might report on a sample of youth in a System of Care, while another might report on a sample from a "public health system serving youth" or a "public behavioral health system." Some studies might focus on specific subsets of youth served in a System of Care (SOC), including youth with mental health issues in a child welfare system, or youth who receive wrap-around services in a SOC (e.g. Painter, 2012). For example, there is a large body of literature evaluating the effectiveness of the Wraparound intervention (Bruns & Walker, 2008), an intervention which was developed specifically for youth in a SOC. The wraparound intervention emphasizes a focus on strengths, and the wraparound website (http://www.nwi.pdx.edu/) and manual (http://nwi.pdx.edu/NWI-book/index.shtml) provide resources for using strengths in clinical work with youth in SOC. However, the Wraparound intervention is not necessarily delivered in every SOC, and not every youth receiving services in a Systems of Care with a Wraparound program might be involved in the Wraparound program.

Since SOCs are context-specific entities, they may or may not share critical characteristics for comparison. For example, SOCs may differ on their implementation strategies and fidelity to SOC values and principles (e.g. Stephens et al., 2004) or on the interventions in use (CMHI, 2011). These studies may use different measures, and thus may operationalize "strengths" or "mental health outcomes" differently. Studies may also use different analysis strategies, different outcomes, and different sample sizes. Some studies may use composite scores versus individual items, or some may use valid scales that measure a specific strength (e.g. a self-esteem scale or a family competence scale).

The site-specific studies reviewed in this paper all used a pre-posttest design without a comparison group, making it impossible to know if the reported gains were related to contact with the System of Care or would have occurred naturally over the course of time. Also, the reported gains are generally modest, as it took multiple months or even years (e.g. CMHI, 2011) to achieve meaningful clinical change. Any conclusions are limited by issues related to the nature of administrative data (data collected for non-research purposes but may be used for research). For example, it is likely that regression to the mean, while a common phenomenon in health service research (e.g. Lambert & Bickman, 2004), is at least partially responsible for the symptom reduction seen in these studies.

Strengths of Youth with Emotional and Behavioral Problems in Systems of Care (SOCs).

Overall, the studies reviewed here suggest that youth enter Systems of Care (SOCs) with at least some strengths. Since 2002, the National Evaluation of Systems of Care, conducted by the Children's Mental Health Initiative (CMHI), has used the Behavioral and Emotional Rating

Scale Parent and Youth scales (BERS-2-P; BERS-2-Y; Buckley & Epstein, 2004) to evaluate youth emotional and behavioral strengths. (The current study uses the Child and Adolescent Needs and Strengths, CANS, Lyons, 2009), described in detail in the Methods section of this paper.) Both the BERS-2 Parent and Youth scales are designed to measure five domains of social-emotional, individual-level strengths, including interpersonal strengths, family involvement, intrapersonal strengths, school functioning, and affective strengths. Each domain's score can be reported individually or compiled into a BERS Strengths Index Score, a composite score of the five domains (Buckley & Epstein, 2004). The five subscales each range from 0-20 and have a mean standard score of 10 (standard deviation: 3). The Strength Index ranges from 0-200 and has a mean of 100 (standard deviation: 15), including below 70 (very poor strengths); 70 to 79 (poor strengths); 80 to 89 (below average strengths); 90 to 110 (average strengths); 111 to 120 (above-average strengths); 121 to 130 (superior strengths); above 130 (very superior strengths). All scales are written at a fifth-grade reading level and take approximately ten minutes to complete.

The BERS was normed on a nationally representative sample and has been claimed to demonstrate acceptable reliability and validity (e.g. Buckley & Epstein, 2004). For example, internal consistency had coefficients ranging from .95-.97 (Buckley & Epstein, 2004, pg. 24). Two studies reported that the BERS-2 scales correlated negatively with a self-report scale measuring youth emotional and behavioral problems (Epstein, 2004; Reid et al., 2000). Reid et al. (2000) reported that the BERS-2 significantly differentiated between youth with and without "emotional disturbance" on all the BERS subscales and the Strengths Index Score (n=418).

Youth Strengths Upon Entry to a System of Care. CMHI's 2011 Annual Report reported on youth entry characteristics for 2008-2009, using the BERS-2 (Appendix G, pgs. 38-39). Between 2008-2009, caregivers (n=566) reported that youth entered with an average Strengths Index Score of 78.1 (SD: 16.8), twelve points below the national average of 90 points (out of 200 total points). Interestingly, youth (n=414) self-reported that they entered with an average Strengths Index Score of 94.4 (SD: 17.5), above the national average (though the standard deviation is large). The fact that youth self-reported higher levels of strengths than caregivers at entry to a System of Care (SOC) suggests that strengths may be a useful focal point for youth in SOCs.

Other studies have also reported that youth enter Systems of Care with some strengths, though each study defines "strengths" differently. For example, Radigan et al. (2013) described prevalence of strengths using the Child and Adolescent Needs and Strengths (CANS) tool (Lyons, 2009), reporting that for youth entering a System of Care (n=799), 61% entered with a family strength they could use in a strengths-based plan, 54% with relationship permanence strength, 59% with education strength, 48% with spiritual/religious strength, 45% with interpersonal peer relationships, and 49% with talents/interests strength. Whitson et al. (2010), using the BERS-2 Parent Scale (Buckley & Epstein, 2004), reported that, for youth who entered a SOC (n=194), caregivers reported a below average strengths score (85.63 out of 200 possible points on the Strengths Index Score), similar to that reported in the national 2011 study. Anderson (2008), using the BERS-2 Parent Scale (Buckley & Epstein, 2006), reported that youth entered a System of Care (n=351) with an average strengths index score of 89.09 (out of 200 possible points). (For a discussion of the reported validity of the BERS-2 scales, see above).

Do strengths increase while youth are in contact with a System of Care? The studies reviewed here reported that youth generally experience a self-reported increase in strengths while in contact with a System of Care (SOC), though gains are modest. For example, the CMHI

(2011) Annual Report explored change in strengths for youth who received an initial assessment in a System of Care community between 2008-2009 (pgs. 18-19). Both youth and parents reported an increase in strengths for youth twelve months after initial assessment, though youth reported smaller increases than caregivers. Caregivers reported that youth experienced an average growth on the Strengths Index Score of 2.9 points (SD: 17.5) (out of 200 points) between initial and six month assessment (n=198), and an average growth of 4.7 points (SD: 16.5) (out of 200 points) between six month and twelve month assessment (n=56). Youth self-reported an average growth on the Strengths Index Score of 3.1 points (SD: 17.5) between initial assessment and six month assessment (n=130), but did not report any growth between six month and twelve month assessment (n=33). Since the Annual Report (2011) does not provide information on why this difference might have occurred, future work should attempt to contextualize and understand the difference between youth and caregiver report of strengths. Also, the large standard deviations suggest there is variation in the data, and further exploration of extreme groups might be useful.

Other studies have also reported a modest growth in strengths over time. Whitson et al (2010) reported that the Overall Strengths Index Score (measured by the BERS-2 Parent scale, Buckley & Epstein, 2004) increased by 1.83 ($p < .01$) points (out of a total of 200 possible points) every 6 months over the course of the study (n=194). Since youth in this study reported an initial Strengths Index Score of 89.9 points, an increase of 1.83 points suggests that youth crossed from below average (under 90 points) to average strengths (a score of 90-110 points) at around 18 months. Anderson (2008) (n=351) reported that youth crossed from below average to average on the BERS-2 Strengths Index Score by 12-18 months after beginning intervention in a System of Care. (Since these changes are so small, the clinical significance is questionable.)

Is level of initial strengths associated with rate of change (improvement in mental health outcomes) over time? Overall, the studies reviewed here suggest that youth who entered a System of Care (SOC) with higher levels of strengths tended to experience more improvement in mental health outcomes over time, though the gains are modest and many studies do not evaluate the relationship between strengths and problems over time. Sieracki et al. (2008), using the Child and Adolescent Strengths and Needs tool (Lyons, 2009), reported that rate of improvement in mental health problems was associated with a one point increase in youth strengths (out of a total possible eighteen points), reported at initial assessment (beta=.003, SE: .001, p<.05) (pg. 803). The National Evaluation of Systems of Care (CMHI, 2011) data does not report on this.

Is change in strengths associated with change in mental health outcomes over time? This author found one study that examined this question in a System of Care sample. Dunleavey et al. (2011) reported on youth (n=77) who entered a System of Care over a five year period and were rated as having high antisocial personality traits (a score of 2 or 3 on a 0-3 likert scale) on a version of the Child and Adolescent Needs and Strengths (CANS; Lyons, 2009). The study explored clinician self-reported change in outcomes between initial assessment and final assessment (but the amount of time between these two assessments varied for each youth and was unreported in this study). A successful outcome was defined as the "resolution of antisocial behavior" (a binary dummy variable for anti-social behavior was created with 1 = acute and 0 = non-acute), measured by the antisocial behavior item moving from a 1 to a 0 between initial and final assessment. Thirty-nine out of 77 youth (50.6%) reported resolution (anti-social behavior score moved from a 1 at initial assessment to a 0 at final assessment). Youth who gained on the wellbeing (p<.001) or family functioning strength item (p<.001) between initial assessment and final assessment in a System of Care were more likely to report resolution. A limitation of this

analysis is that a pre-post test design was used (resulting in threats to internal validity), clustering was not accounted for, and time order cannot be established. Also, the well-being item has a confusing interpretation as it is perhaps more representative of a psychological factor, not a symptom.

Other covariates reported as associated with mental health outcomes over time in the System of Care studies.

Other variables have been reported as associated with mental health symptom change over time. The national evaluation of Systems of Care (2011) suggests that youth generally enter SOCs with some exposure to potentially traumatic events. For youth who entered a SOC from 2008-2009, 51.4% reported that they had experienced at least one potentially traumatic event in the past six months. Twenty-one percent reported exposure to physical abuse, 14.3% reported exposure to sexual abuse, and 9.3% had attempted suicide (pg. 16). A cumulative effect of exposure on potentially traumatic experiences and mental health outcomes has been reported in youth in SOCs (e.g. Drury et al., 2012; Felliti et al., 1998; Whitson et al., 2010), with higher levels of reported exposure to trauma associated with higher severity of mental health symptoms.

Gender and Strengths. In the studies reviewed here, males are often overrepresented in Systems of Care (SOCs). For example, the national evaluation of SOCs (CMHI, 2011) reported that for youth entering SOCs in 2008-2009 (n=14,431), approximately 2/3 were male and 1/3 were female (pg. 10), and other studies mirror these estimates (e.g. Whitson et al., 2010). Gender differences have been reported in prevalence of mental health symptoms and trauma experiences for youth (e.g. Costello et al.). However, other System of Care studies have not reported gender differences in severity of emotional and behavioral issues at entry or over time (e.g. Dunleavy et al., 2011; Anderson et al., 2008). Gender differences bear further examination as these studies include gender in the analysis but do not discuss related literature on expected gender differences in detail, if at all.

Ethnicity and Strengths. A recent nationally representative, population-level study found that African American youth and Latino youth reported significantly lower levels of severe emotional and behavioral issues when compared to white youth (n=6483) (Kessler et al., 2012). The System of Care studies reviewed often did not report ethnic group differences. One study (n=305) found that African American males reported gaining strengths at a slower rate than other youth (however, they also entered at a higher average strength score than other youth), and white females entered with lower levels of strengths than other youth (Anderson et al., 2008), though the meaning of these findings is unclear. More detailed explorations into ethnic differences on symptom improvement and the role of strengths are warranted, given that the studies reviewed here include ethnicity variables but do not examine them in detail or provide a strong theoretical context for ethnic group differences.

Age and Strengths. In the studies examined here, reported results on age group difference are mixed. A recent, nationally representative, population-level study (e.g. Kessler, 2012) found that as youth got older, they were more likely to report higher levels of emotional and behavioral problems. On the other hand, one study described that younger youth were reported as having more emotional and behavioral problems than older youth (Anderson et al., 2008). In this pre-post study, being younger at initial assessment in a System of Care sample (n=305) was associated with greater functional improvement over a course of intervention by .47 points (out of 50 points) for every year of age ($p<.05$) (e.g. Anderson et al., 2008), though given the large range of 50 points, a .47 change is clinically meaningless. Another study found that age was not

a predictor (e.g. Whitson et al., 2010). A more detailed exploration of the effects of age on rates of change of emotional and behavioral problems are needed.

Clinician, Youth, and Time Variables and Strengths. Common factors (e.g. therapeutic alliance, therapist attention, empathy, and positive regard), have been shown to contribute heartily to positive outcomes (e.g. Duncan et al., 2010). Wampold (2001, 2010) demonstrated that clinician variables make a contribution to intervention success, regardless of the type of intervention delivered. Client characteristics are also related to outcomes (e.g. Bohart & Tallman, 2010). Few studies examine these factors in Systems of Care. More research using SOC samples is needed to explore the contribution of time, clinician, and clients to variation in outcomes.

The Current Study.

The current study attempts to address concerns described in the prior sections through the following research questions:

1) How much change in psychiatric symptom severity, if any, was reported for the average youth at six and twelve month assessment in San Francisco's Children's System of Care, adjusting for initial strengths intensity, initial psychiatric severity, lifetime trauma experience, and youth demographics (gender, age, ethnicity)?

2) To what extent is the change in psychiatric symptom severity, if any, associated with client level factors upon entry to San Francisco's Children's System of Care, including initial youth strengths intensity, initial psychiatric severity, lifetime trauma experiences, or youth demographics (gender, age, and ethnicity)?

3) What is the contribution to variation in psychiatric symptom severity change over twelve months of treatment by the clinician providing the services, if any?

4) To what extent is change in strengths at twelve month assessment in San Francisco's System of Care associated with change in psychiatric symptom severity at twelve month assessment, if any, adjusting for lifetime trauma experience and youth demographics (gender, age, and ethnicity)?

Method

Study Design and Measures.

In early 2008, San Francisco's Children's System of Care (SF-CSOC) adopted the Child and Adolescent Needs and Strengths (SF-CANS) instrument as its primary assessment tool for all youth programs. Consistent with the tool's basis in Communimetrics theory (Lyons, 2009), which emphasizes item-level reliability and validity, different stakeholder groups have used different configurations of strength items in their own settings, and have added to the bank of strengths items used as part of the CANS. SF-CSOC went through a group process, involving consultation with subject experts and family members, in order to adopt the items for the SF-CANS.

The SF-CANS includes fifteen domains with more than 100 items. Fourteen of the domains are designed to identify type and acuity of clinical and psychosocial problems (e.g. psychiatric symptom severity, risk behaviors) and one domain is designed to identify youth strengths (SF CANS 5-18 Scoring Manual, 2010). For each item, assessors assign a rating using a four-point likert scale. The fourteen problem domains are designed to assess severity of problems (instead of counting frequency of problems), while the Strengths domain is designed to track the presence of strengths that could be used in a strengths-based plan. After a case is opened, clinicians have thirty days to submit a CANS initial assessment online. During these

thirty days, clinicians are expected to identify and contact multiple stakeholders in the youth's life (e.g. parents, caretakers, other family, teachers, caseworkers, etc.) to gather as much information as possible. The number of contacts within thirty days is not tracked in SF-CSOC, but could vary from one contact to ten or more (as estimated by a senior SF-CSOC administrator). Though clinicians are encouraged to talk to as many stakeholders as possible in the youth's life, there is no way to know how many stakeholders a clinician spoke with before filling out the SF-CANS online and whether the contacts were in person or by telephone. Clinicians are encouraged to establish a collaborative partnership while filling out the SF-CANS with the youth and family, and the SF-CANS includes a place for a caregiver signature (number completed is not tracked). There is no reliable way to know if clinicians are working collaboratively or whether or not the caregiver signed, let alone whether or not the caregiver felt coerced to sign, secretly felt that the clinician was condescending or rude, or had questions that were still unanswered.

An initial assessment is estimated to take around two hours to enter into the online database, but total amounts of time may vary and are untracked in this system. Re-assessments must be completed at least every six months while the youth's case is marked "open" in the San Francisco Children's System of Care (SF-CSOC) database and when the youth terminates contact with SF-CSOC (actual number of assessments completed and number of contacts prior to re-assessment is unknown). Reasons for termination are not reliably tracked, but could include dropping out, moving out of area, or completing a course of therapeutic intervention successfully, partially successfully, or unsuccessfully. The percentage of youth referred for services, where they were referred from, the percentage who went to services, and the type, quantity, and quality of interventions (if any) are not reliably tracked. Since August 1st, 2010, CANS data have been stored in an online, HIPAA-compliant database. De-identified data were extracted for this study by a data manager and stored by the researcher according to procedures approved by the institutional review company, Ethical and Independent Review Services.

Reliability and Validity of the CANS. The CANS performs reasonably well on some traditional psychometric standards. Versions of the CANS have been shown to be reliable at the item level, for example, one study found an intra-class correlation of .85 between researchers and clinicians in sixty randomly drawn case files. (Anderson et al., 2003). Acceptable item-total correlations were demonstrated in a study with a sample of 6010 youth in a public system, for the strengths domain (Cronbach's alpha = .70) and the psychiatric symptom severity domain (Cronbach's alpha = .71) (Lyons, 2009). Also, CANS scores were shown to be correlated (r=.63) with another self-report measure of children's emotional and behavioral problems, the Child and Adolescent Functional Assessment Scale (e.g. Lyons et al., 2004). A version of the CANS accurately distinguished youth (n=150) who were placed in inpatient care and youth placed in community-based outpatient care youth (Anderson et al., 2001).

Sample.

The sampling frame for this study included youth (n=2402) who received an initial CANS assessment between June 2011 and August 2013 and whose assessing clinician entered the initial assessment into an online database within the timeline required by the city, within 45 days from first of point contact. (Youth who entered into the system but whose clinician did not meet the 0-45 day deadline were excluded from the sample due to worry that the assessment was no longer capturing the youth's presentation upon entering the system). Missing-data patterns were examined. Out of the 2,402 unique youth, only 8% had complete data (185 youth were

recorded as having an initial, six month, and twelve month assessment), 23% (561 youth) were recorded as having initial and six month assessments, 63% (1,513 youth) were recorded as having only an initial assessment, and 6% (143 youth) were recorded as having an initial and a twelve month assessment. Between 0 and 11 youth per clinician were reported as having all three assessments (complete data). The number of youth per clinician ranged from 1 to 66. Though the standard deviations and ranges are large for these clusters of clinicians and youth, multi-level modeling can adequately handle clusters of different size (as discussed in the Statistical Analysis section. The final sample (n=2059) included all youth who were not missing descriptive statistics. (Paper two included an agency variable, which resulted in an extra ten missing youth and thus a sample of 2049; slight variation in descriptive statistics between papers three and two are a result of this.)

Youth were referred for an initial assessment from a variety of sources that are currently untracked in this public system, but which include schools, family resource centers, community centers, and juvenile justice programs. They were typically referred due to alleged emotional, behavioral, and environmental issues that were said to impair functioning in their family, school, or community. Socio-economic status of children and families is not currently tracked in this SOC because 98% of client services are paid for by billing Medicaid.

Out of the 2059 youth whose clinician recorded an initial assessment, 650 were reported to receive a six month assessment and 258 were reported to receive a 12 month assessment. Reasons why a youth may not have received a six or twelve month assessment include: 1) the initial assessment was conducted towards the end of the two year period in which data for this study was collected, or 2) youth terminated contact with the System of Care before the next assessment period (for a multitude of reasons that are currently untracked, including premature drop-out or successful attainment of goals). Though some information on why youth might not have received a six month or twelve month assessment is tracked by clinicians in the SOC, this data was inconsistent, incomplete, and complicated to interpret, and thus was left out of this analysis. (This study uses multi-level modeling techniques to account for unbalanced clusters and missing data, as discussed in the Statistical Analysis section below.)

Youth in this sample (n=2409) had an average age of 12.5 (SD: 3.85). Table 1 shows the other demographics of the sample youth. 468 clinicians were included in this sample. On average, there were 4.5 youth per clinician (SD: 6.34, Range: 1-58).

Table 1

Youth Demographics Upon Entry to the Children's System of Care

Characteristic	n	%
Gender		
Male	1264	61
Female	795	39
Ethnicity		
African-American/Black	687	33
Asian	450	22
Latino	573	28
Multi-Ethnic	72	4
Pacific Islander/Native Hawaiian	45	2
White	164	8
Other	68	3

Statistical Analysis.

For a description of the variables used in these analyses, see Appendix G. The statistical models for questions 1-4 are listed below (descriptions of the interpretations of each coefficient are included in Appendix H).

Model for questions 1-3:

$$
\begin{aligned}
Psych_{ijk} = {} & \beta_0 + \beta_1 Assess6dum_i + \beta_2 Assess12dum_i + \beta_3 famnew_init_{jk} \\
& + B_4 interpnew_init_{jk} + B_5 educnew_init_{jk} + \beta_6 talnew_init_{jk} \\
& + \beta_7 spirnew_{init\,jk} + \beta_8 relnew_init_{jk} + B_9 trauma_init_{jk} + B_{10} Male_{jk} \\
& + \beta_{11} Age_cent_{jk} + \beta_{12} Black_{ijk} + \beta_{13} Asian_{ijk} + \beta_{14} Latino_{ijk} \\
& + \beta_{15} Multi_{ijk} + \beta_{16} Pacisl_{ijk} + \beta_{17} Other_{ijk} \\
& + \beta_{18} famnewinitXAssess6dum_i + \beta_{19} famnewinitXAssess12dum_i \\
& + \beta_{20} interpnewinitXAssess6dum_i \\
& + \beta_{21} interpnewinitXAssess12dum_i + \beta_{22} educnewinitXAssess6dum_i \\
& + \beta_{23} educnewinitXAssess12dum_i + \beta_{24} traumainitXAssess6dum_i \\
& + \beta_{25} traumainitXAssess12dum_i + \beta_{26} maleinitXAssess6dum_i \\
& + \beta_{27} maleinitXAssess12dum_i + \beta_{28} agecentXAssess6dum_i \\
& + \beta_{29} agecentXAssess12dum_i + \beta_{30} asianXAssess6dum_i \\
& + \beta_{31} asianXAssess12dum_i + \zeta_{jk}(2) + \zeta_k(3) \\
& + \epsilon_{ijk}, \quad \zeta_{jk}^{(2)} \sim N\left(0, \psi^{(2)}\right), \quad \zeta_k^{(3)} \sim N\left(0, \psi^{(3)}\right), \quad \epsilon_{ijk} | \zeta_j \sim N(0, \theta)
\end{aligned}
$$

Three models were used to explore research questions 1-3 (see Appendix I for the coding used to run these models in Stata13). First, a model (model 1) with no covariates was constructed to determine the percentage of variance that is a) between clinician, b) within clinician (different youth and assessment) and c) within clinician, within youth (different assessment). Next, to answer research question 1, the covariates $\beta_{1\,to}\ \beta_{17}$ were added (model 2) to assess the contribution to the variation in mean psychiatric severity score over twelve months by these

covariates. Interaction terms were constructed for any of the significant covariates from model 2. Then, to answer research question 2, $\beta_{18}\ to\ \beta_{31}$ (the interaction terms) were added to the model (model 3) to see if mean psychiatric severity varied over six months or twelve months for a given level of the covariate, adjusting for the other covariates in the model. Because it appeared that the residuals of the dependent variable were not perfectly normally distributed (see histograms and boxplots below), "robust standard errors using the sandwich estimator method" (Rabe-Hesketh & Skrondal, 2013) were calculated for all the models to address potential violations of the normality assumption. (As long as the sample is not small, these robust standard errors are considered reliable because they are more conservative than non-robust standard errors, even if the normality assumption might be somewhat violated). A likelihood ratio test was conducted to compare the fit of model 1 vs 2 and 2 vs. 3. The likelihood ratio test was evaluated at the .05 significance level and tests the null hypothesis that the restricted model fits the data as well as the full model against the alternative hypothesis that the full model fits the data better than the restricted model. Then, the estimated intra-class correlations were calculated to further answer research question three.

Model for Research Question Four:

$$
\begin{aligned}
Psych_{ijk} = \ & \beta_0 \\
& + \beta_1 famnew_{ijk} + B_2 interpnew_{ijk} + B_3 educnew_{ijk} + \beta_4\, talnew_{ijk} \\
& + \beta_5 spirnew_{ijk} + \beta_6 relnewi + \beta_7 Assess6dum_i + \beta_8 Assess12dum_i \\
& + \beta_9 mnfamnew_{jk} + B_{10} mninterpnew_{jk} + B_{11} mneducnew_{jk} \\
& + \beta_{12}\, mntalnew_{jk} + \beta_{13} mnspirnew_{jk} + \beta_{14} mnrelnew_{jk} \\
& + B_{15} trauma_init_{jk} + B_{16} Male_{jk} + \beta_{17}\, Age_cent_{jk} + \beta_{18} Black_{ijk} \\
& + \beta_{19} Asian_{ijk} + \beta_{20} Latino_{ijk} + \beta_{21} Multi_{ijk} + \beta_{22} Pacisl_{ijk} \\
& + \beta_{23} Other_{ijk} + \zeta_{jk}^{(2)} + \zeta_k^{(3)} \\
& + \epsilon_{ijk}, \quad \zeta_{jk}^{(2)} \sim N\left(0, \psi^{(2)}\right), \quad \zeta_k^{(3)} \sim N\left(0, \psi^{(3)}\right), \qquad \epsilon_{ijk} | \zeta_j \sim N(0, \theta)
\end{aligned}
$$

To answer research question four, model 4 was constructed (see Appendix I for the coding used to run these models in Stata13). $\beta_1\ to\ \beta_6$ are time-varying youth strength covariates, interpreted as the between-clinician effects since $\beta_9\ to\ \beta_{14}$ (mean-centered strength covariates, representing the within-clinician effects) were included to isolate the within-clinician effects. With the inclusion of $\beta_9\ to\ \beta_{14}$, $\beta_1\ to\ \beta_6$ answer the question of whether change in youth strength covariates over twelve months of intervention predict change in youth psychiatric symptoms over twelve months of intervention, for youth with similar characteristics working with different clinicians. Again, for model 4, robust standard errors were calculated since the residual distribution of the response variable is not perfectly normal (these standard errors are more conservative, or larger, than non-robust errors and so help account for potential violations of the normality assumption).

In addition to dealing with clustering in data (as discussed in paper two), multi-level modeling is useful for handling missing data and for dealing with repeated measures. First, MLM provides more flexibility in handling missing data and unbalanced designs by using information from all of the data, even if the data is not all complete (in part due to the use of maximum likelihood estimators, as discussed in paper two). Additionally, multi-level modeling has several

advantages in dealing with longitudinal data when compared to ANOVA repeated measures designs (Kwok, et al., 2008; Rabe-Hesketh & Skrondal, 2012, pg. 244). If fixed effects approaches are used, each individual is able to serve as his/her own control over time, thus accounting for infinite unmeasured confounders (though this is not the case with random effects approaches). Ordinary Least Squares (OLS) approaches treat every time point as repeated cross-sectional data (where the subjects are drawn independently at each time point), and this does not differentiate between within and between-subject comparisons (and within subject comparisons are ideal because subjects truly act as their own controls). OLS approaches are also incorrect when there is missing data. MLM allows for special models that can differentiate between within and between subjects and model within subject residual covariance in more effective ways (Rabe-Hesketh & Skrondal, 2012, pg. 244).

Table Two reports the means and standard deviations for the time-varying and non-time varying variables included in the analysis. Because the data for this analysis include three levels of clustering (i=assessment, j=youth, and k=clinician), two sets of means and standard deviations are provided, the first part of the table reports clustering for the time-varying variables using youth as the clustering variable (reporting the overall mean over twelve months and then how much the response variable Psych varies between youth and within youth averaged over twelve months). There is more variation between youth than within youth for all the time-varying variables. The second part of the table reports clustering for the non-time varying, level two variables, using clinician as the clustering variable (reporting the overall mean over twelve months and then how much the response variable Psych varies per clinician over twelve months).

Also, seven mean variables (e.g. mean_psych, mean_famnew, etc.) were created for the time-varying variables and included in the second part of the table. The "between" deviation reported on these variables can be interpreted as the average amount of variation of the response variable psychiatric symptom severity between clinicians over twelve months. According to these numbers, there is between-clinician variability in psychiatric severity score. Since there is variation both between and within clinicians and youth on all the variables, this suggests that multi-level modeling is appropriate with this dataset in order to account for the variation due to clustering in the final analyses.

Table 2

Mean and Standard Deviations of Time-Varying and Non-Time Varying Variables

Variation in Time-Varying Level One Variables Within and Between Youth (ignoring clinician)					
Variable		Mean	SD	Min	Max
Psychiatric Symptom Severity Composite	overall	6.41	3.08	0.00	22.00
	Between		3.01	0.00	20.00
	Within		1.06	-2.09	14.91
Family/Caregiver Relationship Strength	overall	1.92	0.80	0.00	3.00
	between		0.77	0.00	3.00
	within		0.31	0.42	3.42
Peer and Non-Family/Caregiver Strength	overall	1.90	0.84	0.00	3.00
	between		0.80	0.00	3.00
	within		0.31	0.23	3.40

School System and Educational Plan Strength	overall	1.80	0.89	0.00	3.00
	between		0.87	0.00	3.00
	within		0.30	0.30	3.30
Extracurricular Activities and Talents Strength	overall	1.92	0.91	0.00	3.00
	between		0.88	0.00	3.00
	within		0.33	0.42	3.59
Spiritual/Religious Beliefs / Involvement Strength	overall	1.82	1.21	0.00	3.00
	between		1.18	0.00	3.00
	within		0.42	-0.18	3.82
Relationship Permanence Strength	overall	2.04	0.90	0.00	3.00
	between		0.87	0.00	3.00
	within		0.30	0.04	3.54

Variation in Non-Time-Varying Level Two Variables Between Clinicians

Variable		Mean	SD	Min	Max
Mean Psychiatric Symptom Severity Composite	overall	6.45	3.01	0.00	20.00
	between		2.59	1.00	19.00
	within		2.15	-1.42	18.53
Mean Family/Caregiver Relationship Strength	overall	1.92	0.77	0.00	3.00
	between		0.60	0.00	3.00
	within		0.60	-0.46	3.72
Mean Peer and Non-Family/Caregiver Strength	overall	1.90	0.80	0.00	3.00
	between		0.64	0.00	3.00
	within		0.63	-0.30	3.76
Mean School System and Educational Plan Strength	overall	1.78	0.87	0.00	3.00
	between		0.65	0.00	3.00
	within		0.69	-0.55	3.72
Mean Extracurricular Activities and Talents Strength	overall	1.93	0.88	0.00	3.00
	between		0.71	0.00	3.00
	within		0.64	-0.64	4.21
Mean Spiritual/Religious Beliefs / Involvement Strength	overall	1.81	1.18	0.00	3.00
	between		0.97	0.00	3.00
	within		0.77	-1.00	4.56
Mean Relationship Permanence Strength	overall	2.03	0.87	0.00	3.00
	between		0.73	0.00	3.00
	within		0.61	-0.42	3.97
Male	overall	4.40	4.09	0.00	26.00
	between		3.91	0.00	22.00
	within		2.65	-5.80	24.25

Age_cent	overall	0.17	3.85	-7.36	9.64
	between		3.18	-7.36	6.64
	within		2.73	-8.49	8.90
White	overall	0.08	0.27	0.00	1.00
	between		0.23	0.00	1.00
	within		0.22	-0.59	1.04
Black	overall	0.33	0.47	0.00	1.00
	between		0.39	0.00	1.00
	within		0.35	-0.54	1.28
Latino	overall	0.28	0.45	0.00	1.00
	between		0.38	0.00	1.00
	within		0.30	-0.65	1.25
Asian	overall	0.22	0.41	0.00	1.00
	between		0.33	0.00	1.00
	within		0.26	-0.74	1.16
Pacific Islander/Native Hawaiian	overall	0.02	0.15	0.00	1.00
	between		0.10	0.00	1.00
	within		0.13	-0.31	1.00
Multi-Ethnic	overall	0.03	0.18	0.00	1.00
	between		0.14	0.00	1.00
	within		0.16	-0.47	1.02
Other	overall	0.03	0.18	0.00	1.00
	between		0.15	0.00	1.00
	within		0.15	-0.57	1.01

Results

The residuals for the response variable, psychiatric symptom severity, were checked to make sure that the multi-level modeling assumptions were met. Empirical Bayes estimators of the residuals were used instead of maximum likelihood estimators because empirical Bayes procedures use the data from all clusters to estimate the parameters and weight by cluster size, thus tending to produce more accurate standard errors, particularly if cluster size happens to be small (in this data, cluster sizes vary considerably). Figure 1 suggests that the empirical Bayes predictions of the clinician-level residuals of the response variable generally meet the assumption of normal distribution. Figure 2 suggests that the empirical Bayes predictions of the youth-level residuals of the response variable generally meet the assumption of normal distribution. Figure 3 provides further evidence that the empirical Bayes predictions of the response variable residuals generally meet the assumption of normality. However, there are some outliers, and though they do not appear overly serious, robust standard errors (which are more conservative, or larger, than non-robust standard errors) were calculated for the models to account for any violations to the normality assumption.

Figure 1.

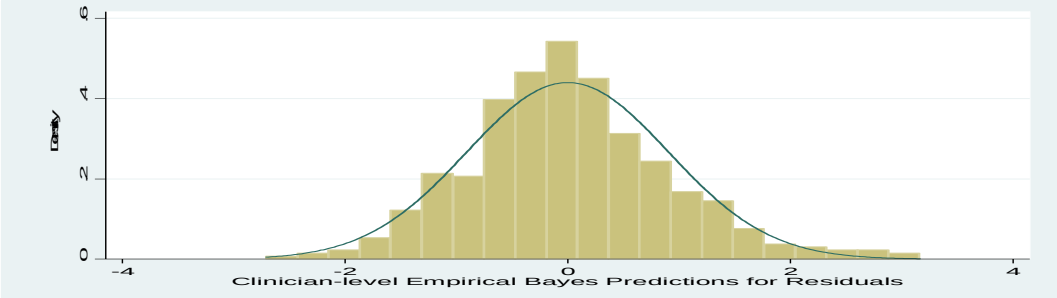


Figure 2.

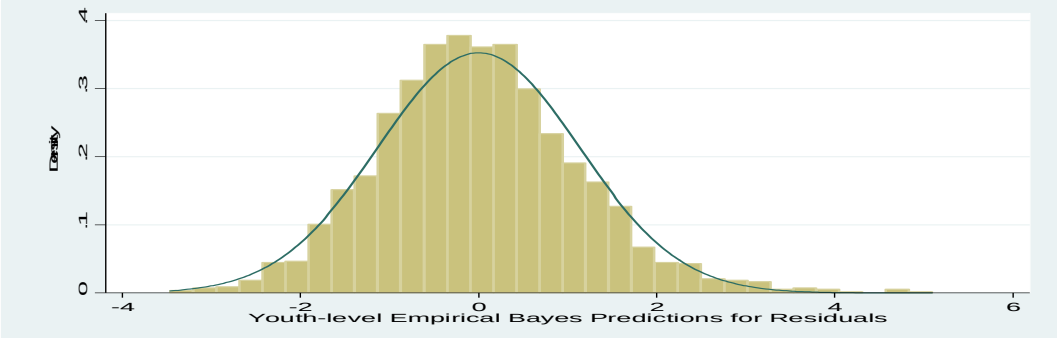
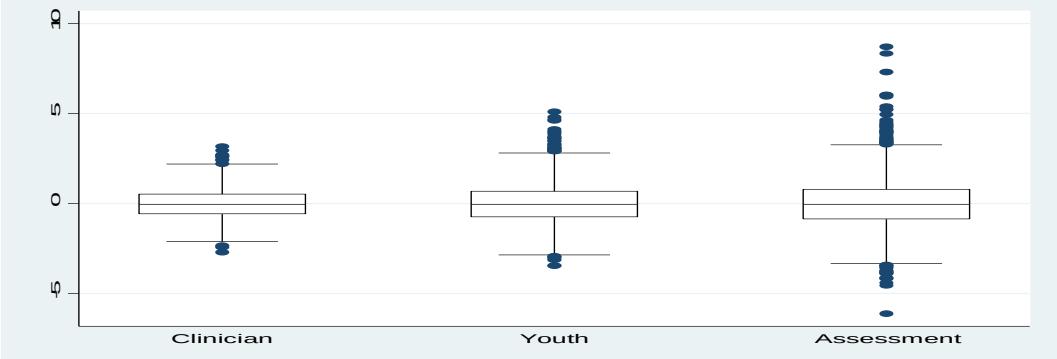


Figure 3.



Research question one was addressed by running a variance component model (model 1) and a random intercept model (model 2). Research question two was addressed by running a random intercept model with cross-level interactions (model 3). Research questions three and four were addressed by running a random intercept model and including the mean-centered versions of the time-varying variables (model 4). Likelihood ratio tests confirmed that for each model, the between-clinician, within-youth variance was significantly different from zero, supporting the use of multi-level modeling estimation techniques in this data (p-values are reported in Table 4 as the "between clinicians standard deviations"). Likelihood ratio tests were also used to confirm that model 3 was a better fit for the data than model 2 or model 1, and model four was a better fit for the data then model 1 (see Table 3).

Table 3
Likelihood Ratio Tests Comparing Nested Models

	$\chi 2$	df	p-value
Model 2 vs. 1	571.92	17	<.001
Model 3 vs. 1	613.87	31	<.001
Model 3 vs. 2	41.95	14	0.0001
Model 4 vs. 1	699.38	23	<.001

Table 5 reports the results from the four models.

Table 5

Multi-level Models of Psychiatric Symptom Severity Over Twelve Months of Intervention in a Children's System of Care.

	Model 1		Model 2		Model 3		Model 4	
	Beta (RobustSE)	*CI*	*Beta (RobustSE)*	*CI*	*Beta (RobustSE)*	*CI*	*Beta (RobustSE)*	*CI*
Fixed Part								
Intercept	6.63(.11)	6.42, 6.84	8.24(.37)***	7.51, 8.96	8.16(0.36)	7.45, 8.88	8.61(.37)***	7.89, 9.32
Six Month Assessment Dummy[a]			-.98(.11)***	-1.19, -.76	-0.36(0.39)	-1.12, 0.40	-.86(.10)***	-1.06, -.65
12 Month Assessment Dummy			-1.15(.17)***	-1.49, -.81	-1.93(0.75)**	-3.41, -0.46	-1.05(.14)***	-1.33, -.76
Initial Family/Caregiver Relationship			-.16(.08)***	-.52, -.20	-0.35(0.08)***	-0.51, -0.19		
Initial Peer and Non-Family/Caregiver			-.48(.08)***	;-.63, -.32	-0.52(0.08)***	-0.68, -0.36		
Initial School System and Educational Plan			-.47(.07)***	-.61, -.32	-0.46(0.08)***	-0.61, -0.31		
Initial Extracurricular Activities and Talents			.02(.08)	-.13, .17	0.02(0.08)	-0.13, 0.17		
Initial Spiritual/Religious Beliefs/Involvement			-.01(.06)	-.12, .10	-0.01(0.06)	-0.12, 0.10		
Initial Relationship Permanence			-.02(.07)	-.16, .12	-0.02(0.07)	-0.16, 0.12		
Time-Varying Strengths[b]								
Family/Caregiver Relationship (TV)							-.29(.17)	-.62, .05
Peer and Non-Family/Caregiver (TV)							-.37(.15)*	-.65, -.08
School System Educational Plan (TV)							-.48(.13)***	-.74, -.23
Extracurricular Activities Talents (TV)							-.11(.14)	-.38, .16
Spiritual/Religious Beliefs/Involvement (TV)							.07(.09)	-.11, .24
Relationship Permanence (TV)							-.36(.12)**	-.59, -.12
Lifetime Exposure to Trauma			.17(.02)***	.13, .20	0.18(0.02)***	0.1 , 0.22	.16(.02)***	.13, .20
Male			.54(.11)***	.33, .74	0.60(0.12)***	0.37, 0.83	.53(.11)***	.32, .73
Age (centered)			.07(.02)***	.04, .10	0.08(0.02)***	0.05, 0.11	.07(.02)***	.03, .10
Ethnicity[c]								
Black			.03(.21)	-.38, .44	0.04(0.21)	-0.36, 0.45	.03(.21)	-.37, .44
Latino			.13(.23)	-.33, .59	0.15(0.24)	-0.31, 0.62	.16(.24)	-.30, .62
Asian			-.77(.24)**	-.91, .42	-0.75(0.25)***	-1.24 , -0.26	-.76(.24)**	-.23, -.29
Multi			.07, .32	-.55, .70	-0.24(0.34)	-0.91, 0.42	-.27(.35)	-.95, .41
Pacisl			-.24(.34)	-.91, .42	0.11.32)	-0.51, 0.73	.05(.32)	-.57, .67
Other			-.01(.35)	-.70, .68	0.02(0.35)	-0.66, 0.70	.04(.35)	-.64, .73
Interactions								
famnew_initXAssess6dum					-0.10(0.14)	-0.36, 0.17		
famnew_initXAssess12dum					0.16(0.24)	-0.30, 0.62		
educnew_initXassess6dum					-0.01(0.13)	-0.27, 0.24		
educnew_initXAssess12dum					-0.05(0.18)	-0.41, 0.31		
interpnew_initXassess6dum					0.03(0.12)	-0.21, 0.28		
interpnew_initXassess12dum					0.61(0.20)***	0.21, 1.01		
trauma_initXassess6dum					-0.07(0.02)**	-0.11, -0.01		
trauma_initXAssess12dum					-0.09(0.04)*	-0.17, -0.01		
maleXassess6dum					-0.33(0.19)	-0.71, 0.04		
maleXassess12dum					-0.09(0.31)	-0.71, 0.53		
age_centXassess6dum					-0.03(0.02)	-0.07, 0.02		
age_centXassess12dum					-0.03(0.04)	-0.11, 0.05		
asianXassess6dum					0.20(0.22)	-0.22, 0.63		
asianXassess12dum					-0.38(0.37)	-1.11, 0.34		
Mean-Centered Variables								
Mean_Family/Caregiver Relationship							-.15(.20)	-.55, .25
Mean_Peer and Non-Family/Caregiver							-.14(.18)	-.49, .20
Mean_School System and Educational Plan							-.01(.15)	-.30, .28
Mean_Extracurricular Activities and Talents							-.11(.16)	-.21, .42
Mean_Spiritual/Religious Beliefs/Involvement							-.09(.11)	-.30, .12
Mean_Relationship Permanence							.32(.14)*	.04, .60
Random Part								
Between Clinicians standard deviation	1.70(.10)***	1.51, 1.91	1.31(.09)***	1.14, 1.51	1.31(.09)***	1.14, 1.51	1.33(.10)***	1.15, 1.53
Between Youth standard deviation	1.75(.08)	1.59, 1.92	1.62(.07)	1.48, 1.77	1.63(.07)	1.50, 1.78	1.65(.07)	1.51, 1.79
Between Assessments standard deviation	1.91	1.78, 2.04	1.76(.06)	1.64, 1.89	1.73(.06)	1.62, 1.85	1.68(.06)	1.57, 1.79

Note

[a]Reference category = Initial Assessment. [b] TV=Time-Varying. [c]Reference category = White/Caucasian.

*p<.05 **p < .01. ***p < .001.

Research question one explored the reported change in psychiatric symptom severity for the average youth client, if any, at six month and twelve month assessment, adjusting for the other covariates. Model two estimated that the average psychiatric severity score at initial assessment was 8.24 points out of 27 possible points (SE:.37) for white females aged 12.5 years with no reported strengths or trauma. At six months, on average, youth improved (had a reduction in their psychiatric severity score) by .98 points on the 27 point scale (SE: .11) (controlling for covariates). Between six months and twelve months, on average, youth improved an additional .17 points (out of 27 points), controlling for covariates.

Higher levels of initial trauma were associated with less reduction in symptoms at twelve month assessment. For every point increase in reported exposure to trauma at initial assessment, the mean psychiatric score at twelve month assessment significantly increased by an additional .17 (SE: .02) points out of 27 points (youth get worse), adjusting for covariates. At a given time, males had a significantly higher average psychiatric score of .54 (SE: .11) points compared to females. For example, at initial assessment, males aged 12.5 years with no strengths or trauma had an average symptom score of 8.78 points (out of a possible 27 points), .54 points higher than the average, comparable female. The mean psychiatric score at a given time was significantly higher by .07 (SE: .02) points for youth a year older in age, adjusting for covariates (though this is so close to zero as to be clinically meaningless). Asian youth had a mean psychiatric score that was, on average, .77 (SE: .24) points lower than white youth at twelve month assessment, adjusting for covariates.

Research question two explored the extent of change in psychiatric symptom severity, if any, associated with client level factors upon entry into the system, including initial youth strengths intensity, initial psychiatric severity, lifetime trauma experiences, or youth demographics (gender, age, and ethnicity). As reported in model 3, it was estimated that the average psychiatric severity score at initial assessment was 8.16 points (SE: .36) for white females aged 12.5 years with no reported strengths or trauma. After controlling for the other covariates, the mean slope of change in psychiatric symptom severity between six months and initial assessment is -.36 for youth without strengths. The six month / trauma interaction term suggested that youth who reported higher levels of exposure to trauma at initial assessment had a significantly greater decrease in psychiatric symptoms over six months. The six-month and youth strength interaction terms suggested that there is little evidence that higher rates of youth strengths at initial assessment resulted in significantly higher or lower rates of change over six months. Similarly, there is little evidence that gender, age, or ethnicity impact rate of change over six months. Given that the average, white, 12.5 year old female with no strengths or trauma had an estimated symptom score of 8.16 points at initial assessment, boys scored an average of .33 points lower than girls, at a given time, for a given age, for a given ethnicity, for a given trauma score, for a given strength, but the difference was not significant at the .05 level. The mean slope of change over twelve months of intervention was 1.93 points per year for male, white youth, aged 12.5 years, with no trauma and no strengths.

The twelve month interaction terms in model three suggest that every point increase on the peer/interpersonal relationships strength item at initial assessment is significantly associated with a smaller decrease in psychiatric score by .61 points, but there is little evidence that higher levels of any of the other strengths at initial assessment predict higher or lower rates of change in psychiatric symptom severity. Youth with higher levels of trauma exposure at initial assessment do report a significantly greater decrease in symptom severity over twelve months. There is little evidence that gender, age, or ethnicity predict rate of change over twelve months.

Question three explored the estimated contribution to variation in psychiatric symptom severity change at six and twelve month assessment by the clinician who provided the services, if any. Two intra-class correlations were calculated based on Model 4. The third level intra-class correlation (ICC3) was estimated at .30. This can be interpreted in two ways: 1) 30% of the residual variance in psychiatric symptom severity is due to unmeasured clinician variables, or 2) the estimated residual correlation on psychiatric symptom severity between two different youth for the same clinician is .30. The fact that 30% of the residual variance in psychiatric symptom severity was estimated as due to clinician factors suggests that clinicians played a large role in mental health outcomes in this public system; this number is aligned with estimates of clinician contribution in other studies (e.g. Duncan et al., 2010; Lutz et al., 2007; Wampold & Brown, 2005). This suggests that, at least in San Francisco's System of Care, unmeasured clinician variables contributed to psychiatric symptom severity.

The second level intra-class correlation (ICC2) is .62. This can be interpreted in two ways: 1) 62% of the residual variation in psychiatric symptom severity for youth with the same clinician is due to unmeasured youth variables, or 2) the estimated residual correlation on psychiatric symptoms severity between two different assessment points for the same youth, same clinician, is .62. The large contribution of unmeasured youth variables to variance in psychiatric score is not unexpected given the large number of client factors that are unmeasured in this dataset and related studies that have found unmeasured client variables as accounting for similar amounts of variation in clinical outcomes (e.g. Bohart & Tallman, 2010).

Research question four explored the extent that change in strengths at six and twelve month assessment in San Francisco's System of Care predicted change in psychiatric symptom severity over twelve months of treatment, if any, adjusting for lifetime trauma experience and youth demographics (gender, age, and ethnicity). This modeled included the time varying strength items and the mean-centered strength items (instead of the initial assessment strength items, as models two and three did). This allowed the analysis of the between youth effect and the within-youth effects separately. The mean-centered strength items represent the estimated within-youth effect on mean psychiatric score of change in each strength item over twelve months (the interpretation of these variables is not as meaningful to the current analysis). However, the between youth effect, estimated by the beta coefficients of the time-varying strength items, addressed the extent to which change in strengths over twelve months of intervention in a public system was associated with change in psychiatric symptom severity.

Model 4 estimated that the average psychiatric severity score at initial assessment was 8.61 points (SE: .37) for white females aged 12.5 years with no reported strengths or trauma. Model four estimated that every point increase in peer and non-family/caregiver strength reported at twelve month assessment was associated with a decrease in average mean psychiatric score by an additional .37 points (out of 27 total points), compared to the overall average rate of change on mean psychiatric score. Every point increase in school system and educational plan strength was associated with an estimated decrease in average mean psychiatric score by an additional .48 points (out of 27 total points) compared to the overall average rate of change on mean psychiatric score. Every additional point increase on relationship permanence at twelve month assessment was associated with a decrease in average estimated mean psychiatric score by an additional .36 points (out of 27 total points). There were no significant differences in rate of change for the other three strength items.

Discussion

This paper explored the relationship between strengths and psychiatric symptom severity over twelve months of intervention in San Francisco's Children's System of Care (SF-CSOC). A meta-analysis or a systematic review of the relationship between strengths and mental health outcomes in Systems of Care does not exist to the knowledge of this author. This paper reviewed some of the existing empirical literature on the relationship between strengths and mental health outcomes in System of Care (SOC) samples, with a focus on the limitations of this literature in order to highlight future areas for research.

Research question one explored the average amount of change in psychiatric symptom severity, if any, at six month and twelve month assessment in San Francisco's System of Care (SF-CSOC). In this sample, the average psychiatric severity score at initial assessment for white females, aged 12.5 years, with no strengths and no trauma, was estimated at 8.24 points out of a possible 27 points (SE:.37). Larger improvement in youth symptoms were reported during the first six months of services in San Francisco's System of Care compared to between six month and twelve month assessment. At six month assessment, on average, clinicians reported that youth had improved (had a reduction in their psychiatric severity score) by an estimated one point (beta=-.98, SE: .11), controlling for other covariates. Between six months and twelve months, on average, youth were estimated by clinicians to improve an additional .17 points (controlling for covariates). Regression to the mean, a common issue in longitudinal public systems data (Lambert & Bickman, 2004) likely contributed to this symptom improvement.

Research question two explored the association between initial levels of youth strengths and change in psychiatric symptom severity at twelve month assessment. These analyses suggested that the initial level of three youth strengths items (family/caregiver, peer relationships, or education strengths) was associated with increased improvement in psychiatric symptom reduction over time, compared to the average improvement. Specifically, youth who entered San Francisco's System of Care (SF-CSOC) one point higher on the family/caregiver item (out of a possible 27 points) had a reduction in psychiatric symptom severity of an average extra .35 points (average reported reduction was 1.15 points) at twelve month assessment. Compared to this average reported reduction of 1.15 points, higher initial levels on the peer relationships item or the education plan item were associated with an average extra improvement of .52 points (SD: .08) and .46 points (SD: .08) over twelve months, respectively.

Research question three explored the amount of unexplained variance that was accounted for by unmeasured time, youth, and clinician variables. Related analyses suggested that a large amount of unexplained variance in psychiatric symptom severity was accounted for by unmeasured youth (approximately 62%) and unmeasured clinician variables (approximately 30%) (see discussion of common factors in Duncan et al., 2010). Other variables (unmeasured) might explain some of the variation (i.e. medication, therapeutic alliance). Future work should make an effort to include other youth and clinician variables.

The contribution of unmeasured clinician variables in these analyses suggests that further attention to other clinician variables may be useful in improving service delivery for youth. Perhaps clinicians who use a strengths-based orientation are more effective at reducing symptoms than clinicians who do not focus on strengths. For example, Cox (2006) found that, in a sample of 84 youth with severe emotional and behavioral problems receiving therapy in a public outpatient program, the strengths-based orientation of the clinician moderated the relationship between receiving a strengths-based assessment and improvement in emotional and

behavioral problems. Future work examining a strength-based orientation, and other related clinician variables, may be a useful next step.

Research question four explored whether change in strengths was associated with rate of change in psychiatric symptom severity at twelve month assessment. These analyses suggested (though directionality cannot be determined) that reported gains on peer relationship, education plan and system, and relationship permanence strengths between initial and twelve month assessment in San Francisco's Children's System of Care (SF-CSOC) were associated with larger reductions in symptoms (adjusting for the other covariates). Youth who gained on these strengths had an increased rate of improvement compared to youth who did not gain on these strengths, though rate of change was modest. The largest rate of improvement was reported for the education plan and system strength, as a one point gain (out of a total of 27 points). This gain was associated with an estimated increased rate of improvement of symptoms of nearly half a point (beta=.48, SD: .13). These findings are associational and limited by regression to the mean, directionality, and other issues related to administrative dataset, as discussed later in detail.

Also, males appear to enter with higher levels of psychiatric severity and to get better more slowly than females. Further analyses could explore whether level of strengths moderated the relationship between gender and psychiatric symptoms. Perhaps some strengths have a larger impact on psychiatric symptoms for males than females, or vice versa. Ethnicity differences have not been explored in these analyses and also warrant further attention.

Limitations. This paper has several limitations, many of which are related to the reliance on administrative data (data collected for non-research purposes, but which may be used for research). It is not possible to determine directionality in this dissertation (e.g. time order cannot be determined) because the data were collected through clinician self-report using a pre-post design. The dataset is missing many variables that might impact both strengths and psychiatric symptoms, such as medication and interventions offered or used. Behavioral theory and related empirical research (e.g. Madden, 2013) suggest that increase in strengths can decrease symptoms. Future analyses could make use of a cross-lagged panel model (e.g. Finkel, 1995) or propensity score matching techniques (e.g. Rosenbaum & Rubin, 1983) to explore directionality further.

And, the analysis is limited by regression to the mean (a statistical phenomenon where extreme scores gravitate towards the group average score over time with a non-random sample). Regression to the mean, or the "clock-setting cure" (Lambert & Bickman, 2004), is a common phenomenon in health services research. At initial assessment in a System of Care, youth are likely experiencing extreme emotions and behaviors. Regression to the mean suggests that these extreme emotions and behaviors would likely abate naturally over time, regardless of intervention. Intervention studies often "set the clock" at an initial assessment, when youth are at their worst, but natural history is likely responsible for at least some of the change. For example, Lambert & Bickman (2004) conducted a statistical simulation (n=10,000 youth) that found that it was possible that initial improvements among youth in the Fort Bragg System of Care evaluation (Bickman, 1996) were likely due, at least in part, to regression to the mean, or the natural history effect. While there are various statistical techniques to partially address regression to the mean in longitudinal statistical models, this paper used a growth model approach (Allison, 1990) where the final estimates are adjusted for the contribution of unmeasured time variables. Even so, caution should be used in interpreting results.

Another limitation is that the SF-CANS was designed as a clinician self-report tool, subject to reporting bias and documentation error. The SF-CANS does not include youth or

parent report to act as a check on the clinician's report. There is no way to track number of contacts, number of hours the clinicians spends on contacting stakeholders before filling out the CANS, and the process the clinician went through to obtain the signature of the primary caregiver. While CANS training emphasizes the importance of collaboration with the youth and family during the assessment process, it is possible that clinicians have not fully elicited or integrated information from all relevant parties. Observation (such as videotaping assessments) has not been used, either in other studies using the CANS or in the collection of the data used in this dissertation. The percentage of youth referred for services, where they were referred from, the percentage who went to services, and the type, quantity, and quality of interventions (if any) are not reliably tracked. Though only 8% of youth in this dataset had complete data (an initial, six month, and twelve month assessment), reasons for termination are not reliably tracked, but could include dropping out, moving out of area, or completing a course of therapeutic intervention successfully, partially successfully, or unsuccessfully. Also, the SF-CANS User Manual is more than 100 pages long, and it is possible that clinicians might not have the time or motivation to consult the manual every time they are scoring an assessment, thus resulting in inaccurate scoring.

Also, since administrative data is not collected with research as the primary goal, the reliability and validity of the data (which is of great concern to researchers) may not have been prioritized by the people collecting the administrative data. For example, the spiritual religious item is confusing because the San Francisco CANS Scoring Manual (2010, pg. 55) does not clarify whether the item can be interpreted as 1) the youth is involved in a religious community or 2) the youth identifies with a spiritual or religious practice and finds strength in this, or both of these. Researchers can only hypothesize as to what the clinician intended when interpreting this item. Future psychometric testing would be helpful here in order to further explore the validity of the SF-CANS items. Even though some validity issues have been examined using some versions of the CANS, since every community has its own version, every community should consider conducting its own validity study in order to understand what the CANS items mean, and San Francisco could lead the way in this. Qualitative work might also be helpful here, for example interviews or focus groups with clinicians who use the SF-CANS could shed light on how clinicians understand, use, and interpret the Strengths domain items, the referral process and how many took up services, reasons for termination, what interventions are provided, if any, and other factors that are not currently reliably tracked in SF-CSOC.

Also, though the SF-CANS may be considered a strengths-based assessment, using the broad definition above, most of the hundreds of possible items are items that measure problems, symptoms, and deficits. There is only one Strengths domain included in a version of the CANS, with eight or ten strength items included at most. Also, some strengths claimed to be related to improved functioning (e.g. self-efficacy, quality of neighborhood or community resources) are not included. The addition of other strengths items would provide a more nuanced picture of the strengths profiles of youth entering SF-CSOC.

Relatedly, this study is unavoidably limited by the number of data items collected by clinicians in the CANS assessment and available for analysis. For example, several risk factors for negative health outcomes in youth were not measured in the CANS assessment (e.g. medications, intensity of exposure to trauma items). Similarly, trauma exposure is only measured at initial assessment, making it impossible to account for the effect of additional trauma exposure during course of treatment. Additionally, this dataset does not contain any information on type or intensity of treatment received, limiting our ability to control for any of these factors over

time. This is a not uncommon limitation of public systems data (e.g. Bickman, 1999; Hoagwood & Kolko, 2009), although it is a problem that requires addressing in future work in order to move the field forward.

Also, types of strengths, types of trauma, and types of psychiatric symptom severity were weighted equally in this study. Research and theory have suggested that the same risk or strength may have different impact in different contexts, when grouped with various other risks and resources, at different developmental stages, and for different populations and outcomes (e.g. Luthar, et al., 2000; Pynoos, 2009). Providing a more fine-grained analysis of these types and characteristics may allow for a clearer understanding of the processes responsible for distinct clinical and resource trajectories. A recent meta-analysis of the Wraparound intervention (Bruns and Suter, 2010) reported that "building assets and strengths" was measured as an outcome in two studies, and sample size was thus so small that they could not calculate an effect size.

A closer examination into the theoretical and empirical literature on gender, age, and ethnic group differences among youth receiving public mental health services is warranted to provide context for these analyses and to further interpretation. For example, group differences between Asian youth and other groups could use contextualization. Also, it might be useful to include dummy variable of age categories (e.g. 5-8 years, 9-12 years) since age was a significant predictor of rate of change in rate of change in psychiatric symptom severity.

In Summary. The analyses in this paper suggest that youth in San Francisco's Children's System of Care are reported to experience modest symptom improvement of around one point (out of 27 total points) between initial and twelve month assessment. Most of the reported improvement occurred between initial and six month assessment, with very little improvement occurring between six month and twelve month assessment. Youth with reported gains on peer relationship, education plan and system, or relationship permanence strengths between initial and twelve month assessment got better faster (had a larger rate of improvement) compared to youth who did not gain on these strengths. (Though gain on family relationships, extracurricular talents and interests, or spiritual/religious strengths were not associated with reduced symptoms.) Unmeasured client and clinicians variables accounted for substantial amounts of unexplained variation in symptom change.

Findings should be interpreted with caution due to the nature of this administrative data and related limitations. Future work exploring whether clinicians using a strengths-based orientation to clinical work are more effective at reducing youth symptoms may be a useful next step. Behavioral theory and related research in applied and empirical settings suggest a directional relationship between strengths and psychiatric symptom severity where increase in strengths decreases symptoms. The nature of the data in this study prohibit determination of the directional model. Researchers should consider non-administrative datasets and statistical techniques (including propensity score matching or cross-lagged panel models) in order to address some of the limitations of this study and take steps towards exploring evidence for causal direction.

CONCLUSION OF THE DISSERTATION

Sarah Accomazzo

School of Social Welfare
University of California, Berkeley

Conclusion of this Dissertation

Overall Themes of this Dissertation

The aim of this dissertation was to explore the relationship, if any, between strengths (including those of the client, the client's family, and the client's environment) and psychiatric symptom severity for "youth" (children and adolescents aged 5-22) in contact with a mental health and/or substance abuse agency in a public System of Care. This dissertation explored a public systems dataset that included administrative data for every client aged 5-22 (n=2049) who entered into San Francisco's Children's System of Care (SF-CSOC) from June 2010-August 2013. Assessment data were collected at initial assessment (entry to the system), six months after initial assessment, and twelve months after initial assessment. Given the kind of data used, time order regarding the role of strengths in decreasing symptoms cannot be determined. Also, several variables, including use of medication and interventions offered or used, were not included in the data. Behavioral/social learning theory and related research in applied and experimental settings (e.g. Goldiamond, 1974; Staats, 2012) suggests that growth in strengths would be associated with reduction in symptoms.

The dissertation introduction provides a background to the entire dissertation and orients to the reader to various terminological and conceptual issues relevant to all three papers, including how problems are framed (e.g. as pathological), behavioral theory and related research, the Child and Adolescent Needs and Strengths (CANS) tool, San Francisco's Children's System of Care (SF-CSOC), the San Francisco CANS, and the benefits and limitations of administrative data.

Paper one reported the prevalence of clinician-reported strengths in San Francisco's public System-of-Care (SF-SOC), explored the psychometric properties of San Francisco's Child and Adolescent Strengths and Needs (SF-CANS) assessment's Strengths section, and described three alternative modeling options used to explore group differences (by gender, age, and ethnicity). Youth assessed at entry to San Francisco's System of Care were reported by clinicians to enter with considerable strengths, similar to other studies reporting prevalence of strengths in a System of Care (e.g. CMHI, 2011; Dunleavey et al., 2011). A confirmatory factor analysis found a hypothesized two factor model was a better fit for the data than a single factor model, though the fit of the two factor model was only adequate and the face validity of the factors were questionable. Individual strength items were assessed as the most informative for future analyses, compared to use of factors or composite scores.

Paper two explored the relationship between clinician-reported strengths and psychiatric symptoms at initial assessment. Higher levels of family relationships, peer relationships, and educational strengths were significantly associated with lower levels of psychiatric symptoms, after accounting for variation due to clinician, program, reported trauma exposure and youth demographics. Oher individual strengths (i.e. relationship permanence, extracurricular talents and interests, and spiritual/religious beliefs or involvement) were not significantly associated with psychiatric symptom severity.

Paper three explored 1) the overall rate of change in psychiatric symptom severity at six and twelve month assessment, 2) whether initial levels of strengths were associated with rate of change in psychiatric symptoms twelve months after initial assessment, 3) whether clinician, youth, or time contributed to the unexplained variation in psychiatric symptom severity, and 4) whether a change in strengths twelve months after initial assessment was associated with a change in psychiatric symptoms. These analyses estimated that, on average, youth psychiatric

symptoms reported by clinicians decreased by a point (out of 27 possible points) between initial and six month assessment, and decreased by an additional .2 points (out of 27 possible points) between six month and twelve month assessment, adjusting for time, clustering, and youth demographics. Higher initial levels of peer relationships, family and caregiver, or school system/education plan strengths were significantly associated with reductions in psychiatric symptom severity over twelve months, adjusting for clustering, time, and youth demographics. Unmeasured clinician and client variables contributed an estimated 30% and 62% of the unexplained variation in psychiatric symptom change, suggesting that future analyses should include related clinician and client variables. Also, it was estimated that an increase in peer relationships, relationship permanence, or school system/education plan strengths over twelve months of intervention was significantly associated with a larger decrease in psychiatric symptoms over twelve months, adjusting for clustering, time, and youth demographics.

Limitations of this Dissertation

This dissertation has several limitations, many of which are related to the fact that this dissertation relied on administrative data (data collected for non-research purposes, but which may be used for research). First, it is impossible to establish directionality between strengths and symptoms due to the nature of the data. Time order cannot be determined, as the clinicians assessed and rated strengths and psychiatric symptom severity at the same time. Data were collected through clinician self-report with no check on the validity of the data. Additionally, the administrative dataset did not provide information on many characteristics that might be partially or fully responsible for the reduction in psychiatric symptom severity (e.g. medication, therapeutic alliance, type, quantity, and quality of intervention), and it is possible that the associations reported in this dissertation might disappear altogether with the inclusion of these other variables. Behavioral theory and related empirical research suggest that the identification and construction of strengths are associated with a reduction in psychiatric symptom severity. However, the data used in this study prohibit estimates regarding the directionality issue. Indeed, regression to the mean (extreme scores gravitate towards the group average score over time) may be responsible for a decrease in symptoms.

Regression to the mean, or the "clock-setting cure" (Lambert & Bickman, 2004), is a common phenomenon in health services research. At initial assessment in a System of Care, youth are likely experiencing extreme emotions and behaviors. Regression to the mean suggests that these extreme emotions and behaviors would likely abate naturally over time, regardless of intervention. Intervention studies often "set the clock" at an initial assessment, when youth are at their worst, but natural history is likely responsible for at least some of the change. For example, Lambert & Bickman (2004) conducted a statistical simulation (n=10,000 youth) that found that initial improvements among youth in the Fort Bragg System of Care evaluation (Bickman, 1996) were likely due, at least in part, to regression to the mean, or the natural history effect. While there are various statistical techniques to partially address regression to the mean in longitudinal statistical models, this dissertation used a growth model approach (Allison, 1990) where final estimates are adjusted for the contribution of unmeasured time variables to variance in outcomes. Even so, caution should be used in interpreting results.

Another limitation is that the San Francisco Child and Adolescent Needs and Strengths (SF-CANS) tool was designed as a clinician self-report tool, subject to reporting bias and documentation error. The SF-CANS does not include youth or parent reports as a check on clinician reports. There is no way to track number of contacts, number of hours the clinicians

spends on contacting stakeholders before filling out the CANS, and the process the clinician went through to obtain the signature of the primary caregiver. While CANS training emphasizes the importance of collaboration with the youth and family during the assessment process, it is possible that clinicians have not fully elicited or integrated information from all relevant parties. Observation (such as videotaping assessments) has not been used, either in other studies using the CANS or in the collection of the data used in this dissertation. The percentage of youth referred for services, where they were referred from, the percentage who went to services, and the type, quantity, and quality of interventions (if any) are not reliably tracked. Though only 8% of youth in paper three had complete data (an initial, six month, and twelve month assessment), reasons for termination are not reliably tracked, but could include dropping out, moving out of area, or completing a course of therapeutic intervention successfully, partially successfully, or unsuccessfully. Also, the SF-CANS User Manual is more than 100 pages long, and it is possible that clinicians might not have the time or motivation to consult the manual every time they are scoring an assessment, thus resulting in inaccurate scoring. Descriptions of items and scoring options are sometimes vague, risking that clinicians may score inconsistently based on definitional confusion.

Also, since administrative data is not collected with research as the primary goal, the reliability and validity of the data (which is of great concern to researchers) may not have been prioritized by the people collecting the data. For example, the spiritual / religious item is confusing because the manual does not clarify whether the item can be interpreted as 1) the youth is involved in a religious community that is a source of strength, or 2) the youth identifies with a spiritual or religious practice and finds strength in this, or both of these. Researchers can only hypothesize as to what the clinician intended when interpreting this item. Future psychometric testing would be helpful here in order to establish the validity of the SF-CANS items. Even though some validity issues have been examined using some versions of the CANS, since every community has its own version, perhaps every community should consider conducting its own validity study in order to understand how CANS items are interpreted. San Francisco could lead the way in this. Qualitative work might also be helpful here, for example interviews or focus groups with clinicians who use the SF-CANS could shed light on how clinicians understand, use, and interpret the Strengths domain items, the referral process and how many took up services, reasons for termination, what interventions are provided, if any, and other factors that are not currently reliably tracked in SF-CSOC.

Also, though the SF-CANS may be considered a strengths-based assessment, using the broad definition suggested, most of the hundreds of possible items measure problems, symptoms, and deficits. There is only one Strengths domain included in a version of the CANS, with eight or ten strength items included at most. Also, some strengths claimed or found to be related to improved functioning (e.g. self-efficacy, quality of neighborhood or community resources) are not included. The addition of other strengths items would provide a more nuanced picture of the strengths profiles of youth entering SF-CSOC.

Relatedly, this study is unavoidably limited by the number of data items collected by clinicians in the CANS assessment and available for analysis. For example, several factors that may be related to psychiatric symptom severity reduction were not measured in the CANS assessment (e.g. medications use, type, and duration, intensity of exposure to trauma items, therapeutic alliance, caregiver and family history of emotional and behavioral problems). Similarly, trauma exposure is only measured at initial assessment, making it impossible to account for the effect of additional trauma exposure during course of treatment. Additionally,

this dataset does not contain any information on type or intensity of treatment received, limiting our ability to control for any of these factors over time. This is a not uncommon limitation of public systems data (e.g. Bickman, 1999; Hoagwood & Kolko, 2009), although it is a problem that requires addressing in future work in order to move the field forward.

Also, types of strengths, types of trauma, and types of psychiatric symptom severity were weighted equally in this study. Research and theory have suggested that the same risk or strength may have different impacts in different contexts, when grouped with various other risks and resources, at different developmental stages, and for different populations and outcomes (e.g. Luthar, et al., 2000; Pynoos, 2009). Providing a more fine-grained analysis of these types and characteristics may allow for a clearer understanding of the operative processes responsible for distinct clinical and resource trajectories. A recent meta-analysis of the Wraparound intervention (Suter & Bruns, 2009) reported that "building assets and strengths" was measured as an outcome in two studies, and sample size was thus so small that they could not calculate an effect size.

Additionally, a closer examination into the theoretical and empirical literature on gender, age, and ethnic group differences among youth receiving public mental health services is warranted to provide context for these analyses and to further interpretation. For example, group differences between Asian youth and other groups could use contextualization. Also, it might be useful to include dummy variable of age categories (e.g. 5-8 years, 9-12 years) since age was a significant predictor of rate of change in rate of change in psychiatric symptom severity.

Future Directions for Research

Some areas for future research are indicated. Within this dissertation, causal direction between strengths and symptoms cannot be critically tested due to the nature of the data. Future work could attempt to explore this. For example, future analyses could make use of statistical techniques that explore causality, such as a cross-lagged panel model (e.g. Finkel, 1995) or propensity score matching techniques (e.g. Rosenbaum & Rubin, 1983). Another idea might be to examine the impact of resource loss and/or resource maintenance. For example, a study exploring the impact of an intervention focused on avoiding loss of strengths (instead of reducing psychiatric symptom severity or other outcomes) might be useful. Or, resource loss, resource gain, and resource maintenance groups could be explored for group differences over time.

Future studies might examine the relationship between youth strengths and psychiatric symptoms in a non-administrative dataset, and preferably with randomization procedures. Use of different measurement tools would allow for the collection of a wider variety of variables. For example, medication might contribute to a reduction in psychiatric symptom severity but is not tracked in the current dataset. Also, a different measurement tool would allow for more and different strengths items. This may be helpful in determining whether some resources are more important than others in particular contexts and individuals. It may also be useful to conduct analyses with extreme groups, since standard deviations were generally large in the analyses in the document, indicating large variation in the data.

Implications for Social Work and Others

This study has implications for the field of social work, though these must be considered in light of the limitations discussed above. Social workers are on the front-lines in public mental health systems, including Systems of Care, serving diverse, impoverished youth. This dissertation has demonstrated and attempted to articulate some of the strengths and limitations of

administrative data, which may be useful to social work practitioners, researchers, and others working in similar settings, or with similar datasets, in the future.

Also, with a few exceptions (e.g. Gambrill, 2013a,b; Thyer, 2005; Wong et al., 2004), behavioral theory and related research in experimental and applied settings is rarely mentioned in conjunction with approaches that involve strengths. On the other hand, many interventions that focus on strengths draw on behavioral theory. This dissertation attempted to articulate the contribution of behavioral theory and related empirical research to clinical work that focuses on strengths. Social workers interested in the role of the environment in shaping behavior may benefit from the decades of behavioral theoretical and empirical work in experimental and applied settings. This work illustrates how to use existing strengths, construct new ones, and use strengths to reduce distress and change behavior (e.g. Gambrill, 2013b; Goldiamond, 1974; Madden, 2013; Staats, 2012). The "strengths-based perspective in social work" has been identified as something unique that social workers bring to clinical work (Saleebey, 2009). Many social work schools teach some version of a strengths-based perspective in their curriculums. However, the emphasis on a strengths-based perspective in social work practice and education is generally value-based: behavioral theory and related research is typically ignored. However, social workers are ethically mandated to use the most up-to-date evidence to make informed decisions about how to intervene with clients. The contribution of unmeasured clinician variables to variation in psychiatric symptom change in these analyses (estimated at 30%) suggests that further attention to other clinician variables, including clinicians' use of a strengths-based orientation, might be a useful next step.

This dissertation also has implications for policymakers and researchers from other disciplines interested in health policy for youth and in social and behavior change more broadly. Psychosocial distress and misbehavior (deviant or socially disruptive behavior) have been attributed to biological, internally-derived disorders of the brain and body, while environmental and social factors have been downplayed (e.g. Conrad, 2007; Szasz, 2007). This dissertation explores a focus on strengths in Systems of Care as an alternative to the status quo - a focus on pathology.

References

Achenbach, T. M., & Rescorla, L. (2000). *Manual for the ASEBA preschool forms & profiles: An integrated system of multi-informant assessment*. Burlington, VT: ASEBA.

American Psychological Association (2013). *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5®)*. Washington, D.C: American Psychiatric Publishing.

American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders: DSM-IV-TR*. Washington, DC: American Psychiatric Association.

Anderson, R. L., & Estle, G. (January 01, 2001). Predicting level of mental health care among children served in a delivery system in a rural state. *The Journal of Rural Health, 17,* 3, 259-65.

Anderson, R. L., Lyons, J. S., Giles, D. M., Price, J. A., & Estle, G. (2003). Examining the reliability of the child and adolescent needs and strengths-mental health (CANS-MH) scale. *Journal of Child and Family Studies*, *12*, 1573–2843.

Anderson, J.A., Wright, E.R., Kelley, K., & Kooreman, H. (2006). Predicting functional improvement over time in a system of care. *Families in Society: The Journal of Contemporary Social Services, 87,* 438-446.

Anderson, J. A., Wright, E. R., Kelley, K., & Kooreman, H. (2008). Patterns of Clinical Functioning over Time for Young People Served in a System of Care. *Journal of Emotional and Behavioral Disorders, 16,* 90-104.

Allison, Paul D (1990). Change scores as dependent variables in regression analysis. *Sociological methodology*, 93-114.

Barth, R. P., Lee, B. R., Lindsey, M. A., Collins, K. S., Strieder, F., Chorpita, B. F., Becker, K. D., Sparks, J. A. (2012). Evidence-Based Practice at a Crossroads: The Timely Emergence of Common Elements and Common Factors. *Research on Social Work Practice, 22,* 108-119.

Bearslee, W. R., Wright, E. J., Gladstone, T. R., & Forbes, P. (2007). Long-term effects from a randomized trial of two public health preventive interventions for parental depression. *Journal of Family Psychology, 21,* 703-13.

Bertram, R. M., Suter, J. C., Bruns, J., & O'Rourke, K. E. (2011). Implementation research and wraparound literature: Building a research agenda. Journal of Child and Family Studies, 20, 713–725.

Bickman, L. (1996). A continuum of care: More is not always better. *American Psychologist*, *51*(7), 689.

Bickman L, Noser K, & Summerfelt WT (1999). Long-term effects of a system of care on children and adolescents. Journal of Behavioral Health Services Research, 26, 198–202.

Bickman, L., Lambert, E. W., Andrade, A. R., & Penaloza, R. V. (2000). The Fort Bragg continuum of care for children and adolescents: mental health outcomes over 5 years. *Journal of Consulting and Clinical Psychology*, *68*, 710.

Bohart, A. & Tallman, K. (2010). Clients: The neglected factor in psychotherapy. In Duncan, B., Miller, S., Wampold, B. & Hubble, M. (Eds), The heart and soul of change: Delivering what works in therapy (2nd ed.), pgs. 83-111.

Bruns, E. & Walker, J. (2008). *The resource guide to Wraparound.* Portland, OR: National Wraparound Initiative, Research and Training Center for Family Support and Children's Mental Health, Portland State University. Available from http://www.rtc.pdx.edu/NWI-book.

Buckley, J. A., & Epstein, M. H. (2004). The Behavioral and Emotional Rating Scale-2 (BERS-2): Providing a Comprehensive Approach to Strength-Based Assessment. *California School Psychologist, 9,* 21-27.

Buckley, J., Ryser, G., Reid, R., & Epstein, M. (2006). Confirmatory Factor Analysis of the Behavioral and Emotional Rating Scale–2 (BERS-2) Parent and Youth Rating Scales. Journal of Child and Family Studies, 15(1), 27–37.

Card, D., Chetty, R., Feldstein, M., & Saez, R. (2010). Expanding access to administrative data for research in the United States. National Science Foundation White Paper. Retrieved at on November 17th, 2014, at http://eml.berkeley.edu/~saez/card-chetty-feldstein-saezNSF10dataaccess.pdf.

Catalano, R. F., Berglund, M. L., Ryan, J. A. M., Lonczak, H. S., & Hawkins, J. D. (2004). Positive Youth Development in the United States: Research Findings on Evaluations of Positive Youth Development Programs. *The Annals of the American Academy of Political and Social Science, 591,* 98-124.

Center for Mental Health Services (U.S.). (2007). *Promotion and prevention in mental health: Strengthening parenting and enhancing child resilience : Report to Congress requested in Senate report 109-103 and Conference report 109-337*. Rockville, MD: U.S. Dept. of Health and Human Services, Substance Abuse and Mental Health Services Administration, Center for Mental Health Services.

Chor, K. H. B., McClelland, G. M., Weiner, D. A., Jordan, N., & Lyons, J. S. (2012). Predicting outcomes of children in residential treatment: A comparison of a decision support algorithm and a multidisciplinary team decision model. *Children and Youth Services Review, 34,* 2345-2352.

CMHI Annual Report (2011). The Comprehensive Community Mental Health Services for Children and Their Families Program Evaluation Findings. Rockville, MD: U.S. Dept. of Health and Human Services, Substance Abuse and Mental Health Services Administration, Center for Mental Health Services.

CMHI Annual Report (2011). The Comprehensive Community Mental Health Services for Children and Their Families Program Evaluation Findings. Rockville, MD: U.S. Dept. of Health and Human Services, Substance Abuse and Mental Health Services Administration, Center for Mental Health Services.

CMHI Annual Report (2005). The Comprehensive Community Mental Health Services for Children and Their Families Program Evaluation Findings. Rockville, MD: U.S. Dept. of Health and Human Services, Substance Abuse and Mental Health Services Administration, Center for Mental Health Services.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Conrad, P. (2007). *The medicalization of society: On the transformation of human conditions into treatable disorders*. Baltimore: Johns Hopkins University Press.

Corcoran, J., Dattalo, P., Crowley, M., Brown, E., & Grindle, L. (2011). A systematic review of psychosocial interventions for suicidal adolescents. *Children and Youth Services Review. 33*, 2112–2118.

Costello, E. J., Erkanli, A., Fairbank, J. A., & Angold, A. (2002). The prevalence of potentially traumatic events in childhood and adolescence. *Journal of Traumatic Stress, 15,* 99-112.

Cowen, E. L. (1994). The Enhancement of Psychological Wellness: Challenges and Opportunities. *American Journal of Community Psychology, 22,* 149.

Cox, K. (2006). Investigating the Impact of Strength-Based Assessment on Youth with Emotional or Behavioral Disorders. *Journal of Child and Family Studies, 15,* 278-292.
Cox, K. (2008). A roadmap for building on youth strengths. In E. J. Bruns & J. S. Walker (Eds.), The resource guide to wraparound. Portland, OR: National Wraparound Association.
Crews, S. D., Bender, H., Vanderwood, M., Cook, C. R., Gresham, F. M., & Kern, L. (2007). Risk and Protective Factors of Emotional and/or Behavioral Disorders in Children and Adolescents: A Mega-Analytic Synthesis. *Behavioral Disorders, 32,* 64-77.
Cress, C. J., Synhorst, L., Epstein, M. H., & Allen, E. (2012). Confirmatory Factor Analysis of the Preschool Behavioral and Emotional Rating Scale (PreBERS) with Preschool Children with Disabilities. *Assessment for Effective Intervention, 37,* 203-211.
Dolan, L. J., Kellam, S. G., Brown, C. H., Werthamer-Larsson, L., Rebok, G. W., Mayer, L. S., Laudolff, J., Wheeler, L. (1993). The Short-Term Impact of Two Classroom-Based Preventive Interventions on Aggressive and Shy Behaviors and Poor Achievement. *Journal of Applied Developmental Psychology, 14,* 317-346.
Drury, S. S., Theall, K., Gleason, M. M., Smyke, A. T., De, V. I., Wong, J. Y., Fox, N. A., ... Nelson, C. A. (2012). Telomere length and early severe social deprivation: linking early adversity and cellular aging. *Molecular Psychiatry, 17,* 719-27.
Dowdy, E., Chin, J. K., Twyford, J. M., & Dever, B. V. (2011). A factor analytic investigation of the BASC-2 Behavioral and Emotional Screening System Parent Form: Psychometric properties, practical implications, and future directions. *Journal of School Psychology*, 49, 265–280.
Duncan, B., Miller, S., Wampold, B. & Hubble, M. (2010). The heart and soul of change: Delivering what works in therapy (2nd ed.), Washington, D.C.: American Psychological Association.
Dunleavy, A. M., & Leon, S. C. (2011). Predictors for resolution of antisocial behavior among foster care youth receiving community-based services. *Children and Youth Services Review, 33,* 2347-2354.
Epstein, M. H. (2004). *Behavioral and Emotional Rating Scale-2nd Edition: A strengths-based approach to assessment*. Austin, TX: PRO-ED.
Epstein, M. H., Mooney, P., Ryser, G., & Pierce, C. D. (2004). Validity and Reliability of the Behavioral and Emotional Rating Scale (2nd Edition): Youth Rating Scale. *Research on Social Work Practice*, *14*, 358–367.
Fergus, S., & Zimmerman, M. A. (2005). Adolescent resilience: a framework for understanding healthy development in the face of risk. *Annual Review of Public Health, 26,* 399-419.
Finkel, S. E. (1995). Causal analysis with panel data. Thousand Oaks, CA: Sage.
Fisher, W., Groff, R., & Roane, H. (2011). Applied behavioral analysis: History, philosophy, principles, and basic methods. In W. Fisher, C. Piazza, & H. Roane (2011). *Handbook of applied behavior analysis*. New York: Guilford Press.
Foster, E. M., Stephens, R., Krivelyova, A., & Gamfi, P. (2007). Can system integration improve mental health outcomes for children and youth? *Children and Youth Services Review, 29,* 1301-1319.
Frances, A. (2010). First draft of the DSM-V: If accepted, will fan the flames of false positives. *British Medical Journal,* 340, 1168.
Franz, J. (2008). ADMIRE: Getting practical about being strengths-based. In E. J. Bruns & J. S. Walker (Eds.), The resource guide to wraparound. Portland, OR: National Wraparound

Initiative, Research and Training Center for Family Support and Children's Mental Health, Portland State University. Available from http://www.rtc.pdx.edu/NWI-book.
Fraser, M. W., Galinsky, M. J., & Richman, J. M. (1999). Risk, protection, and resilience: Toward a conceptual framework for social work practice. *Social Work Research*, *23*, 131–143.
Fraser, M. (2011). Child mental health policy. In J. Jenson & M. Fraser (Eds.), *Social policy for children and families: A risk and resilience perspective*. Thousand Oaks, Calif: SAGE.
Friedman, R. & Israel, N. (2008). Research and evaluation implications. In B. Stroul & G. Blau (Eds.), *The system of care handbook: Transforming mental health services for children, youth, and families*. Baltimore, Md: Paul H. Brookes Pub. Co, pgs. 689-705.
Furedi, F. (2006). *Culture of fear revisited: Risk-taking and the morality of low expectation*. London: Continuum Development.
Gambrill, E. D. (2005). *Critical thinking in clinical practice: Improving the quality of judgments and decisions*. Hoboken, N.J: Wiley
Gambrill, E. (2013a). Birds of a Feather: Applied Behavior Analysis and Quality of Life. *Research on Social Work Practice, 23,* 2, 121-140.
Gambrill, E. D. (2013b). *Social work practice: A critical thinker's guide*. New York: Oxford University Press.
Garcia, J. A., & Weisz, J. R. (2002). When youth mental health care stops: therapeutic relationship problems and other reasons for ending youth outpatient treatment. *Journal of Consulting and Clinical Psychology, 70,* 439-43.
Garland, A. F., Bickman, L., & Chorpita, B. F. (2010). Change what? Identifying quality improvement targets by investigating usual mental health care. *Administration and Policy in Mental Health and Mental Health Services Research*, *37*, 15-26.
Garrett, E., Barnes, H., & Dibben, C. (2010). Health administrative data: Exploring potential for academic research. St. Andrews: Administrative Data Liaison Service.
Goldiamond, I. (1974). Toward a constructional approach to social problems: ethical and constitutional issues raised by applied behavior analysis. *Behaviorism, 2,* 1-84.
Gomory, T., Wong, S. E., Cohen, D., & Lacasse, J. R. (2011). Clinical social work and the biomedical industrial complex. *Journal of Sociology and Social Welfare, 38,* 135-165.
Graves, K. N., & Shelton, T. L. (2007). Family empowerment as a mediator between family-centered systems of care and changes in child functioning: Identifying an important mechanism of change. *Journal of Child and Family Studies*, *16*, 556-566.
Griffin, G., Martinovich, Z., Gawron, T., & Lyons, J. (2009). Strengths moderate the impact of trauma on risk behaviors in child welfare. *Residential Treatment for Children and Youth, 26,* 105-118.
Harrington, D. (2009). *Confirmatory Factor Analysis.* New York, New York: Oxford University Press.
Hu, L.-t. & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covatiance structur analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1-55.
Hawkins, J. D., Kosterman, R., Catalano, R. F., Hill, K. G., & Abbott, R. D. (2008). Effects of Social development intervention in child-hood fifteen years later. A*rchives of Pediatrics & Adolescent Medicine*, 162, 1133.
Hernandez, M., & Hodges, S. (2003). Building upon the Theory of Change for Systems of Care. *Journal of Emotional and Behavioral Disorders, 11,* 19-26.

Hoagwood, K. E. (2013). Don't mourn: Organize. reviving mental health services research for healthcare quality improvement. *Clinical Psychology: Science and Practice, 20*, 120-126.

Hoagwood, K., & Kolko, D. (2009). Introduction to the Special Section on Practice Contexts: A Glimpse into the Nether World of Public Mental Health Services for Children and Families. Administration and Policy in Mental Health and Mental Health Services Research, 36, 35–36.

Hodges, S., Ferreira, K., Israel, N., Mazza, J., & Systems of Care. (2010). Systems of care, featherless bipeds, and the measure of all things. *Evaluation and Program Planning, 33,* 4-10.

Hodges, S., & Ferreira, K. (2010). Special issue: Systems of care. *Evaluation and Program Planning, 33*.

Horner, R. H., Sugai, G., & Anderson, C. M. (2010). Examining the Evidence Base for School-Wide Positive Behavior Support. *Focus on Exceptional Children, 42,* 8.

Houts, A. (2006). Behavioral Treatment for Enuresis. In A. Kazdin & J. Weisz (Eds.), *Evidence-based psychotherapies for children and adolescents,* pgs. 389-406.

Illouz, E. (2008). Saving the modern soul: Therapy, emotions, and the culture of self-help. Berkeley: University of California Press.

Insel, T. (2013). Blog post. Retrieved on December 5th, 2014, from http://www.nimh.nih.gov/about/director/2013/transforming-diagnosis.shtml.

Jenson, J. M., & Fraser, M. W. (2011). *Social policy for children and families: A risk and resilience perspective*. Thousand Oaks, CA: SAGE.

Jenson, J. M., Alter, C., Nicotera, N., Anthony, E., Forrest-Bank, S. (2013). *Risk, resilience, and positive youth development: Developing effective community programs for at-risk youth : lessons from the Denver Bridge Project*. New York: Oxford University Press.

Johnson, Richard, and Wichern, (2002). *Applied Linear Statistical Models, 5th edition,* Prentice-Hall.

Jörg, F., Ormel, J., Reijneveld, S. A., Jansen, D. E., Verhulst, F. C., & Oldehinkel, A. J. (2012). Puzzling findings in studying the outcome of "real world" adolescent mental health services: The TRAILS study. *PloS one*, *7*, e44704.

Kahn, J. H., & Schneider, W. J. (2013). It's the Destination and It's the Journey: Using Multilevel Modeling to Assess Patterns of Change in Psychotherapy. *Journal of clinical psychology*, *69*(6), 543-570.

Kam, C.-M., Greenberg, M. T., & Kusché, C. A. (2004). Sustained Effects of the PATHS Curriculum on the Social and Psychological Adjustment of Children in Special Education. *Journal of Emotional and Behavioral Disorders*, *12*, 66–78.

Kellam, S. G., Brown, C. H., Poduska, J. M., Ialongo, N. S., Wang, W., Toyinbo, P., Petras, H., Wilcox, H. C. (2008). Effects of a universal classroom behavior management program in first and second grades on young adult behavioral, psychiatric, and social outcomes. *Drug and Alcohol Dependence, 95.*

Kelley, S., Bickman, L., & Norwood, E. (2010). Evidence-based treatments and common factors in youth psychotherapy. In Duncan, B., Miller, S., Wampold, B. & Hubble, M. (Eds), The heart and soul of change: Delivering what works in therapy (2nd ed.), (pp. 325-355). Washington, DC, US: American Psychological Association,.

Kessler R.C, Gruber M.J, Petukhova M, Sampson N.A, Zaslavsky A.M, McLaughlin K.A, … Costello J. (2012). Severity of 12-month DSM-IV disorders in the National Comorbidity

Survey Replication Adolescent Supplement. *Arch. Gen. Psychiatry Archives of General Psychiatry*, *69*, 381–389.
Keyes, C. L. (2007). Promoting and protecting mental health as flourishing: a complementary strategy for improving national mental health. *The American Psychologist, 62,* 2.
Kim, E. K., Furlong, M. J., Dowdy, E., & Felix, E. D. (2014). Exploring the Relative Contributions of the Strength and Distress Components of Dual-Factor Complete Mental Health Screening. *Canadian Journal of School Psychology, 29,* 127-140.
Kim, D., Wampold, B. E., & Bolt, D. M. (2006). Therapist effects in psychotherapy: A random-effects modeling of the National Institute of Mental Health Treatment of Depression Collaborative Research Program data. Psychotherapy Research, 16, 161−172.
Kirk, S. A. (2005). *Mental disorders in the social environment: Critical perspectives*. New York: Columbia University Press.
Kirk, S. A., Gomory, T., & Cohen, D. (2013). *Mad science: Psychiatric coercion, diagnosis, and drugs*. Transaction Publishers.
Kline, R.B. (2005), Principles and Practice of Structural Equation Modeling (2nd Edition ed.). New York: The Guilford Press.
Knitzer, J., & Olson, L. (1982). *Unclaimed children: The failure of public responsibility to children and adolescents in need of mental health services*. Washington, D.C: Children's Defense Fund.
Koren, P. E., DeChillo, N., & Friesen, B. J. (1992). Measuring empowerment in families whose children have emotional disabilities: A brief questionnaire. *Rehabilitation Psychology*, *37*, 305–321.
Kutchins, H., & Kirk, S. A. (1997). *Making us crazy: DSM : the psychiatric bible and the creation of mental disorders*. New York: Free Press.
Kwok, O. M., Underhill, A. T., Berry, J. W., Luo, W., Elliott, T. R., & Yoon, M. (2008). Analyzing longitudinal data with multilevel models: An example with individuals living with lower extremity intra-articular fractures. *Rehabilitation Psychology*, *53*, 370.
Lacasse, J.R., & Gomory, T. (2003). Is graduate social work education promoting a critical approach to mental health practice? *Journal of Social Work Education*, *39*, 383–408.
Lambert, W., & Bickman, L. (2004). The ‘‘clock-setting’’ cure: How children’s symptoms might improve after ineffective treatment. Psychiatric Services, 55, 381–382.
Layng, T. V. J. (2009). The Search for an Effective Clinical Behavior Analysis: The Nonlinear Thinking of Israel Goldiamond. *Behavior Analyst, 32,* 1, 163-184.
LeBuffe, P. A., Shapiro, V. B., & Naglieri, J. A. (2009). *The Devereux Student Strengths Assessment (DESSA) assessment*. Lewisville, NC: Kaplan.
Liu, W. M. (2011). *Social class and classism in the helping professions: Research, theory, and practice*. Thousand Oaks: SAGE.
Lourie, I. S., & Hernandez, M. (2003). A Historical Perspective on National Child Mental Health Policy. *Journal of Emotional and Behavioral Disorders*, *11*, 5–9.
Luthar, S. S. (1991). Vulnerability and Resilience: A Study of High-Risk Adolescents. *Child Development, 62,* 3, 600-616.

Luthar, S. S., Cicchetti, D., & Becker, B. (2000). The Construct of Resilience: A Critical Evaluation and Guidelines for Future Work. *Child Development, 71,* 543-562.

Luthar, S. S., & Brown, P. J. (2007). Maximizing resilience through diverse levels of inquiry: Prevailing paradigms, possibilities, and priorities for the future. *Development and Psychopathology, 19,* 931-955.

Lutz, W., Leon, S. C., Martinovich, Z., Lyons, J. S., & Stiles, W. B. (2007). Therapist Effects in Outpatient Psychotherapy: A Three-Level Growth Curve Approach. *Journal of Counseling Psychology, 54,* 32-39.

Lyons, J. S. (2004). *Redressing the emperor: Improving our children's public mental health system*. Westport, CT: Praeger.

Lyons, J. (2009). *Communimetrics: A theory of measurement for human service enterprises.* New York: Springer.

Lyons, J. S., Griffin, G., Quintenz, S., Jenuwine, M., & Shasha, M. (2003). Clinical and forensic outcomes from the Illinois mental health juvenile justice initiative. *Psychiatric Services, 54,* 1629-34.

Lyons, J. S., MacIntyre, J. C., Lee, M. E., Carpinello, S., Zuber, M. P., & Fazio, M. L. (2004). Psychotropic medications prescribing patterns for children and adolescents in New York's public mental health system. *Community Mental Health Journal, 40,* 101-18.

Lyons, J., Uziel-Miller, N., Reyes, F., & Sokol, P. (2000). Strengths of Children and Adolescents in Residential Settings: Prevalence and Associations with Psychopathology and Discharge Placement. *Journal of the American Academy of Children and Adolescent Psychiatry, 39,* 176-181.

Madden, G. J. (2013). *APA handbook of behavior analysis*. Washington, DC: American Psychological Association.

Manteuffel, B., Stephens, R. L., & Santiago, R. (2002). Overview of the national evaluation of the Comprehensive Community Mental Health Services for Children and Their Families Program and summary of current findings. *Children's Services: Social Policy, Research, and Practice*, *5*, 3-20.

Masten, A. S. (2001). Ordinary magic. Resilience processes in development. *The American Psychologist, 56,* 227-38.

Masten, A. S. (2007). Resilience in developing systems: Progress and promise as the fourth wave rises. *Development and Psychopathology, 19,* 921-930.

McCammon, S. L. (2012). Systems of care as asset-building communities: Implementing strengths-based planning and positive youth development. *American Journal of Community Psychology*, *49*, 556-565.

McCroskey, J., & Meezan, W. (1997). *Family preservation & family functioning*. Washington, DC: CWLA Press.

Merikangas, K. R., He, J., Burstein, M., Swendsen, J., Avenevoli, S., Case, B., & Olfson, M. (2011). Service Utilization for Lifetime Mental Disorders in U.S. Adolescents: Results of the National Comorbidity Survey–Adolescent Supplement (NCS-A). *Journal of the American Academy of Child & Adolescent Psychiatry*, *50*, 32–45.

Miles, J., Espiritu, R. C., Horen, N. M., Sebian, J., & Waetzig, E.(2010). *A public health approach to children's mental health: A conceptual framework: expanded executive summary.* Washington, DC: Georgetown University Center for Child and Human Development.

National Wraparound Initiative. (2014). Retrieved on November 20th, 2014, from: www.nwi.pdx.edu/wraparoundbasics.shtml.

O'Connell, M. E., Boat, T. F., Warner, K. E., & National Research Council (Institute on Medicine) (2009). *Preventing mental, emotional, and behavioral disorders among young people: Progress and possibilities*. Washington, D.C: National Academies Press.

Ozer, E. (2005). The Impact of Violence on Urban Adolescents. *Journal of Adolescent Research, 20,* 167-192

President's New Freedom Commission on Mental Health (2003). *Achieving the promise: Transforming mental health care in America: executive summary: final report*. Rockville, MD: United States Government.

Painter, K. (2012). Outcomes for youth with severe emotional disturbance: A repeated measures longitudinal study of a wraparound approach of service delivery in systems of care. *Child & Youth Care Forum, 41*, 407-425.

Praed Foundation (2014). Retrieved on November 20th, 2014, from http://www.praedfoundation.org/About%20the%20Foundation.html.

Prinz, R. J., Sanders, M. R., Shapiro, C. J., Whitaker, D. J., & Lutzker, J. R. (2009). Population-based prevention of child maltreatment: the U.S. Triple p system population trial. *Prevention Science, 10,* 1-12.

Probst, B. (2009). Contextual meanings of the strengths perspective for social work practice in mental health. *Families In Society*, *90*, 162-166.

Pynoos, R.S., Steinberg, A.M., Layne, C.M., Briggs, E.C., Ostrowski, S.A., & Fairbank, J. (2009). DSM-V PTSD diagnostic criteria for children and adolescents: A developmental perspective and recommendations. *Journal of Traumatic Stress, 22*, 391–398.

Rabe-Hesketh, S., & Skrondal, A. (2012). *Multilevel and longitudinal modeling using Stata*. STATA press.

Radigan, M., & Wang, R. (2013). Relationships Between Youth and Caregiver Strengths and Mental Health Outcomes in Community Based Public Mental Health Services. *Community mental health journal*, *49*(5), 499-506.

Rapley, M., Moncrieff, J., & Dillon, J. (2011). Carving nature at its joints? DSM and medicalization of everyday life. In M. Rapley, M. Moncrieff, & J. Dillon (Eds.), *Demedicalization of Misery: Psychology, Psychiatry, and the Human Condition* (pgs. 1-9). New York: Macmillan.

Reid, R., Epstein, M. H., Pastor, D. A., Ryser, G. R. (2000). Strengths-based assessment differences across students with LD and EBD. *Remedial and Special Education, 21*, 346-355.

Rosenbaum, Paul R. and Donald B. Rubin. 1983. The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*, 70, 41-55.

Saleebey, D. (2009). *The strengths perspective in social work practice*. Boston: Allyn & Bacon.

Salzer, M. S., & Bickman, L. (1997). Delivering Effective Children's Services in the Community: Reconsidering the Benefits of System Interventions. *Applied and Preventive Psychology, 6,* 1-14.

San Francisco CANS (5-18) User Manual (2010). San Francisco Children's System of Care. San Francisco Department of Public Health: San Francisco, CA.

Scheff, T. J. (1999). *Being mentally ill: A sociological theory*. New York: Aldine de Gruyter.

Schwartz, A., & Goldiamond, I. (1975). *Social casework: A behavioral approach*. New York: Columbia University Press.

Sieracki, J. H., Leon, S. C., Miller, S. A., & Lyons, J. S. (2008). Individual and

provider effects on mental health outcomes in child welfare: A three level growth curve approach. *Children and Youth Services Review, 30,* 800-808.
Smith, T. (2010). Early and Intensive Behavioral Intervention in Autism. In Weisz, J.R. (Ed.), *Evidence-Based Psychotherapies for Children and Adolescents*. New York: Guilford Publications, pgs. 312-326.
Snyder, C. R., Ritschel, L. A., Rand, K. L., & Berg, C. J. (2006). Balancing psychological assessments: Including strengths and hope in client reports. *Journal of Clinical Psychology*, *62*, 33–46.
Snyder, F. J., Roberts, Y. H., Crusto, C. A., Connell, C. M., Griffin, A., Finley, M. K., ... & Kaufman, J. S. (2012). Exposure to traumatic events and the behavioral health of children enrolled in an early childhood system of care. *Journal of traumatic stress*, *25*(6), 700-704.
Staats, A. W. (2012). *The marvelous learning animal: What makes human nature unique*. Amherst, N.Y: Prometheus Books.
StataCorp. 2013. Stata: Release 13. Statistical Software User Manual. College Station, TX: StataCorp LP.
Stephens, R. L., Holden, E. W., & Hernandez, M. (2004). System of Care Practice Review Scores as Predictors of Behavioral Symptomatology and Functional Impairment. *Journal of Child and Family Studies, 13,* 179-191.
Stroul, B.A., & Friedman, R.M. (1986). *A system of care for seriously emotionally disturbed children and youth: What are the results?* Washington, DC: Georgetown University Child Development Center.
Stroul, B.A., & Friedman, R.M. (1994). *A system of care for seriously emotionally disturbed children and youth: What are the results?* Washington, DC: Georgetown University Child Development Center.
Stroul, B. & Blau, G. (2008) *The system of care handbook: Transforming mental health services for children, youth, and families*. Baltimore, Md: Paul H. Brookes Pub. Co.
Stroul, B. A., & Blau, G. M. (2010). Defining the system of care concept and philosophy: To update or not to update? *Evaluation and Program Planning, 33,* 59-62.
Stroul, B., Blau, G. & Friedman, R. (2010). *Updating the system of care concept and philosophy.* Washington, DC: Georgetown University Center for Child and Human Development and Substance Abuse and Mental Health Services Administration.
Stroul, B., Blau, G., & Sondheimer, D. (2008). Systems of Care: A strategy to transform children's mental health care. In B. Stroul & G. Blau (Eds.), *The system of care handbook: Transforming mental health services for children, youth, and families*. Baltimore, Md: Paul H. Brookes Pub. Co, pgs. 3-23.
Suter, J. C., & Bruns, E. J. (2009). Effectiveness of the wraparound process for children with emotional and behavioral disorders: a meta-analysis. *Clinical Child and Family Psychology Review, 12,* 336-51.
Suter, J. C., & Bruns, E. J. (2008). Narrative review of wraparound outcome studies. In E. J. Bruns & J. S. Walker (Eds.), Theresource guide to wraparound. Portland, OR: National Wraparound Initiative, Research and Training Center for Family Support and Children's Mental Health, Portland State University. Available from http://www.rtc.pdx.edu/NWI-book.
Surgeon General's Conference on Children's Mental Health: Developing a National Action Agenda. 2000. *Report of the Surgeon General's Conference on Children's Mental Health*

a national action agenda. [Washington, D.C.]: Dept. of Health and Human Services, Public Health Service. http://purl.access.gpo.gov/GPO/LPS17397.
Szasz, Thomas S. *The myth of mental illness: Foundations of a theory of personal conduct*. HarperPerennial, 1974.
Szasz, T. S. (1974). *The myth of mental illness: foundations of a theory of personal conduct*. New York: Harper & Row.
Szasz, T. (1987). *Insanity: The idea and its consequences*. New York: Wiley.
Szasz, T. (2001). *Pharmacracy: Medicine and politics in America*. Westport, Conn: Praeger.
Szasz, T. (2007). *The medicalization of everyday life: Selected essays*. Syracuse, N.Y: Syracuse University Press.
TAP (Technical Assistance Partnership for Children and Family Mental Health), a program of Substance Abuse and Mental Health Services Administration). Systems of Care. Retrieved on November 19th, 2014: http://www.tapartnership.org/systemsOfCare.php.
Tedeschi, R. G., & Kilmer, R. P. (2005). Assessing Strengths, Resilience, and Growth to Guide Clinical Interventions. *Professional Psychology: Research and Practice*, *36*, 230–237.
Thyer, B. (2005). The misfortunes of behavioral social work: misprized, misread and misconstrued. *Mental disorders in the social environment: Critical perspectives*. New York: Columbia University Press.
Trask, E. & Garland, A. (2012). Are Children Improving? Results from Outcome Measurement in a Large Mental Health System. *Administration and Policy in Mental Health and Mental Health Services Research 39,* 210-220
Toche-Manley, L., Nankin, M., & Dietzen, L. (2004). Using Outcome Data to Create Clinical Change: Foundations and Example. *Child and Youth Care Forum, 33,* 51-67.
U.S. Department of Health and Human Services. Mental Health: A Report of the Surgeon General—Executive Summary. Rockville, MD: U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Center for Mental Health Services, National Institutes of Health, National Institute of Mental Health, 1999.
Wampold, B. E., & Brown, G. S. (2005). Estimating variability in outcomes attributable to therapists: a naturalistic study of outcomes in managed care. *Journal of Consulting and Clinical Psychology, 73,* 914-23.
Wampold, B. (2010). The research evidence for the common factors model: a historical perspective. In B Duncan (Ed.), *The heart & soul of change: Delivering what works in therapy*. Washington, DC: American Psychological Association.
Walrath, C. M., Ybarra, M. L., Sheehan, A. K., Holden, E. W., & Burns, B. J. (2006). Impact of Maltreatment on Children Served in Community Mental Health Programs. *Journal of Emotional and Behavioral Disorders, 14,* 143-156.
Wraparound Initiative, Research and Training Center for Family Support and Children's Mental Health, Portland State University. Available from http://www.rtc.pdx.edu/NWI-book.
Werner, E. & Smith, R. (1992). *Overcoming the odds: High risk children from birth to adulthood.* Ithaca: Cornell University Press.
Weisz, J. R., Sandier, I. N., Durlak, J. A., & Anton, B. S. (2005). Promoting and Protecting Youth Mental Health Through Evidence-Based Prevention and Treatment. *American Psychologist, 60,* 6.
Whitson, M. L., Connell, C. M., Bernard, S., & Kaufman, J. S. (2012). An

Examination of Exposure to Traumatic Events and Symptoms and Strengths for Children Served in a Behavioral Health System of Care. *Journal of Emotional and Behavioral Disorders, 20,* 193-207.

Wong, S.E., Wilder, D.A., Schock, K., & Clay, C. (2004). Behavioral interventions with severe and persistent mental disorders. In H. Briggs & T. Rzepnicki (Eds.). *Using evidence in social work practice: Behavioral perspectives*. Chicago: Lyceum Books.

World Health Organization (2014). Definition of mental health. Retrieved from http://www.who.int/features/factfiles/mental_health/en/.

Wyman, P. Sandler, I., Wolchik, S., & Nelson, K. (2000). Resilience as cumulative competence promotion. In D. Cicchetti, J. Rappaport, I. Sanders, & R. Weissberg (Eds.), *The promotion of wellness in children and adolescents*. Washington, DC: CWLA Press.

Appendix A

The *San Francisco Child and Adolescent Needs and Strengths (5-18) Initial Assessment Scoring Manual (2010)* was in use in San Francisco's Children's System of Care during the time the data for this dissertation was collected. The scoring directions for the six strengths items (pgs. 51-56) are excerpted below*:*

FAMILY

Family refers to all biological or adoptive relatives with whom the child or youth remains in contact along with other individuals in relationships with these relatives.

0 Significant family strengths. This level indicates a family with much love and mutual respect for each other. Family members are central in each other's lives. Child is fully included in family activities.

1 Moderate level of family strengths. This level indicates a loving family with generally good communication and ability to enjoy each other's company. There may be some problems between family members. Child is generally included.

2 Mild level of family strengths. Family is able to communicate and participate in each other's lives; however, family members may not be able to provide significant emotional or concrete support for each other. Child is often not included in family activities.

3 This level indicates a child with no known family strengths. Child is not included in normal family activities.

Section 4: Scoring Manuals 51

INTERPERSONAL

This rating refers to the interpersonal skills of the child or youth both with peers and adults.

0 Significant interpersonal strengths. Child is seen as well liked by others and has significant ability to form and maintain positive relationships with both peers and adults. Individual has multiple close friends and is friendly with others.

1 Moderate level of interpersonal strengths. Child has formed positive interpersonal relationships with peers and/or other non-caregivers. Child may have one friend, if that friendship is a healthy „best friendship model.

2 Mild level of interpersonal strengths. Child has some social skills that facilitate positive relationships with peers and adults but may not have any

current relationships, but has a history of making and maintaining healthy friendships with others.

3 This level indicates a child with no known interpersonal strengths. Child currently does not have any friends nor has he/she had any friends in the past. Child does not have positive relationships with adults.

EDUCATIONAL

This rating refers to the strengths of the school system and may or may not reflect any specific educational skills possessed by the child or youth.

0 This level indicates a child who is in school and is involved with an educational plan that appears to exceed expectations. School works exceptionally well with family and caregivers to create a special learning environment. A child in a mainstream educational system who does not require an individual plan would be rated here.

1 This level indicates a child who is in school and has a plan that appears to be effective. School works fairly well with family and caregivers to ensure appropriate educational development.

2 This level indicates a child who is in school but has a plan that does not appear to be effective.

3 This level indicates a child who is either not in school or is in a school setting that does not further his/her education.

TALENTS / INTERESTS

This rating should be based broadly on any talent, creative or artistic skill a child or adolescent may have including art, theatre, music, athletics, etc.

0 This level indicates a child with significant creative/artistic strengths. A child/youth who receives a significant amount of personal benefit from activities surrounding a talent would be rated here.

1 This level indicates a child with a notable talent. For example, a youth who is involved in athletics or plays a musical instrument, etc. would be rated here.

2 This level indicates a child who has expressed interest in developing a specific talent or talents even if they have not developed that talent to date.

3 This level indicates a child with no known talents, interests, or hobbies.

SPIRITUAL / RELIGIOUS

This rating should be based on the child or adolescent's and their family's involvement in spiritual or religious beliefs and activities.

0 This level indicates a child with strong moral and spiritual strengths. Child may be very involved in a religious community or may have strongly held spiritual or religious beliefs that can sustain or comfort him/her in difficult times.

1 This level indicates a child with some moral and spiritual strengths. Child may be involved in a religious community.

2 This level indicates a child with few spiritual or religious strengths. Child may have little contact with religious institutions.

3 This level indicates a child with no known spiritual or religious involvement.

RELATIONSHIP PERMANENCE

This rating refers to the stability of significant relationships in the child or youth's life. This likely includes family members but may also include other individuals.

0 This level indicates a child who has very stable relationships. Family members, friends, and community have been stable for most of his/her life and are likely to remain so in the foreseeable future. Child is involved with both parents.

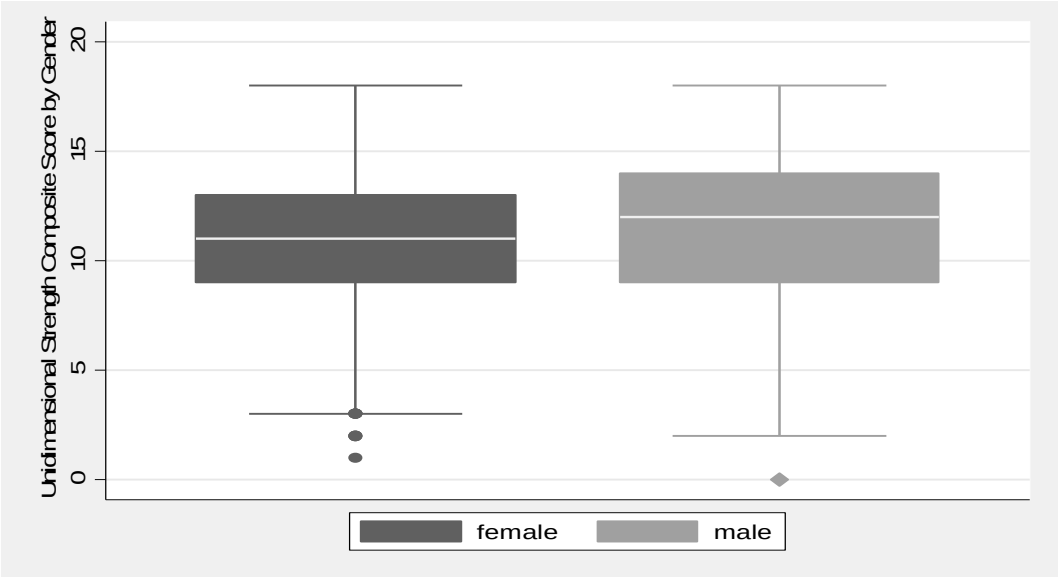
1 This level indicates a child who has had stable relationships but there is some concern about instability in the near future (one year) due to transitions, illness, or age. A child who has a stable relationship with only one parent may be rated here.

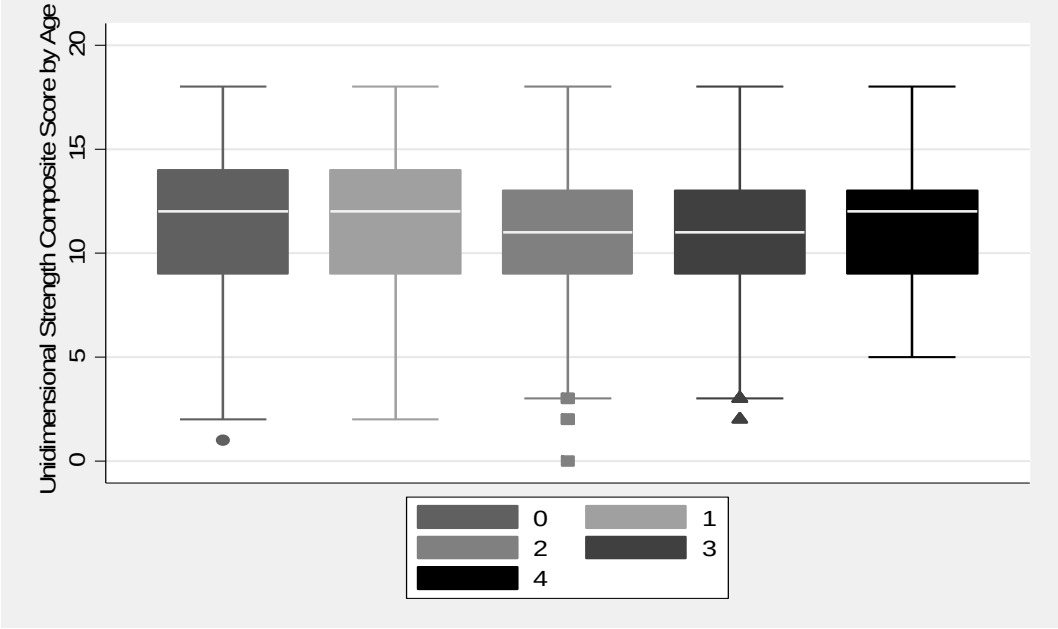
2 This level indicates a child who has had at least one stable relationship over his/her lifetime but has experienced other instability through factors such as divorce, moving, removal from home, and death.

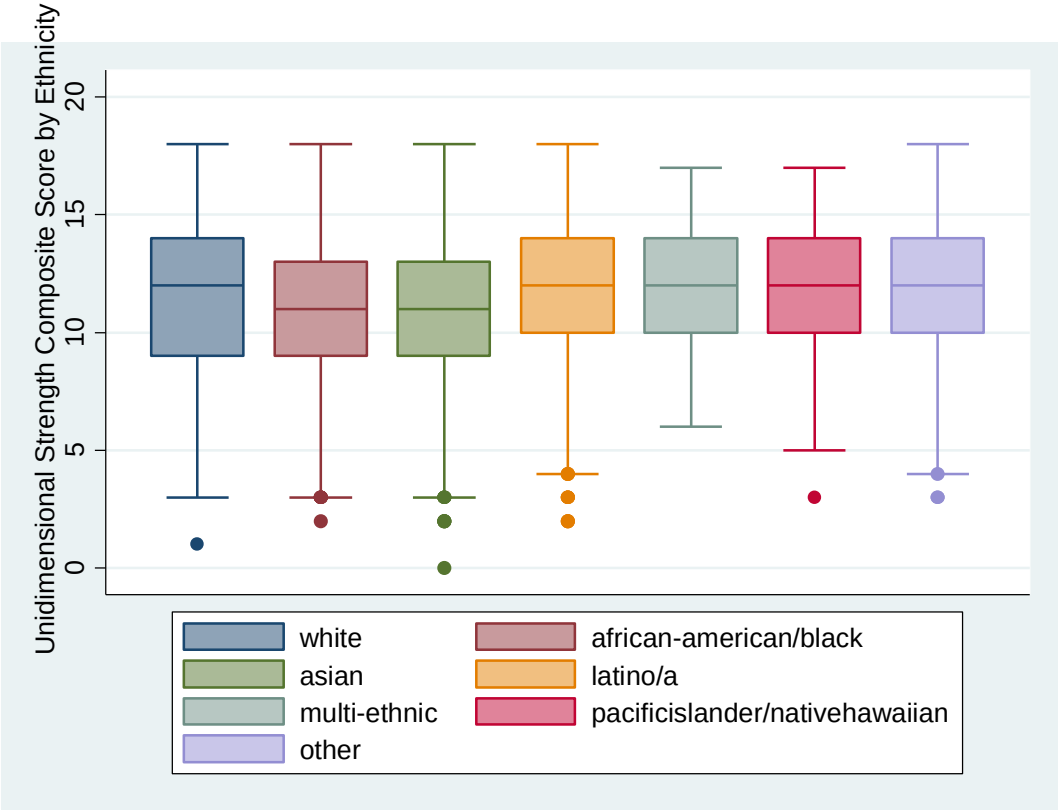
3 This level indicates a child who does not have any stability in relationships.

Appendix B

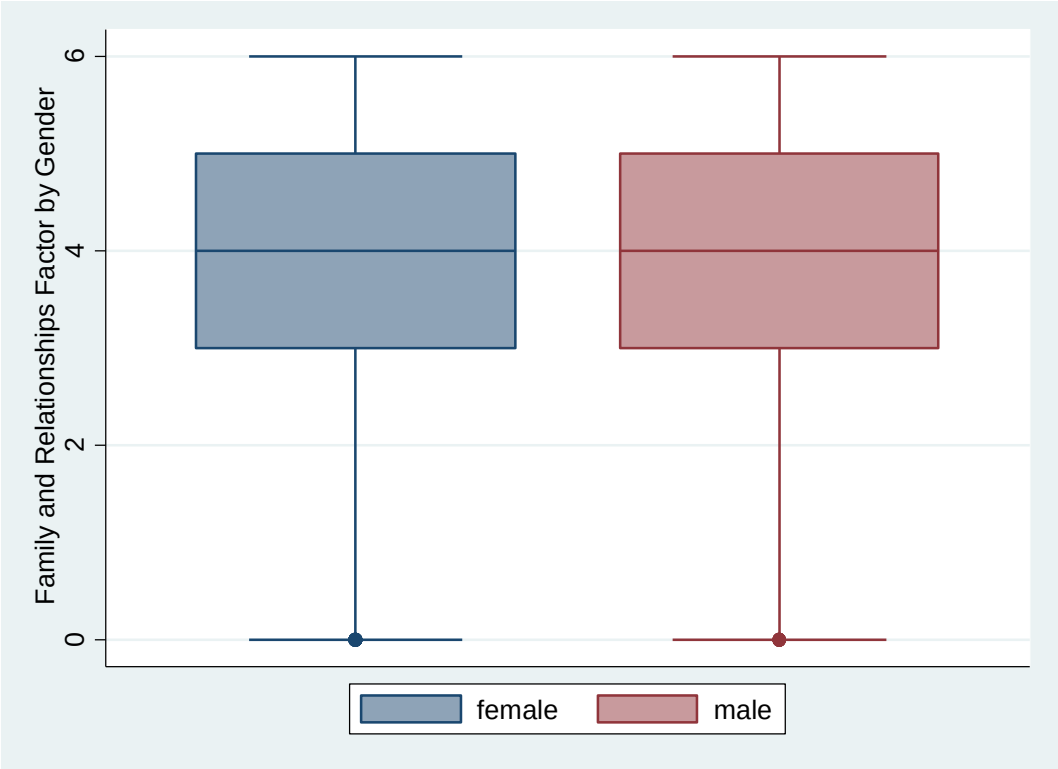
Box plots for Strengths Composite Score:

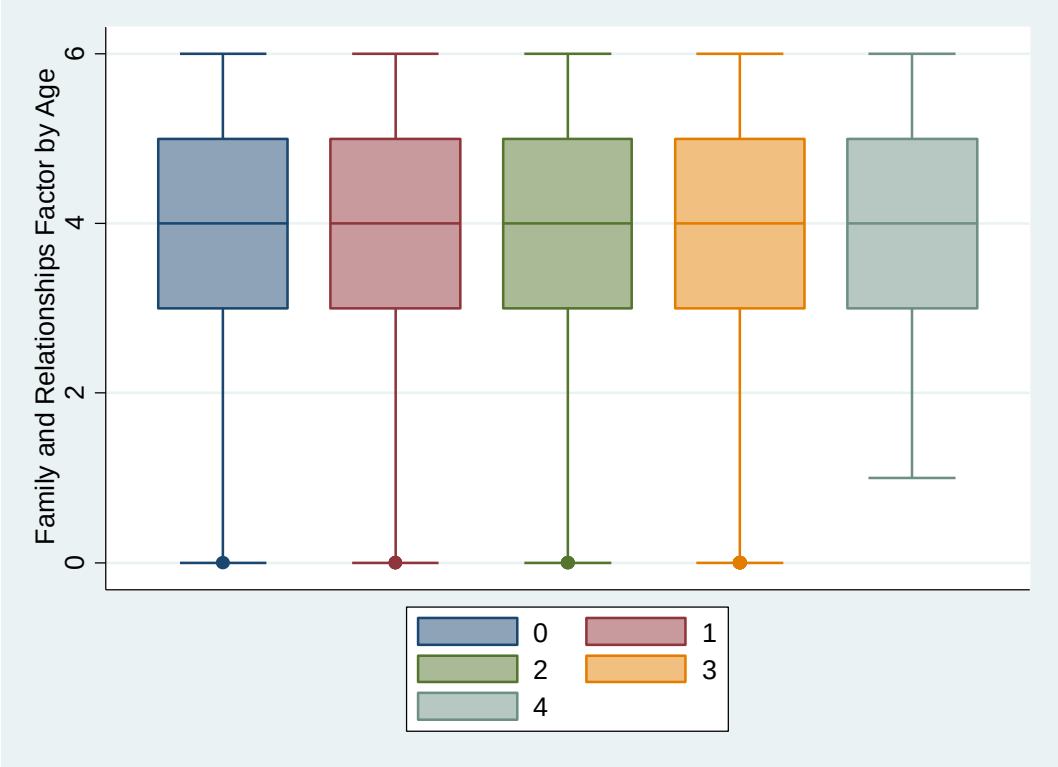


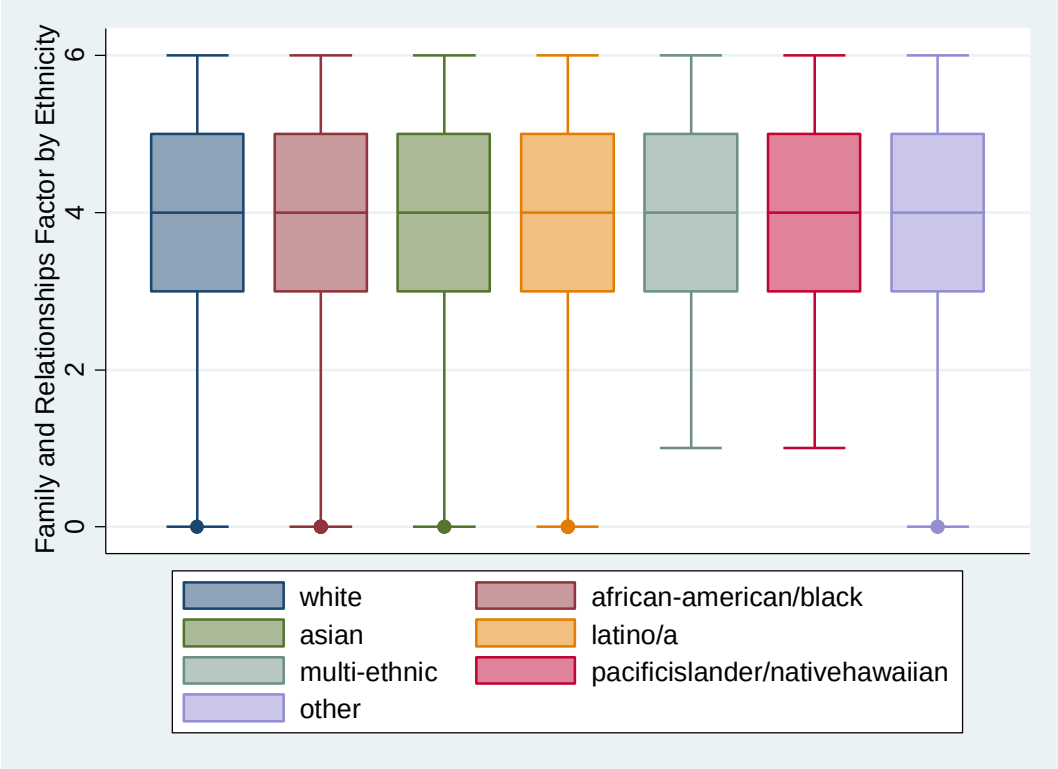


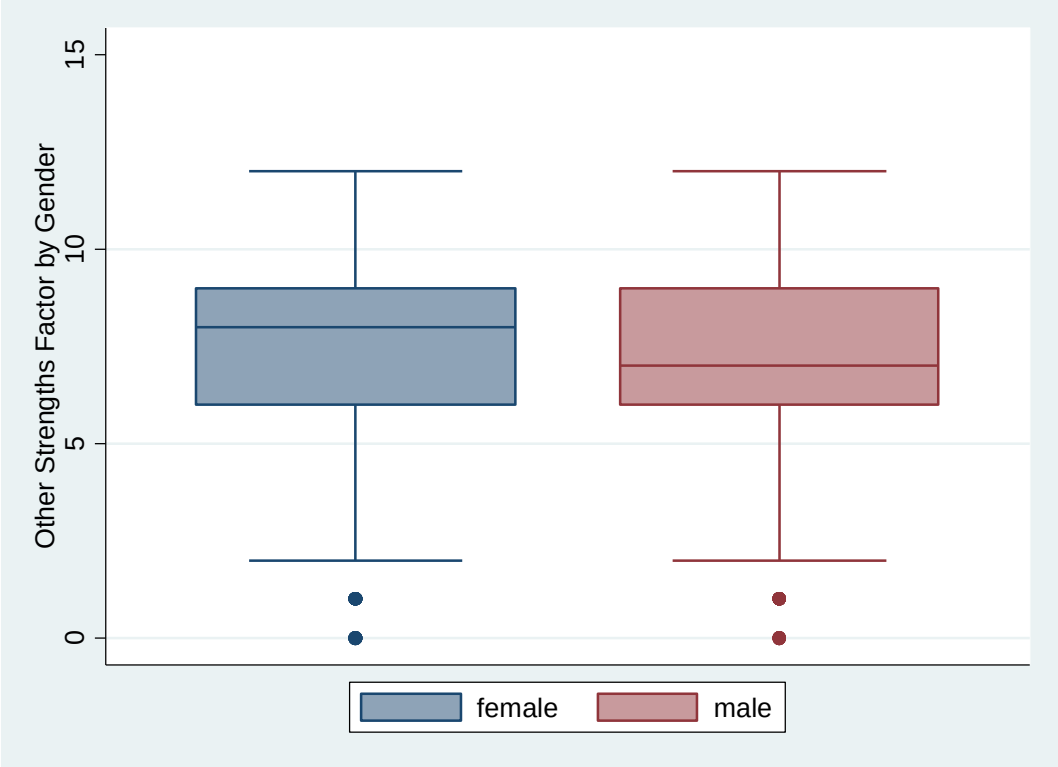

Unidimensional Strength Composite Score by Ethnicity
0
5
10
15
20
white
african-american/black
asian
latino/a
multi-ethnic
pacificislander/nativehawaiian
other

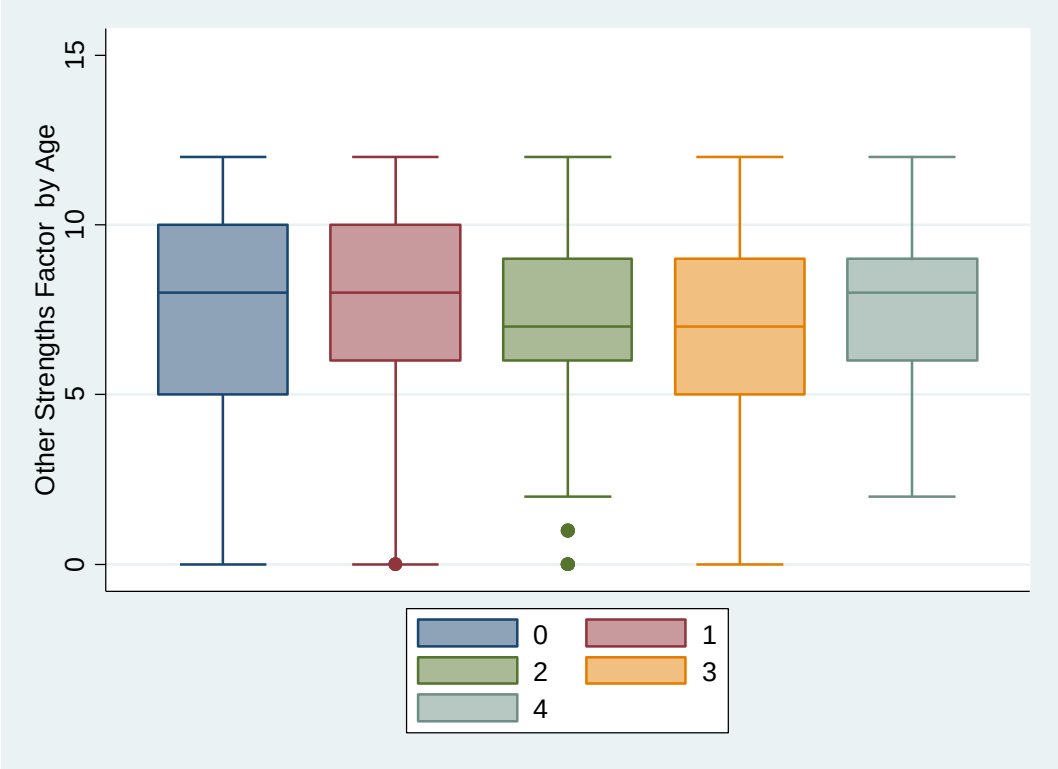
Box plots for Factor Scores:

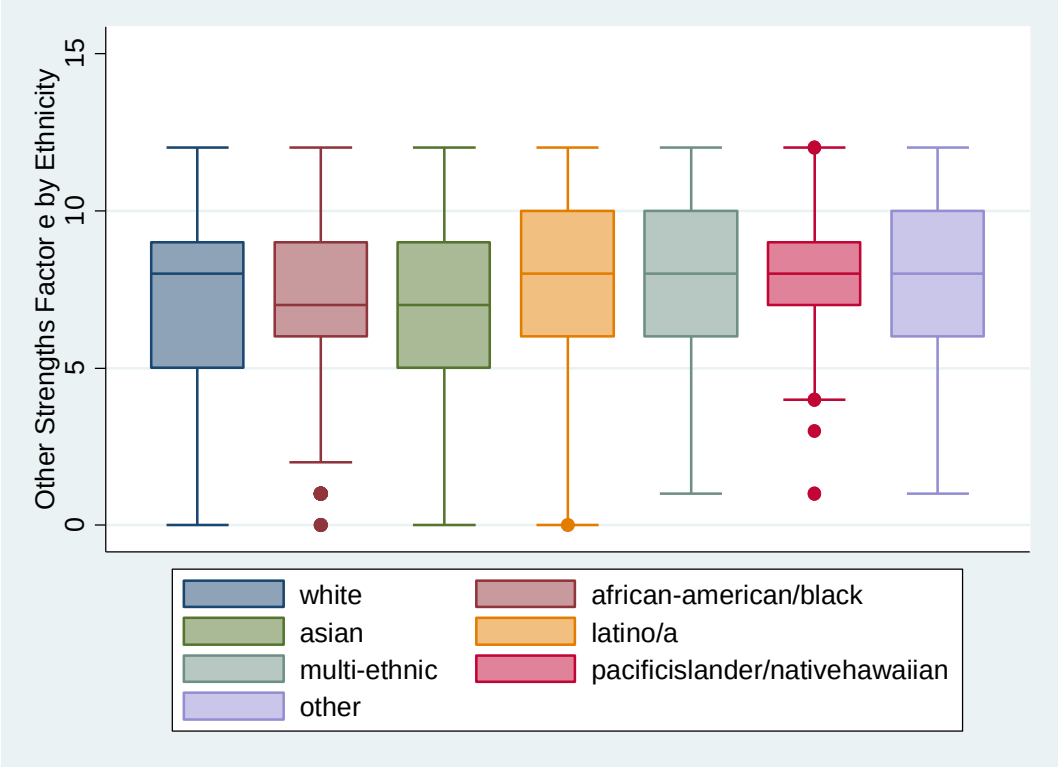





Family and Relationships Factor by Ethnicity
0
2
4
6
white
african-american/black
asian
latino/a
multi-ethnic
pacificislander/nativehawaiian
other


Other Strengths Factor by Gender
0
5
10
15
female
male


Other Strengths Factor by Age
0
5
10
15
0
1
2
3
4


Other Strengths Factor e by Ethnicity
0
5
10
15
white
african-american/black
asian
latino/a
multi-ethnic
pacificislander/nativehawaiian
other

Appendix C

Appendix C

Correlation Table for all Strengths Modeling Options

	Strength Composite	Family Relationship Item	Peer and Non-Caregiver Item	Education System and Environment Item	Extracurricular Talents and Interests Item	Spiritual and Religious Item	Relationship Permanence Item	Family and Relationship Factor	Other Strengths Factor
Strength Composite	1								
Family Relationship Item	0.467	1							
Peer and Non-Caregiver Item	0.6144	0.1797	1						
Education System and Environment Item	0.5937	0.1372	0.3984	1					
Extracurriculr Talents and Interests Item	0.6671	0.107	0.3313	0.2931	1				
Spiritual and Religious Item	0.6364	0.0374	0.2051	0.1827	0.4062	1			
Relationship Permanence Item	0.5419	0.3606	0.1422	0.1409	0.1789	0.1549	1		
Family and Relationship Strengths Factor	0.6138	0.8005	0.1936	0.1686	0.1758	0.1208	0.8477	1	
Other Strengths Factor	0.9106	0.1562	0.6555	0.6431	0.7298	0.7208	0.2238	0.2326	1

Appendix D

The Stata13 (Statacorps, 2013) command codes for the analyses in paper one are included here:

```
*Paper 1 Do File */

*Assumption Checking for Research Questions

*for the individual items (#1)
summarize famnew1, detail //to check skew and kurtosis score
graph box famnew1, over(male) ytitle(Family Strength by Gender) asyvars
graph box famnew1, over(agedum) ytitle(Family Strength by Age) asyvars
graph box famnew1, over(racefinal) ytitle(Family Strength by Ethnicity) asyvars

summarize interpnew1, detail //to check skew and kurtosis score
graph box interpnew1, over(male) ytitle(Peer and Non-Caregiver Strength by Gender)
asyvars
graph box interpnew1, over(agedum) ytitle(Peer and Non-Caregiver Strength by Gender by
Age) asyvars
graph box interpnew1, over(racefinal) ytitle(Peer and Non-Caregiver Strength by Gender
by Ethnicity) asyvars

summarize educnew1, detail //to check skew and kurtosis score
graph box educnew1, over(male) ytitle(Education System Strength by Gender) asyvars
graph box educnew1, over(agedum) ytitle(Education System Strength Score by Age)
asyvars
graph box educnew1, over(racefinal) ytitle(Education System Strength by Ethnicity)
asyvars

summarize talnew1, detail //to check skew and kurtosis score
graph box talnew1, over(male) ytitle(Talents and Interests Strength by Gender) asyvars
graph box talnew1, over(agedum) ytitle(Talents and Interests Strength by Age) asyvars
graph box talnew1, over(racefinal) ytitle(Talents and Interests Strength by Ethnicity)
asyvars

summarize spirnew1, detail //to check skew and kurtosis score
graph box spirnew1, over(male) ytitle(Spiritual and Religious Strength by Gender)
asyvars
graph box spirnew1, over(agedum) ytitle(Spiritual and Religious Strength by Age)
asyvars
graph box spirnew1, over(racefinal) ytitle(Spiritual and Religious Strength by
Ethnicity) asyvars

summarize relnew1, detail //to check skew and kurtosis score
graph box relnew1, over(male) ytitle(Relationship Permanence Strength by Gender)
asyvars
graph box relnew1, over(agedum) ytitle(Relationship Permanence Strength by Age)
asyvars
graph box relnew1, over(racefinal) ytitle(URelationship Permanence Strength by
Ethnicity) asyvars

*for the sub-scale factors (#2)
summarize famrelpfactor, detail //to check skew and kurtosis score
graph box famrelpfactor, over(male) ytitle(Family and Relationships Factor by Gender)
asyvars
graph box famrelpfactor, over(agedum) ytitle(Family and Relationships Factor by Age)
asyvars
graph box famrelpfactor, over(racefinal) ytitle(Family and Relationships Factor by
Ethnicity) asyvars

summarize otherfactor, detail //to check skew and kurtosis score
```

```
graph box otherfactor, over(male) ytitle(Other Strengths Factor by Gender) asyvars
graph box otherfactor, over(agedum) ytitle(Other Strengths Factor  by Age) asyvars
graph box otherfactor, over(racefinal) ytitle(Other Strengths Factor e by Ethnicity)
asyvars


*for Overall Composite Score (#3)
summarize strength, detail //to check skew and kurtosis score
graph box strength, over(male) ytitle(Unidimensional Strength Composite Score by
Gender) asyvars
graph box strength, over(agedum) ytitle(Unidimensional Strength Composite Score by
Age) asyvars
graph box strength, over(racefinal) ytitle(Unidimensional Strength Composite Score by
Ethnicity) asyvars

/*RESEARCH QUESTION ONE*/

*Table 2 Prevalence Statistics:
*prevalence data for the individual items
tabstat famnew1, statistics(mean sd n sk k r)
tabstat interpnew1, statistics(mean sd n sk k r)
tabstat educnew1, statistics(mean sd n sk k r)
tabstat talnew1, statistics(mean sd n sk k r)
tabstat spirnew1, statistics(mean sd n sk k r)
tabstat relnew1, statistics(mean sd n sk k r)

*create composite scores of sub-scales for prevalence data
generate famrelpfactor = famnew1 + relnew1
label variable famrelpfactor "famnew1 + relnew1"
generate otherfactor = interpnew1 + educnew1 + talnew1 + spirnew1
label variable otherfactor "interpnew1 + educnew1 + talnew1 + spirnew1 "

*prevalence data for the subscales (#2)
tabstat famrelpfactor, statistics(mean sd n sk k r)
tabstat otherfactor, statistics(mean sd n sk k r)

*prevalence data for Overall Strength Composite (#3)
tabstat strength, statistics(mean sd n r sk k)

/* RESEARCH QUESTION TWO */

*Table 3 CFA code
*one factor model
sem (STRENGTHS ->famnew1 interpnew1 educnew1 talnew1 spirnew1 relnew1), stand
estimates store m1 /* for lrtest comparing this as the nested model*/
estat gof, stats(all)

*two factor model
sem (STRENGTH1 ->interpnew1 educnew1 talnew1 spirnew1)(STRENGTH2 ->famnew1 relnew1),
stand
estimates store m2 /*for lrtest comparing this as the full model*/
estat gof, stats(all)

*post-estimation using predict command to get factor scores
predict factorscore1 factorscore2, latent

lrtest m1 m2 /* to compare the two models */

/* RESEARCH QUESTION THREE */

*Table 4 Gender and ttests/effect sizes
*gender and individual items (#1)
ttest famnew1, by(male)
```

```
ttest interpnew1, by(male)
ttest educnew1, by(male)
ttest talnew1, by(male)
ttest spirnew1, by(male)
ttest relnew1, by(male)

*gender and factor scores (#2) (using variables from predict command above)
ttest factorscorestrength1, by(male)
ttest factorscorestrength2, by(male)

*gender and overall composite score (#3)
ttest strength, by(male)

*Table 5.1 & 5.2: Age and Anovas/Effect Sizes/bonferroni tests
*age and individual items (#1)
tabstat famnew1, by(agedum) statistics(mean sd n sk k r)
tabstat interpnew1, by(agedum) statistics(mean sd n sk k r)
tabstat educnew1, by(agedum) statistics(mean sd n sk k r)
tabstat talnew1, by(agedum) statistics(mean sd n sk k r)
tabstat spirnew1, by(agedum) statistics(mean sd n sk k r)
tabstat relnew1, by(agedum) statistics(mean sd n sk k r)

oneway famnew1 agedum, b
esizei 4 2044 5.29
oneway interpnew1 agedum, b
esizei 4 2044 1.12 /* not necessary cause anova didn't show sig pvalue but coded it
anyways */
oneway educnew1 agedum, b
esizei 4 2044 8.23
oneway talnew1 agedum, b
esizei 4 2044 6.65
oneway spirnew1 agedum, b
esizei 4. 2044 1.19
oneway relnew1 agedum, b
esizei 4 2044 1.85

*age and factor scores for subscales(#2)
tabstat factorscorestrength1, by(agedum) stat(mean sd range sk k min max)
tabstat factorscorestrength2, by(agedum) stat(mean sd range sk k min max)

oneway factorscorestrength1 agedum, b /* oneway anova */
esizei 4 2044 2.92 /*for table4.1 effect size */

oneway factorscorestrength2 agedum, b /* oneway anova*/
esizei 4 2044 4.41

*age and overall composite score (#3)
tabstat strength, by(agedum) statistics(mean sd n r sk k)
tabstat strength, by(agedum) statistics(mean sd n r sk k)
tabstat strength, by(agedum) statistics(mean sd n r sk k)

oneway strength agedum, b
esizei 4 2044 3.75

*Table 6.1 & 6.2: Ethnicity and Anovas/Effect Sizes/bonferroni tests
*ethnicity and individual items (#1)
tabstat famnew1, by(racefinal) statistics(mean sd n sk k r)
tabstat interpnew1, by(racefinal) statistics(mean sd n sk k r)
tabstat educnew1, by(racefinal) statistics(mean sd n sk k r)
tabstat talnew1, by(racefinal) statistics(mean sd n sk k r)
tabstat spirnew1, by(racefinal) statistics(mean sd n sk k r)
tabstat relnew1, by(racefinal) statistics(mean sd n sk k r)
```

```
oneway famnew1 racefinal, b
esizei 6 2042 4.12
oneway interpnew1 racefinal, b
esizei 6 2042 5.02
oneway educnew1 racefinal, b
esizei 6 2042 2.86
oneway talnew1 racefinal, b
esizei 6 2042 3.74
oneway spirnew1 racefinal, b
esizei 6 2042 11.29
oneway relnew1 racefinal, b
esizei 6 2042 14.73

*ethnicity and factor scores for sub-scales (#2)
tabstat factorscorestrength1, by(racefinal) stat(mean sd range sk k min max)
tabstat factorscorestrength2, by(racefinal) stat(mean sd range sk k min max)

oneway factorscorestrength1 racefinal, b /* oneway anova */
esizei 6 2042 4.71

oneway factorscorestrength2 racefinal, b /* oneway anova*/
esizei 6 2042 8.30

*ethnicity and overall composite score (#3)
tabstat strength, by(racefinal) statistics(mean sd range sk k min max)

oneway strength racefinal, b
esizei 6 2042 4.12
```

Appendix E

$$
\begin{aligned}
Psych_{ijk} = {} & \beta_1 + \beta_2 Famnew1_{ijk} + \beta_3 Interpnew1_{ijk} + \beta_4 Educnew1_{ijk} \\
& + \beta_5 Talnew1_{ijk} + \beta_6 Spirnew1_{ijk} + \beta_7 Relnew_{ijk} \\
& + \beta_8 Trauma_{ijk} + B_9 Male_{ijk} + \beta_{10} Age_{ijk} + \beta_{11} Black_{ijk} \\
& + \beta_{12} Asian_{ijk} + \beta_{13} Latino_{ijk} + \beta_{14} Multi_{ijk} + \beta_{15} Pacisl_{ijk} \\
& + \beta_{16} Other_{ijk} + \zeta_{jk}^{(2)} + \zeta_k^{(3)} \\
& + \epsilon_{ijk}, \quad \zeta_{jk}^{(2)} \sim N\left(0, \psi^{(2)}\right), \quad \zeta_k^{(3)} \sim N\left(0, \psi^{(3)}\right), \\
& \epsilon_{ijk} | \zeta_j \sim N(0, \theta)
\end{aligned}
$$

Psychiatric symptom severity for a particular youth (i) for a particular clinician (j) in a particular agency (k) is modeled as a function of an average youth-specific intercept (β_1) the child-level covariates (β_2 ***to*** β_{12}) and three error components ($\zeta_{jk}(2)$, $\zeta_k(3)$, and ϵ_{ijk}). β_1 is the mean intercept representing the mean psychiatric severity score for female, white youth, at age 12.5, across clinicians and agencies with no trauma history (trauma = 0) and scores of 0 for all strengths. β_2 to β_7 are the coefficients representing the mean change in psychiatric symptom severity for every one point increase in the youth strengths items, adjusting for the other covariates. β_8 is the mean change in psychiatric severity score for every one point increase in lifetime trauma exposure, adjusting for the other covariates. β_9 is the difference in mean psychiatric severity score between males and females, adjusting for the other covariates. β_{10} is the mean change in psychiatric severity score for every year increase in youth age. β_{11} to β_{16} are the coefficients representing the difference in mean psychiatric severity score between white youth and the each of the other ethnicities, adjusting for the other covariates.

The superscripts of the random intercept residual error components $\zeta_{jk}^{(2)}$, $\zeta_k^{(3)}$, and ϵ_{ijk} denote the levels at which the random intercepts vary. These error components are assumed to have a mean of zero and to be mutually uncorrelated so that their variances add up to the total variance (Rabe-Hesketh & Skrondal, 2012, pg. 388). $\zeta_k^{(3)}$ is the random intercept for agency *k*, or how the estimated mean psych score at agency k deviates from the grand mean. $\zeta_k^{(3)}$has a variance $\psi^{(3)}$, interpreted as the level-three variance or the between-agency variance. $\zeta_{jk}^{(2)}$ is the random intercept for clinician *j* in agency *k*, or how clinician j in agency k deviates from the estimated mean psych score of agency k. $\zeta_{jk}^{(2)}$ has a variance $\psi^{(2)}$, interpreted as the level two variance, or the between-clinicians, within-agency variance. ϵ_{ijk} is the random intercept for youth *i* in clinician *j* in agency *k*, or how each youth i within clinician j within agency k deviates the clinician mean. ϵ_{ijk} has a variance component θ, interpreted as the level one residual, or the between youth, within clinician, and within agency variance.

Appendix F

The Stata13 (Statacorps, 2013) command codes for the analyses in paper two are included here:

```
/* Paper Two Do-File*/

*Descriptives
 /*to run descriptives, I attained the final sample by dropping all youth ///
 who had any missing data on psych famnew1 interpnew1 educnew1 talnew1 ///
 spirnew1 relnew1 trauma male age black asian latino multi pacisl other */

save paper1_missingremovedfordescriptives /

 *dropped 353 youth who were missing data for a final n of 2049 */
 drop if (psych>=.) | (famnew1>=.) | (interpnew1>=.) | (educnew1>=.) ///
 | (talnew1>=.) | (spirnew1>=.) | (relnew1>=.)| (trauma>=.) | (male>=.) | ///
 (age>=.) | (black>=.) | (white>=.) | (asian>=.) | (latino>=.) | (multi>=.) | ///
 (pacisl>=.) | (other>=.) | (clinician>=.) | (agency>=.) | (youth>=.)

 drop if (youth>=.)

*generate the mean-centered age variable age_cent
egen grandmean_age=mean(age)
gen age_cent = age-grandmean_age

 /*4.1.a */
 tab male
 tab racefinal
 summarize age

 tab famnew1
 tab interpnew1
 tab educnew1
 tab talnew1
 tab spirnew1
 tab relnew1

 summarize psych
 summarize trauma

 xtsum psych trauma famnew1 interpnew1 educnew1 talnew1 spirnew1 ///
 relnew1, i(clinician)

 xtsum psych trauma famnew1 interpnew1 educnew1 talnew1 spirnew1 ///
 relnew1, i(agency)


/* 4.1.b (re-ran all this on the ///
paper1_missingremovedfordescriptives dataset */
sort clinician
 egen clinicianid = group(clinician) //generate an ID for clinicians
 //starting at 1.
 list clinician psych clinicianid in 1/10
 sort agency
 egen agencyid = group(agency)
 list agency psych agencyid in 1/10

 egen tag_agency = tag(agency) //dummy variable taking value one once per
 *agency and zero otherwise.
 tab tag_agency, missing //none are missing
```

```
 egen tag_clinician = tag(clinician) // dummy variable taking value once
 *per clinician and zero otherwise.
 tab tag_clinician, missing //none are missing

 /*Sample Sizes: How many clusters are in the data? How many clinicians in each agency?
 How many youth from each clinician?) Generate count variables: */
 bysort agency: gen agencysize = _N //number of youth from each agency
bysort clinician: gen cliniciansize = _N // number of youth from each clinician
egen agencyclasssize = total(tag_clinician), by(agency) //number of clincians
 *from each agency

 *Summaries (reported in the table in 4.1.a):
 summ agencysize if tag_agency
 summ cliniciansize if tag_clinician
 summ agencyclasssize if tag_agency
 display _N

*For Covariates at different levels:
*Youth Level 1
summ psych
*Clinician Level 2
summ psych if tag_clinician
*Agency level 3
summ psych if tag_agency

/*4.1.c Residual Diagnostics and Predictions */
*run full model
xtmixed psych famnew1 interpnew1 educnew1 talnew1 spirnew1 relnew1 ///
 trauma male age black asian latino multi pacisl other || agency: || clinician:, mle

 predict zeta_agency zeta_clinician, reffects //level two and level three resids

 histogram zeta_agency if tag_agency, normal xtitle("Agency-level Empirical Bayes
Predictions for Residuals")
 histogram zeta_clinician if tag_clinician, normal xtitle("Clinician-level Empirical
Bayes Predictions for Residuals")

predict resid, residuals //level one residuals

replace zeta_agency=. if tag_agency!=1 // i am just graphing one from each agency
replace zeta_clinician=. if tag_clinician!=1 // i am just graphing one clinician from
each agency
graph box zeta_agency zeta_clinician resid, ascategory box(1, bstyle(outline)) ///
yvaroptions(relabel(1 "Agency" 2 "Clinician" 3 "Youth")) ///
medline(lcolor(black))
title("Box Plots of Empirical Bayes Predictions For Random Intercepts at the Agency,
Clinician, and Youth Level")

/* Multi-level Models */

*Model 1
xtmixed psych || agency: || clinician:, mle vce(robust)

*Model 2
xtmixed psych famnew1 interpnew1 educnew1 talnew1 ///
spirnew1 relnew1 || agency: || clinician:, mle vce(robust)

*Model 3
xtmixed psych famnew1 interpnew1 educnew1 talnew1 ///
spirnew1 relnew1 trauma || agency: || clinician:, mle vce(robust)

*Model 4
xtmixed psych famnew1 interpnew1 educnew1 talnew1 spirnew1 relnew1 ///
```

```
 trauma male age_cent black asian latino multi pacisl other ///
 || agency: || clinician:, mle vce(robust)

/* Likelihood ratio tests to test if between-methods within-subjects
variance is significantly different from zero? */

*Model 1
xtmixed psych || agency: || clinician:, mle
estimates store threelev1
xtmixed psych || agency:, mle
estimates store twolev1
lrtest threelev1 twolev1

*Model 2
xtmixed psych famnew1 interpnew1 educnew1 talnew1 spirnew1 ///
relnew1 || agency: || clinician:, mle
estimates store threelev2
xtmixed psych famnew1 interpnew1 educnew1 talnew1 spirnew1 ///
relnew1 || agency:, mle
estimates store twolev2
lrtest threelev2 twolev2

*Model 3
xtmixed psych famnew1 interpnew1 educnew1 talnew1 spirnew1 relnew1 ///
 trauma || agency: || clinician:, mle
 estimates store threelev3
 xtmixed psych famnew1 interpnew1 educnew1 talnew1 spirnew1 relnew1 ///
 trauma || agency:, mle
 estimates store twolev3
 lrtest threelev3 twolev3



*Model 4
xtmixed psych famnew1 interpnew1 educnew1 talnew1 spirnew1 relnew1 ///
 trauma male age black asian latino multi pacisl other///
 || agency: || clinician:, mle
estimates store threelev4
xtmixed psych famnew1 interpnew1 educnew1 talnew1 spirnew1 relnew1 ///
 trauma male age black asian latino multi pacisl other || agency:, mle
estimates store twolev4
lrtest threelev4 twolev4

/*Likelihood Ratio Tests Comparing the Models */
xtmixed psych || agency: || clinician:, mle
estimates store model1

xtmixed psych famnew1 interpnew1 educnew1 talnew1 ///
spirnew1 relnew1 || agency: || clinician:, mle
estimates store model2

xtmixed psych famnew1 interpnew1 educnew1 talnew1 spirnew1 relnew1 ///
 trauma || agency: || clinician:, mle
 estimates store model3

 xtmixed psych famnew1 interpnew1 educnew1 talnew1 spirnew1 relnew1 ///
 trauma male age black asian latino multi pacisl ///
 other || agency: || clinician:, mle
 estimates store model4

 lrtest model2 model1
 lrtest model3 model2
 lrtest model4 model3
```

```
 /* Calculated ICC3 and ICC2 */
 *ICC3
 display (1.65)^2 / ((1.65)^2 + (1.23)^2 + (2.48)^2)

*ICC2
display ((1.65)^2 + (1.23)^2) / ((1.65)^2 + (1.23)^2 + (2.48)^2)
```

Appendix G

The variables for the analyses in paper three are described below:

Dependent Variable.

Psychiatric Symptom Severity (Psych). The "Psychiatric Symptoms" domain is designed to track psychiatric symptom severity in the last thirty days. This domain consists of 9 items: psychosis, impulse/hyper, depression, anxiety, oppositional, conduct, substance use, somatization, and anger control. Each item receives a score on a four point likert scale (0 = no evidence of any needs, 1 = a dimension that requires monitoring watchful waiting, or preventive activities, 2= a dimension that requires action to ensure that this identified need is addressed/child meets criteria for a DSM diagnosis in this category, 3 = a dimension that requires immediate or intensive action). This variable was created by adding up the 0-3 score for the nine items in the domain. Psych is a continuous variable ranging from 0-27. Psych was measured at each time point (Assessment 1, 6, and 12) and
I s a time-varying variable.

Level One Time Varying Covariates (assessment points = i).

Time is a categorical variable measured in three dummy variables: Initial Assessment (Assess1dum = 1, other=0), Six Month Assessment (Assess6dum=1, other=0) or Twelve Month Assessment (Assess12dum=1, other=0). The data is considered "balanced," meaning that every youth who received a six month assessment received one during the same time frame, and similarly for the twelve month assessment, and thus why dummy variables are used in this analysis.

Strength. The six items listed below are part of the Youth Strengths domain. The Youth Strengths domain is designed to track the presence of accessible youth strengths in the past thirty days that could be used as the basis of a strengths-based plan. Each item is rated on a 0-3 likert scale and were reverse coded to assist in interpretation (with reveres coding: 3 = strengths exist that can be used as a centerpiece for a strength-based plan, 2 = strengths exist but require some strength building efforts in order for them to serve as a focus of a strength-based plan, 1 = strengths have been identified but they require significant strength building efforts before they can be effectively utilized in as a focus of a strength-based plan, 0 = efforts are needed in order to identify potential strengths for strength building efforts). Each strength was measured at every time point (1 = initial assessment, 6 = six month assessment, and 12 = twelve month assessment) and is a time-varying variable.

Family/Caregiver Relationship Strength (famnew).
Peer and Non-Family/Caregiver Interpersonal Relationships Strength (interpnew).
School System and Educational Plan Strength (educnew).

Extracurricular Activities and Talents Strength (talnew).
Spiritual/Religious Beliefs and/or Involvement Strength (spirnew).
Relationship Permanence Strength (relnew).

Level Two Non-Time-Varying Covariates (youth = j).

Youthfinal. Youthfinal is a unique identification number consecutively ordered from 1-n. Youthfinal is the level two variable, and so youth will be considered nested in clinicians.

Strength_init. The following six items are part of the Youth Strengths domain. The Youth Strengths domain is designed to track the presence of accessible youth strengths in the past thirty days that could be used as the basis of a strengths-based plan. Each item is rated on a 0-3 likert scale (0 = strengths exist that can be used as a centerpiece for a strength-based plan, 1 = strengths exist but require some strength building efforts in order for them to serve as a focus of a strength-based plan, 2 = strengths have been identified but they require significant strength building efforts before they can be effectively utilized in as a focus of a strength-based plan, 3 = efforts are needed in order to identify potential strengths for strength building efforts). These variables are included to control for initial strengths scores in one of the models.

Initial Strength Composite (strength_init)
Initial Family/Caregiver Relationship Strength (famnew_init).
Initial Peer and Non-Family/Caregiver Interpersonal Relationships Strength (interpnew_init).
Initial School System and Educational Plan Strength (educnew_init).
Initial Extracurricular Activities and Talents Strength (talnew_init).
Initial Spiritual/Religious Beliefs and/or Involvement Strength (spirnew_init).
Initial Relationship Permanence Strength (relnew_init).

Mean Strengths. These variables were constructed by adding together the 0-3 scores at each time point for each of the strengths variables and dividing by three (for example: [strength1 + strength6 + strength12 / 3]) for each youth, effectively creating an average strengths score over the 12 months of treatment. This allows for isolating the within-youth variation and the between-youth variation and makes for a more meaningful interpretation of the strength (famnew, interpnew, educnew, talnew, spirnew, relnew) coefficients. Mean-centering the covariates effectively eliminates the intercepts for the clinicians (similar to creating dummy variables for all the clinicians) so that only the within-clinician variation is explained by the covariates (Rabe-Hesketh & Skrondal, 2012, pg. 146). Each clinician serves as his/her own control here. By including these mean-centered strength terms, the within effects are isolated so that the other strength items in the model can be interpreted as the between-clinician effects. (Interpretation of the mean centered strength variables and the resulting interpretation of the other strength variables in model 4 is based on Rabe-Hesketh

& Skrondal, 2012, pg. 143). (For the other variables that were not mean centered in this model, the interpretation can be either a comparison of youth between different clinicians or between youth with the same clinician.)

Mean Strength Composite (mn_strength)
Mean Family/Caregiver Relationship Strength (mn_famnew).
Mean Peer and Non-Family/Caregiver Interpersonal Relationships Strength (mn_interpnew).
Mean School System and Educational Plan Strength (mn_educnew).
Mean Extracurricular Activities and Talents Strength (mn_talnew).
Mean Spiritual/Religious Beliefs and/or Involvement Strength (mn_spirnew).
Mean Relationship Permanence Strength (mn_relnew).

Lifetime Trauma Experience Severity at Initial Assessment (Trauma_init). This variable is designed to track lifetime exposure to trauma at initial assessment. This continuous composite variable was created by adding together the score for each of the twelve trauma types items on the CANS initial assessment. This domain includes twelve items (physical abuse, sexual abuse, emotional abuse, neglect, medical trauma, witness family violence, witness community violence, school violence, natural/man-made disaster, traumatic grief/separation, war affected, terrorism affected) rated on a 0-3 likert scale, with 0 = a dimension where there is no evidence of any trauma of this type, 1 = a dimension where a single event trauma occurred or suspicion exists of trauma experiences, 2 = a dimension on which the child has experienced multiple traumas, and 3 = a dimension which describes repeated and severe trauma with medical and physical consequences.

Male is a categorical variable for gender, coded male = 1 and female = 0.

Race is a categorical variable with seven categories: African-American/Black, Asian, Latino/a, Multi-ethnic, Native Hawaiian or Other Pacific Islander, White, or Other. White is used as the reference category and dummy variables were constructed for the others, including African-American/Black (AfAm = 1, all others = 0), Latino (Latino=1, all others = 0), Asian (Asian=1, all others = 0), Multi-Ethic (Multi-Ethnic = 1, all others = 0), HPI (Hawaii / Other Pacific Islander=1, all others = 0), and Other (Other=1, all others = 0).

Age. Age at initial assessment is a continuous variable ranging from 5-22.

Age_cent. Age_cent was created by mean centering the Age variable above. Age_cent helps with interpretation of the intercept (now interpreted as the outcome for a youth of average age, adjusting for other covariates).

Level Three Covariates (clinician = j).

Clinician. Each clinician has a clinician identification number (CIN) assigned to him or her. Clinician is the level three variable, so time is nested in youth is nested in clinician.

Appendix H

The models for questions 1-4 are listed below. Detailed descriptions of the interpretations of each coefficient are included in Appendix.

Model for questions 1-3:

$$
\begin{aligned}
Psych_{ijk} = {} & \beta_0 + \beta_1 Assess6dum_i + \beta_2 Assess12dum_i + \beta_3 famnew_init_{jk} \\
& + B_4 interpnew_init_{jk} + B_5 educnew_init_{jk} + \beta_6 talnew_init_{jk} \\
& + \beta_7 spirnew_{init_{jk}} + \beta_8 relnew_init_{jk} + B_9 trauma_init_{jk} + B_{10} Male_{jk} \\
& + \beta_{11} Age_cent_{jk} + \beta_{12} Black_{ijk} + \beta_{13} Asian_{ijk} + \beta_{14} Latino_{ijk} \\
& + \beta_{15} Multi_{ijk} + \beta_{16} Pacisl_{ijk} + \beta_{17} Other_{ijk} \\
& + \beta_{18} famnewinitXAssess6dum_i + \beta_{19} famnewinitXAssess12dum_i \\
& + \beta_{20} interpnewinitXAssess6dum + _i \beta_{21} interpnewinitXAssess12dum_i \\
& + \beta_{22} educnewinitXAssess6dum_i + \beta_{23} educnewinitXAssess12dum_i \\
& + \beta_{24} traumainitXAssess6dum_i + \beta_{25} traumainitXAssess12dum_i \\
& + \beta_{26} maleinitXAssess6dum_i + \beta_{27} maleinitXAssess12dum_i \\
& + \beta_{28} agecentXAssess6dum_i + \beta_{29} agecentXAssess12dum_i \\
& + \beta_{30} asianXAssess6dum_i + \beta_{31} asianXAssess12dum_i + \zeta_{jk}^{(2)} \\
& + \zeta_k^{(3)} + \epsilon_{ijk}, \quad \zeta_{jk}^{(2)} \sim N\left(0, \psi^{(2)}\right), \quad \zeta_k^{(3)} \sim N\left(0, \psi^{(3)}\right), \\
& \epsilon_{ijk} | \zeta_j \sim N(0, \theta)
\end{aligned}
$$

Psychiatric symptom severity at a particular time point (i) for a particular youth (j) for a particular clinician (k) is modeled as a function of an average intercept (β_0), the time-varying covariates (with coefficients $\beta_1 - \beta_2$), the non-time varying youth-level covariates (with coefficients $\beta_3 - \beta_{17}$), fourteen cross-level interaction terms (with coefficients$\beta_{18} - \beta_{31}$), and three error components ($\zeta_{jk}^{(2)}$, $\zeta_k^{(3)}$, ϵ_{ijk}). β_0 is the mean intercept representing the mean psychiatric severity score at initial assessment for female, white youth, aged 12.5, with no trauma history (trauma_init = 0), who entered the SOC with no initial strengths (famnew,_init = interpnew_init = educnew_init = talnew_init = sprinew_init = relnew_init = 0) across clinicians and agencies. β_1 *and* β_2 are the coefficents representing the mean difference in psychiatric symptom severity between, respectively, initial assessment and six-month assessment, and initial assessment and twelve month assessment, adjusting for the other covariates. $\beta_3 - \beta_8$ are the coefficients representing the difference in mean psychiatric severity score at the initial assessment between youth differing in their strength item by one unit, adjusting for the other covariates. β_9 is the difference in mean psychiatric severity score at the initial assessment for youth differing in the lifetime trauma exposure by one point, adjusting for the other covariates. β_{10} is the difference in mean psychiatric severity score at the initial assessment between males and females, adjusting for the other covariates. β_{11} is the difference in mean psychiatric score at the initial assessment between youth differing in age by one year, adjusting for the other covariates. $\beta_{12} - \beta_{17}$ is the difference in mean psychiatric severity score between white youth and each other ethnicity at a given time-point, adjusting for the other beta covariates. β_{18} is the difference in the mean increase (from initial assessment to six months) in psychiatric severity score between youth differing by one unit on the initial family strength score, adjusting for other covariates. β_{19} is the difference in the mean increase (from initial assessment to 12 months) in

psychiatric severity score for youth differing in their initial family strength score by one unit, adjusting for other covariates. β_{20} and β_{21} are the mean difference in psychiatric severity score between, respectively, 6 month and initial assessment, and 12 month and initial assessment, for youth with a given initial interpersonal strength score, for a particular clinician, adjusting for other covariates. β_{22} and β_{23} are the mean difference in psychiatric severity score between, respectively, 6 month and initial assessment, and 12 month and initial assessment, for youth with a given initial education strength score, for a particular clinician, adjusting for other covariates. β_{24} and β_{25} are the mean difference in psychiatric severity score between, respectively, 6 month and initial assessment, and 12 month and initial assessment, for youth with a given initial trauma score, for a particular clinician, adjusting for other covariates. β_{26} and β_{27} are the mean difference in psychiatric severity score between, respectively, 6 month and initial assessment, and 12 month and initial assessment, for males (compared to females), for a particular clinician, adjusting for other covariates. β_{28} and β_{29} are the mean difference in psychiatric severity score between, respectively, 6 month and initial assessment, and 12 month and initial assessment, for every year increase in age in youth, for a particular clinician, adjusting for other covariates. β_{30} and β_{31} are the mean difference in psychiatric severity score between, respectively, 6 month and initial assessment, and 12 month and initial assessment, for Asian youth (compared to White youth), for a particular clinician, adjusting for other covariates.

The superscripts of the random intercept residual error components $\zeta_{jk}^{(2)}$, $\zeta_{k}^{(3)}$ and ϵ_{ijk} denote the levels at which the random intercepts vary. These error components are assumed to have a mean of zero and to be mutually uncorrelated so that their variances add up to the total variance (Rabe-Hesketh & Skrondal, 2012, pg. 388). $\zeta_{k}^{(3)}$ is the random intercept for clinician *k,* or how the estimated mean psych score for clinician k deviates from the grand mean. $\zeta_{k}^{(3)}$ has a variance $\psi^{(3)}$, interpreted as the level-three variance or the between-clinician variance. $\zeta_{jk}^{(2)}$ is the random intercept for youth *j* in clinician *k*, or how youth j in clinician k deviates from the estimated mean psych score of the clinician k. $\zeta_{jk}^{(2)}$ has a variance $\psi^{(2)}$, interpreted as the level two variance, or the between-youth, within- clinician variance. ϵ_{ijk} is the random error for assessment *i* in youth *j* in clinician *k*, or how each assessment i within youth j within clinician k deviates around the youth mean. ϵ_{ijk} has a variance component θ, interpreted as the level one variance, or the between- assessment, within-youth, and within-clinician variance.

Three models were used to answer research questions 1-3. First, a model (model 1) with no covariates was constructed to determine the percentage of variance that is a) between clinician, b) within clinician (different youth and assessment) and c) within clinician, within youth (different assessment). Next, to answer research question 1, the covariates $\beta_{1\ to}$ β_{17} were added (model 2) to assess the contribution to the variation in mean psychiatric severity score over twelve months by these covariates. Interaction terms were constructed for any of the significant covariates from model 2. Then, to answer research question 2, β_{18} to β_{31} (the interaction terms) were added to the model (model 3) to see if mean psychiatric severity varied over six months or twelve months for a given level of the covariate, adjusting for the other covariates in the model. Because it appeared that the residuals of the dependent variable, psychiatric symptom severity, was not perfectly normally distributed (see histograms and boxplots below), robust standard errors were calculated for all the models using the sandwich

estimator method (as long as the sample is not small, these robust standard errors are considered reliable even if the normality assumption might be somewhat violated). A likelihood ratio test was conducted to compare the fit of model 1 vs 2 and 2 vs. 3. The likelihood ratio test was evaluated at the .05 significance level and tests the null hypothesis that the restricted model fits the data as well as the full model against the alternative hypothesis that the full model fits the data better than the restricted model. Then, the estimated intraclass correlations were calculated to further answer research question three.

Model for Research Question Four:

$$
\begin{aligned}
Psych_{ijk} = \ & \beta_0 \\
& + \beta_1 famnew_i + B_2 interpnew_i + B_3 educnew_i + \beta_4 talnew_i \\
& + \beta_5 spirnew_i + \beta_6 relnew_i + \beta_7 Assess6dum_i + \beta_8 Assess12dum_i \\
& + \beta_9 mnfamnew_{jk} + B_{10} mninterpnew_{jk} + B_{11} mneducnew_{jk} \\
& + \beta_{12} mntalnew_{jk} + \beta_{13} mnspirnew_{jk} + \beta_{14} mnrelnew_{jk} \\
& + B_{15} trauma_init_{jk} + B_{16} Male_{jk} + \beta_{17} Age_cent_{jk} + \beta_{18} Black_{ijk} \\
& + \beta_{19} Asian_{ijk} + \beta_{20} Latino_{ijk} + \beta_{21} Multi_{ijk} + \beta_{22} Pacisl_{ijk} \\
& + \beta_{23} Other_{ijk} + \zeta_{jk}^{(2)} + \zeta_k^{(3)} \\
& + \epsilon_{ijk}, \quad \zeta_{jk}^{(2)} \sim N\left(0, \psi^{(2)}\right), \quad \zeta_k^{(3)} \sim N\left(0, \psi^{(3)}\right), \qquad \epsilon_{ijk} | \zeta_j \sim N(0, \theta)
\end{aligned}
$$

Psychiatric symptom severity at a particular time point (i) for a particular youth (j) for a particular clinician (k) is modeled as a function of an average intercept (β_0), the time-varying covariates (with coefficients $\beta_1 - \beta_9$), the non-time varying youth-level covariates (with coefficients $\beta_{10} - \beta_{23}$), and three error components ($\zeta_{jk}(2)$, $\zeta_k(3)$, ϵ_{ijk}). β_0 is the mean intercept representing the mean psychiatric severity score at initial assessment for female, white youth, aged 12.5, with no trauma history (trauma_init = 0), who entered the SOC with no initial strengths (famnew_init = interpnew_init = educnew_init = talnew_init = sprinew_init = relnew_init = 0) across clinicians. β_1 *to* β_6 are the coefficients representing the mean difference in psychiatric symptom severity at a given time between two youth (who share similar covariates) who are working with different clinicians, one of whom is a point higher = on the strength item than the other one (for example, at six months, one youth may have increased on the strength item from 0 to 1 and the other one may have stayed at 0). β_7 *and* β_8 are the coefficients representing the mean change in psychiatric severity score between six months and initial assessment, or 12 months and initial assessment, adjusting for the covariates. β_9 *to* β_{14} are the coefficients representing the mean difference in psychiatric symptom severity between two youth, differing in their youth strength item by one point, given the level-2 covariates. β_{15} is the mean increase in psychiatric symptom severity at a given time for every point increase in lifetime trauma exposure, adjusting for the other covariates. B_{16} is the difference in mean psychiatric severity score at a given time between males and females, adjusting for the other covariates. B_{17} is the mean difference in psychiatric score at a given time for every one point increase in age, adjusting for the other covariates. $B_{18} - B_{23}$ is the mean difference in psychiatric severity score at a given time between white youth and each of the other ethnic groups, adjusting for the other covariates. $\zeta_k(3)$ is the random intercept for clinician *k*, or how the estimated mean psych score for clinician k varies from the grand mean. $\zeta_k(3)$ has a variance component $\psi^{(3)}$, interpreted as the level-three residual or the between-clinician

variance. $\zeta_{jk}^{(2)}$ is the random intercept for youth *j* in clinician *k*, or how youth j in clinician k varies from the estimated mean psych score of the clinician k. $\zeta_{jk}^{(2)}$ has a variance component $\psi^{(2)}$, interpreted as the level two residual, or the between youth, within clinician variance. ϵ_{ijk} is the random intercept for assessment *i* in youth *j* in clinician *k*, or how each assessment i within youth j within clinician k varies around the youth mean. ϵ_{ijk} has a variance component θ, interpreted as the level one residual, or the between assessment, within youth, and within clinician variance.

To answer research question four, model 4 was constructed. $\beta_1 \ to \ \beta_6$ are time-varying youth strength covariates, interpreted as the between-clinician effects since $\beta_9 \ to \ \beta_{14}$ (mean-centered strength covariates, representing the within-clinician effects) were included to isolate the within-clinician effects. With the inclusion of $\beta_9 \ to \ \beta_{14}$, $\beta_1 \ to \ \beta_6$ answer the question of whether change in youth strength covariates over twelve months of intervention predict change in youth psychiatric symptoms over twelve months of intervention, for youth with similar characteristics working with different clinicians. Again, for model 4, robust standard errors were calculated since the residual distribution of the response variable is not perfectly normal.

In addition to dealing with clustering in data (as discussed in paper two), multi-level modeling is useful for handling missing data and for dealing with repeated measures. First, MLM provides more flexibility in handling missing data and unbalanced designs by using information from all of the data, even if the data is not all complete (in part due to the use of maximum likelihood estimators, as discussed in paper two). Additionally, multi-level modeling has several advantages in dealing with longitudinal data when compared to ANOVA repeated measures designs (Kwok, et al., 2008; Rabe-Hesketh & Skrondal, 2012, pg. 244). If fixed effects approaches are used, each individual is able to serve as his/her own control over time, thus accounting for infinite unmeasured confounders (though this is not the case with random effects approaches). Ordinary Least Squares (OLS) approaches treat every time point as repeated cross-sectional data (where the subjects are drawn independently at each time point), and this does not differentiate between within and between-subject comparisons (and within subject comparisons are ideal because subjects truly act as their own controls). OLS approaches are also incorrect when there is missing data. MLM allows for special models that can differentiate between within and between subjects and model within subject residual covariance in more effective ways (Rabe-Hesketh & Skrondal, 2012, pg. 244).

Appendix I

The Stata13 (Statacorps, 2013) command codes for the analyses in paper one are included here:

```
/*PAPER 3 DO FILE*/

/*Data Cleaning*/
*get all data composites created now for each timepoint
generate psych12 = (psychosis_12 + impulse_hyper_12 + depression_12 +  ///
anxiety_12 + oppositional_12 + conduct_12+ substance_use_12_A + ///
somatization_12 + anger_control_12)
label variable psych12 "Psychiatric Symptom Severity TwelveMonthAssess"
tab psych12, missing


generate psych6 = (psychosis_6 + impulse_hyper_6 + depression_6 +  ///
anxiety_6 + oppositional_6 + conduct_6+ substance_use_6_A + ///
somatization_6 + anger_control_6)
label variable psych6 "Psychiatric Symptom Severity SixMonthAssess"
tab psych6, missing

rename riskbeh riskbeh1

generate riskbeh6 = (suicide_risk_6 + other_self_harm_6 + danger_to_others_6 +
sexual_aggression_6 + ///
  runaway_6 + delinquency_6 + fire_setting_6 + social_behavior_6 )
  label variable riskbeh6 "Risk Behaviors Severity Sixmonths"


generate riskbeh12 = (suicide_risk_12 + other_self_harm_12 + danger_to_others_12 +
sexual_aggression_12 + ///
  runaway_12 + delinquency_12 + fire_setting_12 + social_behavior_12 )
  label variable riskbeh12 "Risk Behaviors Severity Twelve Months"


*reverse code 6moStrengths items and make a 6mstrengths composite
generate famnew6 = 3 - family_strength_6
list family_strength_6 famnew6 in 1/10
tab famnew6, missing
tab family_strength_6, missing

generate interpnew6 = 3 - interpersonal_6
list interpersonal_6 interpnew6 in 1/10
tab interpnew6, missing

generate educnew6 = 3 - educational_6
list educational_6 educnew6 in 1/10
tab educnew6, missing

generate talnew6 = 3 - talents_interests_6
list talents_interests_6 talnew6 in 1/10
tab talnew6, missing

generate spirnew6 = 3 - spiritual_religious_6
list spiritual_religious_6 spirnew6 in  1/10
tab spirnew6, missing

generate relnew6 = 3 - relationship_perm_6
list relationship_perm_6 relnew6 in 1/10
tab relnew6, missing

generate strength6 = (famnew6 + interpnew6 + educnew6 + talnew6 + spirnew6 + relnew6)
```

```
list famnew6 interpnew6 educnew6 talnew6 spirnew6 relnew6 strength6 in 1/10
label variable strength6 "Child Strengths Initial Composite, Reverse Coded"
tab strength6, missing

*reversecode 12mostrengthns and make a composite
generate famnew12 = 3 - family_strength_12
list family_strength_12 famnew12 in 1/10
tab famnew12, missing
tab family_strength_12, missing

generate interpnew12 = 3 - interpersonal_12
list interpersonal_12 interpnew12 in 1/10
tab interpnew12, missing

generate educnew12 = 3 - educational_12
list educational_12 educnew12 in 1/10
tab educnew12, missing

generate talnew12 = 3 - talents_interests_12
list talents_interests_12 talnew12 in 1/10
tab talnew12, missing

generate spirnew12 = 3 - spiritual_religious_12
list spiritual_religious_12 spirnew12 in 1/10
tab spirnew12, missing

generate relnew12 = 3 - relationship_perm_12
list relationship_perm_12 relnew12 in 1/10
tab relnew12, missing

generate strength12 = (famnew12 + interpnew12 + educnew12 + talnew12 + spirnew12 +
relnew12)
list famnew12 interpnew12 educnew12 talnew12 spirnew12 relnew12 strength12 in 1/10
label variable strength12 "Child Strengths 12month Composite, Reverse Coded"
tab strength12, missing


*going between long and wide data
reshape long psych psychosis_ impulse_hyper_ depression_ anxiety_ ///
oppositional_ conduct_ substance_use_youth_ somatization_ anger_control_ ///
riskbeh suicide_risk_  other_self_harm_ danger_to_others_  ///
sexual_aggression_  runaway_  delinquency_  fire_setting_  ///
social_behavior_ strength famnew interpnew educnew talnew ///
spirnew relnew, i(youthfinal) j(assessment) /*when i reshape to long
youthfinal moves automatically to first variable in the list */

*generate the mean-centered age variable age_cent
egen grandmean_age=mean(age)
gen age_cent = age-grandmean_age

*generate assessment dummy variables
tabulate assessment, generate(r) //(Rabe-Hesketh & Skrondal,2012, pg 143)
rename r1 assess1dum
rename r2 assess6dum
rename r3 assess12dum

drop assess1dum /*drop these before reshape wide because they are only for long
form data, add back in when using long data */
drop assess6dum
drop assess12dum
```

```
reshape wide psych psychosis_ impulse_hyper_ depression_ anxiety_ ///
oppositional_ conduct_ substance_use_youth_ somatization_ anger_control_ ///
riskbeh suicide_risk_  other_self_harm_ danger_to_others_  ///
sexual_aggression_  runaway_  delinquency_  fire_setting_  ///
social_behavior_ strength famnew interpnew educnew talnew ///
spirnew relnew, i(youthfinal) j(assessment) /*had to go back to wide to create
any non-time varying variables like strength_init, famnew_init interpnew_init
educnew_init talnew1_init spirnew_init relnew_init and mean_strength below*/

*generate strength_init
generate strength_init = strength1
tab strength_init strength1, missing /*both are missing in 14 cases */

generate famnew_init = famnew1
generate interpnew_init = interpnew1
generate educnew_init = educnew1
generate talnew_init = talnew1
generate spirnew_init = spirnew1
generate relnew_init = relnew1

*generate interaction terms
generate famnew_initXAssess6dum = famnew_init*assess6dum
generate famnew_initXAssess12dum = famnew_init*assess12dum
generate educnew_initXassess6dum = educnew_init*assess6dum
generate educnew_initXAssess12dum = educnew_init*assess12dum
generate interpnew_initXassess6dum = interpnew_init*assess6dum
generate interpnew_initXassess12dum= interpnew_init*assess12dum
generate trauma_initXassess6dum = trauma_init*assess6dum
generate trauma_initXAssess12dum = trauma_init*assess12dum
generate strength_initXassess6dum = strength_init*assess6dum
generate strength_initXassess12dum = strength_init*assess12dum
generate maleXassess6dum = male*assess6dum
generate maleXassess12dum = male*assess12dum
generate age_centXassess6dum = age_cent*assess6dum
generate age_centXassess12dum = age_cent*assess12dum
generate asianXassess6dum = asian*assess6dum
generate asianXassess12dum = asian*assess12dum

*generate mean variables to isolate the between effects
egen mns = mean(strength), by(youthfinal) /*this gives us purely
 the between clinician effects, (Rabe-Hesketh & Skrondal, 2012, pg 1444)  */
tab mns, missing /*missing 39 values */

egen mn_famnew = mean(famnew), by(youthfinal) /* missing 39 */
egen mn_interpnew = mean(interpnew), by(youthfinal) /* missing 39 */
egen mn_educnew = mean(educnew), by(youthfinal) /* missing 39 */
egen mn_talnew = mean(talnew), by(youthfinal) /* missing 39 */
egen mn_spirnew = mean(spirnew), by(youthfinal) /* missing 39 */
egen mn_relnew = mean(relnew), by(youthfinal) /* missing 39 */


*Check missing data
xtset youthfinal assessment
xtdescribe if psych<. /*shows patter of missing data*/

egen numobs2 = count(psych), by(youthfinal) /*count the number of observations
per youth where the response variable is not missing */
generate comp1 = numobs2==3 /* creates dummy var comp1 for youth having complete
data*/
generate any =  numobs2>0 /*creates dummy variable for youth with any data */
egen pick_youth = tag(youth) /*creates dummy variable taking the value one exactly
once per youth and zero otherwise */
```

```
egen numcomp3 = total(comp1*pick_youth), by(clinician) /*gives me number of youth with
c
with all three assessments by clinician) */
egen numany3 = total(any*pick_youth), by(clinician) /* gives me number of youth with
any assessments by clinician */
generate rate = numcomp3/numany3 /*calculates proportion of children with complete
data in each school */
egen pick_clinician = tag(clinician) /*dummy variable to pick out one observation
per school */
summarize numany3 numcomp3 rate if pick_clinician==1 /*summarizes the clinician
level variables */


/*Descriptives*/
*drop youth who were missing data
generate insample = e(sample) // use directly after an estimation,
* 1 = obs used in estimation
keep if insample // now data only keeps people with no missing data, but you have
*to reload the database  to get the data back...so careful...
tabulate insample /*This paper has 2059 youth */

tab male if pick_youth==1, missing
tab racefinal if pick_youth==1
summarize age if pick_youth==1

*to summarize level 1 variables
*the next two commands summarize the variability of the time-varying variables
*within and between youth, ignoring clinicians
quietly xtset youthfinal
xtsum psych famnew interpnew educnew talnew spirnew relnew

*to summarize the level-2 variables
egen mn_psych = mean(psych), by(youthfinal)/*create the mean psych score, use xtsum
which will tell us how much it varies between clinicians*/
quietly xtset clinician
xtsum mn_psych mn_famnew mn_interpnew mn_educnew mn_talnew mn_spirnew mn_relnew ///
trauma_init male age_cent white black latino asian pacisl multi other ///
 if pick_youth==1 /*this finds means and within and between school SDs
of the child-level variables */

sort clinician
egen clinicianid = group(clinician) //generate ID for clinician starting at 1
list clinician psych clinicianid in 1/10

sort youth
egen youthid = group(youth)
list youth psych youthid in 1/10

egen tag_clinician = tag(clinician) // dummy variable taking value one once per
clinician and 0 otherwise
tab tag_clinician, missing // none are missing
egen tag_youth = tag(youth) // dummy var taking one once per youth and zero otherwise
tab tag_youth, missing

/* sample sizes: how many clusters are in the data?
how many youth in each clinician? how many youth at each timepint? */
bysort clinician: gen cliniciansize = _N // number of youth
bysort youth: gen youthsize = _N
egen clinicianyouthsize = total(tag_youth), by(clinician)

*summaries
summ cliniciansize if tag_clinician //number of clinicians
summ youthsize if tag_youth // number of youth
```

```
summ clinicianyouthsize if tag_clinician
display _N

*Residual Diagnostics and Predictions
*run full model
xtmixed psych assess6dum assess12dum famnew_init interpnew_init ///
educnew_init talnew_init spirnew_init relnew_init trauma_init ///
male age_cent black latino asian pacisl multi other ///
famnew_initXAssess6dum famnew_initXAssess12dum ///
educnew_initXassess6dum educnew_initXAssess12dum ///
interpnew_initXassess6dum interpnew_initXassess12dum ///
trauma_initXassess6dum trauma_initXAssess12dum maleXassess6dum ///
maleXassess12dum age_centXassess6dum age_centXassess12dum asianXassess6dum ///
asianXassess12dum|| clinician:, || youth:, mle vce(robust)

predict zeta_clinician zeta_youth, reffects // level two and level three resids
histogram zeta_clinician if tag_clinician, normal xtitle("Clinician-level Empirical
Bayes Predictions for Residuals")
histogram zeta_youth if tag_youth, normal xtitle("Youth-level Empirical Bayes
Predictions for Residuals")

predict resid, residuals // level one residuals

replace zeta_clinician=. if tag_clinician!=1 // i am just graphing one from each
clinician
replace zeta_youth=. if tag_youth!=1 // i am just graphing one youth from each
clinician
graph box zeta_clinician zeta_youth resid, ascategory box(1, bstyle(outline)) ///
yvaroptions(relabel(1 "Clinician" 2 "Youth" 3 "Assessment")) ///
medline(lcolor(black))
title("Box Plots of Empirical Bayes Predictions For Random Intercepts at the
Clinician, Youth, and Assessment Level")

/*Models*/
*Model 1
xtmixed psych || clinician:, || youth:, mle vce(robust)

*Model 2
xtmixed psych assess6dum assess12dum famnew_init interpnew_init ///
educnew_init talnew_init spirnew_init relnew_init trauma_init ///
male age_cent black latino asian pacisl multi other || clinician:, ///
|| youth:, mle vce(robust)

*Model 3
xtmixed psych assess6dum assess12dum famnew_init interpnew_init ///
educnew_init talnew_init spirnew_init relnew_init trauma_init ///
male age_cent black latino asian pacisl multi other ///
famnew_initXAssess6dum famnew_initXAssess12dum ///
educnew_initXassess6dum educnew_initXAssess12dum ///
interpnew_initXassess6dum interpnew_initXassess12dum ///
trauma_initXassess6dum trauma_initXAssess12dum maleXassess6dum ///
maleXassess12dum age_centXassess6dum age_centXassess12dum asianXassess6dum ///
asianXassess12dum|| clinician:, || youth:, mle vce(robust)

*Model 4
xtmixed psych famnew interpnew educnew talnew spirnew relnew ///
 assess6dum assess12dum mn_famnew mn_interpnew mn_educnew mn_talnew ///
 mn_spirnew mn_relnew trauma_init male age_cent black latino asian pacisl ///
 multi other || clinician:, || youth:, mle vce(robust) /*i think including the
 mn variables also estimates within clinician effects */

/*Likelihood Ratio Tests to test if between-methods within-subjects variance is
is significantly different from zero (I could also run lincom mn_psych - psych
```

```
see lab 8)*/
*Model 1
xtmixed psych || clinician:, || youth:, mle
estimates store threelev1
xtmixed psych || clinician:, mle
estimates store twolev1
lrtest threelev1 twolev1

*Model 2
quietly xtmixed psych assess6dum assess12dum famnew_init interpnew_init ///
educnew_init talnew_init spirnew_init relnew_init trauma_init ///
male age_cent black latino asian pacisl multi other || clinician:, ///
|| youth:, mle
estimates store threelev2
quietly xtmixed psych assess6dum assess12dum famnew_init interpnew_init ///
educnew_init talnew_init spirnew_init relnew_init trauma_init ///
male age_cent black latino asian pacisl multi other || clinician:, mle
estimates store twolev2
lrtest threelev2 twolev2

*Model 3
quietly xtmixed psych assess6dum assess12dum famnew_init interpnew_init ///
educnew_init talnew_init spirnew_init relnew_init trauma_init ///
male age_cent black latino asian pacisl multi other ///
famnew_initXAssess6dum famnew_initXAssess12dum ///
educnew_initXassess6dum educnew_initXAssess12dum ///
interpnew_initXassess6dum interpnew_initXassess12dum ///
trauma_initXassess6dum trauma_initXAssess12dum maleXassess6dum ///
maleXassess12dum age_centXassess6dum age_centXassess12dum asianXassess6dum ///
asianXassess12dum|| clinician:, || youth:, mle
estimates store threelev3
quietly xtmixed psych assess6dum assess12dum famnew_init interpnew_init ///
educnew_init talnew_init spirnew_init relnew_init trauma_init ///
male age_cent black latino asian pacisl multi other ///
famnew_initXAssess6dum famnew_initXAssess12dum ///
educnew_initXassess6dum educnew_initXAssess12dum ///
interpnew_initXassess6dum interpnew_initXassess12dum ///
trauma_initXassess6dum trauma_initXAssess12dum maleXassess6dum ///
maleXassess12dum age_centXassess6dum age_centXassess12dum asianXassess6dum ///
asianXassess12dum|| clinician:, mle
estimates store twolev3
lrtest threelev3 twolev3

*Model 4
quietly xtmixed psych famnew interpnew educnew talnew spirnew relnew ///
 assess6dum assess12dum mn_famnew mn_interpnew mn_educnew mn_talnew ///
 mn_spirnew mn_relnew trauma_init male age_cent black latino asian pacisl ///
 multi other || clinician:, || youth:, mle /*i think including the
 mn variables also estimates within clinician effects */
 estimates store threelev4
 quietly xtmixed psych famnew interpnew educnew talnew spirnew relnew ///
 assess6dum assess12dum mn_famnew mn_interpnew mn_educnew mn_talnew ///
 mn_spirnew mn_relnew trauma_init male age_cent black latino asian pacisl ///
 multi other || clinician:, mle  /*i think including the
 mn variables also estimates within clinician effects */
 estimates store twolev4
 lrtest threelev4 twolev4

 /*Likelihood ratio tests comparing the Models */
 /*Models*/
*Model 1
quietly xtmixed psych || clinician:, || youth:, mle
estimates store model1
```

```
*Model 2
quietly xtmixed psych assess6dum assess12dum famnew_init interpnew_init ///
educnew_init talnew_init spirnew_init relnew_init trauma_init ///
male age_cent black latino asian pacisl multi other || clinician:, ///
|| youth:, mle
estimates store model2

*Model 3
quietly xtmixed psych assess6dum assess12dum famnew_init interpnew_init ///
educnew_init talnew_init spirnew_init relnew_init trauma_init ///
male age_cent black latino asian pacisl multi other ///
famnew_initXAssess6dum famnew_initXAssess12dum ///
educnew_initXassess6dum educnew_initXAssess12dum ///
interpnew_initXassess6dum interpnew_initXassess12dum ///
trauma_initXassess6dum trauma_initXAssess12dum maleXassess6dum ///
maleXassess12dum age_centXassess6dum age_centXassess12dum asianXassess6dum ///
asianXassess12dum|| clinician:, || youth:, mle
estimates store model3

*Model 4
quietly xtmixed psych famnew interpnew educnew talnew spirnew relnew ///
 assess6dum assess12dum mn_famnew mn_interpnew mn_educnew mn_talnew ///
 mn_spirnew mn_relnew trauma_init male age_cent black latino asian pacisl ///
 multi other || clinician:, || youth:, mle  /*i think including the
 mn variables also estimates within clinician effects */
 estimates store model4

 lrtest model2 model1
 lrtest model3 model1
 lrtest model3 model2
 lrtest model4 model1

 /*Calculate ICC3 and ICC2 based on Model 1 (variance component model)*/
 *ICC3
 display (1.70)^2 / ((1.70)^2 + (1.75)^2 + (1.91)^2)

*ICC2
 display ((1.70)^2 + (1.75)^2) / ((1.70)^2 + (1.75)^2 + (1.91)^2)
```